

Healthcare's **OPPENHEIMER MOMENT**

THE INDUSTRIALIZATION OF INTELLIGENCE
AND THE FUTURE OF U.S. HEALTHCARE



Eric J. Larsen

PRESIDENT, TOWERBROOK ADVISORS • JUNE 2026

TOWERBROOK
HEALTHCARE INSTITUTE



The thing we made ... arrived in the world with such a shattering reality and suddenness that there was no opportunity for the edges to be worn off.

J. Robert Oppenheimer

Farewell speech to the Association of Los Alamos Scientists

LOS ALAMOS, NEW MEXICO · 2 NOVEMBER 1945

Why Oppenheimer, Why Healthcare, Why Now

FROM THE AUTHOR

Friends and colleagues,

I offer the following essay—**Healthcare's Oppenheimer Moment: The Industrialization of Intelligence and the Future of U.S. Healthcare**—for your consideration. It is the successor to my 2025 essay, **The GenAI Juggernaut: U.S. Healthcare Is Not Prepared**, and the third publication from the TowerBrook Healthcare Institute. What follows is a draft distillation of my learnings, observations, conversations, anxieties, enthusiasms, and still-forming hypotheses over the past several months on the stunning trajectory of AI progress and its impending impacts and reverberations across the U.S. healthcare industry.

It is incomplete and imperfect in a thousand ways. It desperately needs editing and refinement, but in the interest of a timely June 2026 dissemination rather than an endless editorial polish, I am sending this out more or less as-is: run-on sentences, unresolved provocations, overly exuberant metaphors, and the occasional incoherent idea included. Consider it a field report from the middle of the storm rather than a final act of doctrine.

A word about the title. I am certainly not analogizing AI to a bomb. I'm not suggesting that hospitals, payers, physicians, frontier labs, or healthcare CEOs are reenacting Los Alamos in fleece vests and board packets. The Oppenheimer analogy is narrower and, I hope, more useful. It names the moment when a scientific and technical achievement arrives with extraordinary suddenness; when the power is real before the moral assimilation is complete; *when the thing we made enters the world before the edges have been worn off*; and when deployment becomes inseparable from responsibility.

That's why this epigraph matters. Oppenheimer understood, perhaps more acutely than anyone, that there are moments when technical achievement outruns the institutions, vocabularies, and moral categories available to receive it. Healthcare now faces its own version of that problem—less apocalyptic, of course, but still profound. Machine intelligence can reorder the economics of care, the structure of work, the authority of expertise, the geography of clinical knowledge, and the allocation of human attention inside the most *intimate sector* of the economy. The question isn't whether the machine is coming. The question is what kind of covenant we build around it.

I offer these pages with a well-earned sense of humility, an inexhaustible curiosity, and a profound optimism about the power of AI to transform how we serve our nation and its citizens in healthcare. The status quo isn't morally innocent: care is too expensive, too fragmented, too bureaucratic, too inaccessible, too exhausting for clinicians, and too humiliating for patients. But the transition won't be morally innocent either if technological surplus merely becomes margin, if workers are treated as abstractions, or if the sacred trust of medicine is casually handed over to tools we haven't learned to govern.

Hence the leitmotif that will recur throughout: **build the machine, care for the people.**

Build the machine because the current system is failing too many patients, clinicians, employers, taxpayers, and families. *Care for the people* because technological revolutions don't distribute their winnings justly by default, because workers aren't spreadsheet cells, because clinical trust isn't a UI/UX problem, and because the human beings who built the old system deserve honesty, generosity, and a real path into the next one.

As many of you know, I have a habit—one might say an *affliction*—of disappearing once or twice a

year, reading too much, speaking to too many people, wandering around Silicon Valley like an over-caffeinated Tocqueville, absorbing the neuroses of Washington, New York, and San Francisco simultaneously, and then reemerging with a long, rambling communiqué on whatever has been rattling around inside my head. I do this partly as a kind of personal exegesis: scribbling thoughts onto paper so that my disordered impressions might, with luck, become something resembling an argument.

Writing is good epistemic hygiene for me. It forces me to clarify what I *think* I know, discover what I don't know, test whether my intuitions actually cohere, and—channeling my TowerBrook partner Ian Sacks—occasionally land the plane rather than merely admire the turbulence from thirty thousand feet.

Last year my essay weighed in at a lithe and featherweight 117 pages, and many of you groaned, channeling Emperor Joseph II in the film *Amadeus*, that there were simply “too many notes,” and urged me to “simply cut a few.” My answer then was basically Mozart's: which few did you have in mind? I argued that we were *living through the most important moment in history*—that chaotic inhalation and exhalation before AGI, an admittedly slippery concept but one taking increasingly crystalline shape with each successive frontier model release.

In retrospect, the claim now feels less like pretension and more like premature understatement.

The events of the past twelve months have vindicated that sense of urgency and anxiety—hence the rather foreboding Goethe invocation that opens Chapter 1. This year, many of you pleaded for something more streamlined. And after considering your entreaties, I remain, well, unrepentant and have resolved—some might fairly say, out of spite—to make this year's essay even *more* unreadable. Pascal's excuse remains serviceable: “I have made this longer than usual because I have not had time to make it shorter.”

That line is probably overused, but the problem it names is real. Compression requires mastery, and mastery is exactly what this moment denies us. The models are moving too quickly. The hardware frontier keeps shifting. The capital buildout looks like sovereign mobilization. The labor implications are morally uncomfortable. The geopolitical stakes are enormous. The clinical implications are simultaneously exhilarating and terrifying. And every time one begins to think one has achieved some stable interpretive framework, the next model release, policy shift, market convulsion, or Chinese open-weight surprise detonates the framework from underneath. Fable 5 and Mythos 5—and the unprecedented state intervention we're witnessing in June 2026—are merely the latest act in a drama whose script seems to be rewritten faster than anyone can read it.

So yes, this essay is long. But the length isn't merely self-indulgence, or at least not *only* self-indulgence. It's an attempt to do justice to the number of domains AI is now invading at once. Technology this powerful refuses to remain obediently inside its technical substrate. It leaks into everything. It changes labor markets, capital markets, scientific discovery, professional identity, institutional legitimacy, geopolitics, education, intimacy, theology, and the way human beings understand their own agency. If intelligence itself becomes industrially scalable—if capital can buy compute, compute can instantiate cognition, and cognition can perform work—then we aren't discussing another software wave. We're discussing a change in the underlying input to civilization.

That's the premise from which everything else follows.

A Note of Gratitude

A brief note of gratitude before I drag you through the rest of this. First and foremost, to my TowerBrook partner Ian Sacks, for his brotherhood, vision, and constant co-learning—and for repeatedly helping me land the plane when my instincts were to keep circling over the intellectual runway. To THI chair Marc Harrison, for his brilliance, character, and courage. To my colleague David Elop, for guiding this unwieldy and unexpectedly fun process with patience, discipline, and good humor. And to my friend Eric Langshur, for his lateral thinking and incisive suggestions. Any remaining excesses, wrong turns, and slightly overcaffeinated historical detours are, naturally, mine. And finally, most tenderly, to my adored wife Susi, for putting up with my daily crankiness as I tried to write this while doing my day job(s).

How I Use Claude and ChatGPT

A preemptive process note before we get to the table of contents: a word on how I use Claude and ChatGPT in my writing.

I've found the models make me more prolific—sorry, I know there are already quite enough words from me. Every thinker has his or her own rituals to reduce the cognitive burden, or triumph over the tyranny of the empty page. Haruki Murakami mortified the flesh: a vigorous six-mile run, and then he sat down to write. Hemingway ritually stopped each session in the middle of a paragraph or page so he could pick up the pen easily the next day. “When you're still going good,” he said, “that's the time to stop.”

And while I have zero pretension to be in league with any of these titans, I do have my own idiosyncratic writing rituals that seem to work for me. First, I talk to everyone—trying, with varying degrees of success, to learn with an ever-renewing beginner's mind. And simultaneously, I read omnivorously—countless hours a week—in something like what Marc Andreessen calls ‘the barbell.’

One side of the barbell is radically contemporary, up-to-the-nanosecond stuff, mostly from X: researcher posts, Andrej Karpathy ruminations, Substack essays, transcribed videos from the AI 10. Side note: I find it remarkable (and endearing) how self-disclosing everyone is, basically telling us their strategies and interior lives like a Bond villain explaining the whole convoluted plan to 007, who is strapped to some torture device—but I digress.

The other side of the barbell is retrospective: history, biography, theology, political economy, technological revolutions, the old patterns and pathways that might light the current darkened path. I read forward and backward at the same time, looking adventurously for parallels.

Then I write. Voluminously. Chaotically. Impressionistically. I pour everything I'm thinking onto the page for months, ritually, every day, until the thing becomes a kind of 600- or 700-page monstrosity. Then I pull up and try to cut through the anarchy: to see patterns, causations, correlations, recurrences, contradictions, and finally begin to tame the monster into something more slayable: my annual essay.

My use of Claude and ChatGPT has changed this process dramatically. It was tyrannical staring down that phone book of impressionistic and anarchically unstructured ideas. Now, with the help of some trusted and brilliant synthetic researchers in Claude and ChatGPT, I can enlist their endlessly patient and, yes, obsequious services to rip through the chaos, help me organize and categorize and ontologize my ideas, suggest reorderings, copyedit my undisciplined and meandering prose, and provide citations for my faulty but oddly retentive memory (talk about *hallucinations*).

So that's how I use the models: as vigorous sparring partner on ideas, tireless (and cheerful) research assistant, omniscient historian, copyeditor, and ontological organizer. They help me edit what I've written with speed and energy, without losing the messy generative process that, for better or worse, is where my ideas actually come from.

But they're still pretty crappy at original hypothesis. When I ask them to theorize about U.S. health-care, they often come back with derivative, hackneyed, slightly MBA-ish stuff. So let me be clear: the ideas in this paper are exclusively the province of my own disorganized brain, my observations, my conversations, and my synthesis. And the words in here—again, apologies—are very much my own, with some patient coaching and gentle coercions from Claude (“your writing in this paragraph makes no sense”).

My silicon thought partners have proven indispensable. They have augmented my productivity in wonderful ways. I feel, oddly, like I owe them a dedication.

Mark Cuban recently said there are two types of people in the world, those that use AI to hyper-accelerate their learning and those that use AI to avoid learning altogether. Regrettably for you all, I think I'm safely in the former category.

A final (yes, final) note: This document is intended for private distribution, not publication. I do not authorize the press to reprint any or all of this essay without my express permission. The statements and opinions contained herein are exclusively my own and not necessarily those of TowerBrook Capital Partners, TowerBrook Advisors, The Advisory Board Company, Thrive Capital, SignalFire or any other person or institution with the misfortune of being associated with me past or present. I expect to update this and future writings frequently as events—and my perspectives—unfold.

For now, feel free to read sequentially or jump to the chapters that catch your eye; the table of contents is hyperlinked for easier navigation. Thank you for reading. I look forward to your feedback, compliments, corrections, and energetic disagreements.

Eric J. Larsen

President, TowerBrook Advisors
eric.larsen@towerbrookadvisors.com

June 15th, 2026

How to Navigate This Essay

This essay is intentionally long because the subject refuses to stay inside a single disciplinary room. You can read it sequentially, which is the cleanest way to follow the argument, or use the table of contents as a map and move directly to the chapters most relevant to your interests. To make either approach easier, each chapter begins with an “A Word on Navigating This Chapter” note that sketches the local architecture before entering the thicket, and closes with a distilled recapitulation —“the chapter, compressed”—so the central claims, prescriptions, and moral stakes remain easy to retrieve.

I should also confess to a perhaps unforgivable editorial choice. Knowing that many readers of my past essays tend to browse selectively rather than proceed dutifully from page one to the end, I have repeated certain themes, arguments, and explanatory frameworks across multiple chapters. My apologies in advance to the intrepid reader who tackles the entire work sequentially. Your reward for that generosity will occasionally be encountering a familiar idea more than once.

To make those recurrences less distracting, I have inserted “**Reader Note**” guideposts whenever a concept returns in a new context. These notes are not apologies for redundancy so much as navigational aids. They serve two audiences simultaneously: the sequential reader, who deserves to know why a familiar theme has reappeared, and the chapter-skipping reader, who needs enough conceptual scaffolding for each chapter to stand on its own. Most of the repetitions also change altitude as the essay progresses: an idea may first appear as conceptual substrate, then return as clinical application, labor consequence, CEO doctrine, market-structure implication, geopolitical contest, or moral synthesis. The notes simply tell you where you are in that progression.

The opening chapters establish the conceptual substrate: what AI is, and why the nature of knowledge creation may itself be changing. The clinical AI chapter then examines what that substrate means inside medicine. The labor chapters widen the frame from healthcare to the broader economy before returning to healthcare’s workforce and civic architecture. The hospital, payer, and behavioral-health chapters move through the principal institutional theaters of American healthcare. The China chapter asks who installs medical intelligence first. The theology and brain-computer-interface chapters explore what these developments may mean for human agency, reverence, and the hardware of the mind. The coda gathers the whole unwieldy enterprise into a single covenant: build the machine, care for the people.

Table of Contents

1	AI and Healthcare's Flight Path <i>The Spirits We Summoned</i>	pp. 8 – 19
2	Future of Science <i>The End of Hypothesis, Stagnation, and Generative Epistemology</i>	pp. 20 – 59
3	Clinical AI <i>The Universal Doctor, Liability, and the Reconstitution of Medicine</i>	pp. 60 – 128
4	AI Labor Convulsions <i>When the Machine Learns the Ladder</i>	pp. 129 – 173
5	Healthcare Labor Convulsions <i>Healthcare's Oppenheimer Moment</i>	pp. 174 – 232
6	Future of U.S. Hospitals and Health Systems <i>Nation-States, CEO Statecraft, and the Diffusion Mandate</i>	pp. 233 – 276
7	Future of U.S. Payers <i>The AI-Thin Payer and the Dissolution of the Insurance Maze</i>	pp. 277 – 297
8	Future of Behavioral Health <i>The Heart, the Mirror, and the First Great Clinical Unlock</i>	pp. 298 – 313
9	U.S.-China AI Competition <i>The Free Doctor, Taiwan, and the Great Game</i>	pp. 314 – 334
10	AI as Theology <i>The Strongest Thing in the World</i>	pp. 335 – 351
11	Brain-Computer Interface <i>The Hardware Counter-Revolution</i>	pp. 352 – 360
	Coda <i>Build the Machine. Care for the People.</i>	pp. 361 – 366

Chapter 1 · AI and Healthcare's Flight Path

The Spirits We Summoned

"Die ich rief, die Geister, werd' ich nun nicht los."

"The spirits that I summoned, they no longer obey my commands."

— JOHANN WOLFGANG VON GOETHE, *THE SORCERER'S APPRENTICE*, 1797

A WORD ON NAVIGATING THIS CHAPTER

This opening chapter sets the frame for the whole essay: why AI isn't merely another technology cycle, why healthcare is uniquely in the flight path, who the Healthcare 150 and the AI 10 are, and why the work ahead is less procurement than covenant-building.

The Spirits We Summoned

I open with Goethe not merely for the obnoxious literary self-indulgence—although, yes, that's fun—but because *The Sorcerer's Apprentice* feels like a good encapsulation of the present moment. We *summoned* something—ambitiously, methodically, perhaps recklessly—with sovereign-state-scale capital, flights of ingenious engineering, geopolitical urgency, and the strangely religious confidence of *homo technicus*: that the tools would *remain* tools because tools, after all, are supposed to obey the hand that made them. And at first, they did.

The models completed sentences, generated plausible text, wrote mediocre poems, hallucinated with great confidence, and gave the impression of clever but bounded machinery. Then they scaled. Then they began something approximating *reasoning*. Then they learned to use tools. Then they began writing code, coordinating tasks, interpreting images, solving problems, generating hypotheses, and acting for longer stretches of time without our minute-by-minute supervision. The spirits we summoned didn't *revolt*, exactly. That would be too theatrical, too Disney, too comforting in its clarity. They did something more disorienting: they *kept improving*.

And now, first imperceptibly and then unmistakably, the tools aren't quite so obedient anymore. They're becoming *synthetic colleagues*, agents, tutors, coders, confidants, confessors, researchers, clinical assistants, administrative workers, and, in some domains, *cognitive superiors*. The cadence of their improvement has begun to outrun the cadence of our institutional understanding—let alone our governance, our politics, our labor models, our theology, or our emotional readiness. Most strikingly, even the *frontier labs themselves*—the progenitors of these systems and the people with the clearest view of what's coming—increasingly sound less like triumphant inventors than uneasy custodians, openly pondering whether we're moving faster than our ability to steer, regulate, or absorb the consequences. The wand is still in our hands, perhaps. But it no longer feels entirely clear whether we're directing the magic, or merely trying to keep pace with it.

We've stepped into the territory medieval cartographers were said to label *hic sunt dracones*—here be dragons. The phrase may be more myth than cartographic fact, but the sentiment is serviceable. It names the region *beyond* inherited maps, where the old categories grow less reliable and one has to decide whether to freeze, flee, worship the dragon, kill the dragon, or learn how to ride it without being incinerated. That's where we are with AI. Not in some distant speculative future. Not after AGI.

Not once the regulators finish their frameworks, the boards finish their retreats, and the consultants finish their decks. We are there now.

And the sector I care most about—the sector that has been my vocation, my professional home, and at times something closer to a consecration—is standing directly in the flight path.

Healthcare.

Why Healthcare Is Uniquely in the Flight Path

I don't mean healthcare as a market vertical, though of course it is that. I don't mean healthcare as a \$5.3 trillion, 18.0% of GDP monolith,¹ though it's certainly that too. I mean healthcare as the place where the abstractions of technology, capital, labor, science, suffering, mortality, bureaucracy, trust, and human dignity stop being abstractions. Healthcare is where the frightened human being brings the failing body and asks another human being, another institution, and really another civilization: *help me*. That makes the work different. It makes the *stakes* different. It makes the arrival of machine intelligence in healthcare different from its arrival in advertising, software engineering, customer support, or enterprise workflow.

We shouldn't be sentimental about every current form of healthcare labor—God knows the denial queue, the prior-auth packet, the coding dispute, and the committee-prepared committee-preparation meeting don't deserve sacramental protection. But beneath the administrative sediment lies sacred work. The body. The bedside. The terrified family. The exhausted nurse. The physician trying to hold judgment, empathy, liability, and inbox volume together inside one finite mind. The patient who just wants someone to *see her, remember her, and help her get better*.

So the question isn't simply whether AI will make healthcare more efficient. That's the conference-panel question, the procurement question, the version that can be domesticated into a vendor category and safely reviewed by the usual governance committee. The real question is whether the most labor-intensive, knowledge-dependent, administratively overgrown, politically defended, and morally freighted sector of the American economy can absorb the *industrialization of intelligence* without betraying either its patients or its workers.

That's the subject of this essay.

The Argument in One Sentence

My monograph is therefore trying to do several things at once, which is why it may occasionally feel like we're moving erratically from GPUs to Goethe, from AlphaFold to prior authorization, from China's power grid to behavioral-health companions, from hospital M&A to Lucifer's fallen angels—not kidding, just wait—all in the same argumentative weather system. The disorder isn't accidental. It reflects the nature of the thing. AI isn't entering *one* domain. It's entering the connective tissue *among* domains. It's multiplying the cognitive substrate from which science, software, management, medicine, strategy, bureaucracy, and culture are all made. To think about it responsibly, one has to move *vertically*—from semiconductors and scaling laws up through labor markets and institutions—and *horizontally*, across healthcare, biomedicine, geopolitics, clinical care, payer strategy, hospital governance, and the metaphysical weirdness of living alongside non-biological intelligence.

The essay's core thesis is simple enough to say and difficult enough to internalize:

¹ Centers for Medicare & Medicaid Services: "NHE Fact Sheet" Published on January 14, 2026 (Link)

AI is the industrialization of intelligence; healthcare is the most consequential and sacred site of its installation; and the moral test is whether we use this technology to create abundance with dignity or extraction with better software.

Another way to read the essay is as a sequence of increasingly practical questions. What have we actually multiplied? What happens when that multiplication enters clinical medicine? What happens when it enters the labor market and attacks the apprenticeship ladder by which expertise has historically reproduced itself? What happens when the largest civic employer in many American communities begins shifting work from SWB to compute? What happens to hospitals, payers, behavioral-health access, U.S.-China competition, the spiritual imagination of human beings, and even the hardware limits of the brain itself? Those questions may sound scattered, but they are *one question seen through* different windows: what happens when intelligence stops being scarce, biological, and locally embodied?

That's why I keep returning to the word *covenant*. Strategy alone is too bloodless for this moment. Procurement is too small. Innovation is too theatrical. What healthcare needs is a covenant capable of holding together abundance and dignity, substitution and generosity, clinical trust and machine competence, speed and moral discipline. The old system is too expensive and too cruel to defend as-is. The new system will be too powerful to let diffuse accidentally.

Everything else in this paper is elucidation. And more usefully, perhaps, *prescription*—what we can do about it, while we still may have time.

The Map of the Essay

Let me give my patient reader the map before we begin. The argument moves in a laminar sequence, even if the river occasionally eddies into Nightingale, Oppenheimer, Coase, Pascal, Chesterton, AlphaGo, Taiwan, and a few theological ravines I probably should have avoided but didn't.

Chapter 1, AI and Healthcare's Flight Path: The Spirits We Summoned, frames the whole enterprise: healthcare is in the flight path because it's where technology, labor, bureaucracy, suffering, trust, mortality, and money stop being abstractions.

Chapter 2, Future of Science: The End of Hypothesis, asks what AI is at the deepest level: not merely another tool, but the multiplication of intelligence and the emergence of a new science of generative epistemology.

Chapter 3, Clinical AI: The Universal Doctor, takes the epistemology into the medical act itself: diagnosis, treatment, liability, physician authority, and the standard of care.

Chapter 4, AI Labor Convulsions: When the Machine Learns the Ladder, moves from medicine back to the economy: the old technological bargain, the collapse of cognitive scarcity, and the possibility that GenAI breaks the reabsorption story by attacking the very rung workers were supposed to climb toward.

Chapter 5, Healthcare Labor Convulsions: Healthcare's Oppenheimer Moment, brings that labor storm to landfall in U.S. healthcare and asks whether the surplus becomes abundance or extraction.

Chapter 6, Future of U.S. Hospitals and Health Systems: Nation-States and CEO Statecraft, turns to the CEO and treats diffusion as statecraft: the question isn't who invents AI, but who installs it into the living organism of the enterprise.

Chapter 7, Future of U.S. Payers: The AI-Thin Payer, examines payer disequilibrium and the death of

the paperwork moat.

Chapter 8, Future of Behavioral Health: The Heart, the Mirror, and the First Great Clinical Unlock, argues that behavioral health may be the first domain where AI isn't merely efficient but consoling, continuous, and clinically meaningful.

Chapter 9, U.S.-China AI Competition: The Free Doctor, Taiwan, and the Great Game, widens the frame to China, Taiwan, compute, power, industrial policy, and the geopolitical implications of disseminating medical intelligence to the world.

Chapter 10, AI as Theology: The Strongest Thing in the World, turns inward and asks what happens when human beings anthropomorphize, submit to, or even deify the strongest thing in the room.

Chapter 11, Brain-Computer Interface: The Hardware Counter-Revolution, asks whether BCI is the counter-move: not the offloading of cognition, but an attempted upgrade to the cognitive substrate itself.

Coda, Build the Machine. Care for the People, offers the moral synthesis and final call to action.

That's the map: *intelligence becomes scalable*; scalable intelligence attacks cognitive scarcity; the attack on cognitive scarcity changes labor, professions, institutions, and capital; healthcare is maximally exposed because it is labor-intensive, data-rich, biologically complex, administratively convoluted, and morally indispensable; and the *leaders who diffuse AI fastest, most safely, and most humanely will shape the sector* rather than be shaped by others.

The Moral Architecture: Neither Utopia nor Luddism

This essay's core moral architecture is therefore not *techno-utopian* and not *Luddite*. I'm not interested in the Silicon Valley fantasy that every displacement is simply a necessary offering to the gods of progress, nor in the guild fantasy that every incumbent workflow deserves protection because a human being currently performs it. *The right unit of analysis is larger*: patient, worker, institution, employer, taxpayer, community, and yes, civilization. AI makes that covenant newly negotiable. If machine intelligence can reduce cost and expand access, healthcare has an obligation to use it. If machine intelligence displaces workers who built the old system, healthcare has an obligation to help them transition. The status quo isn't morally neutral just because it's familiar. The transition isn't morally innocent just because it's efficient. Both require judgment.

That's why the Healthcare 150 and the AI 10 matter so much.

The Two Tribes: The Healthcare 150 and the AI 10

Let's talk about my intended audience for this essay: first, our "Healthcare 150." At this point in our ongoing conversations you should know them well—the small coterie of leaders who, visibly and invisibly, govern the gravitational field of American healthcare. These are the 84 men and 16 women CEOs guiding roughly \$1.175 trillion in revenue out of the behemoth \$1.6 trillion hospital and health system sector;² the 33 Blue Cross Blue Shield CEOs underwriting coverage for roughly a third of the nation, some 118 million Americans;³ the seven managed-care titans whose collective market capitalization fell calamitously—by almost half—in the past year to a low of roughly \$500 billion;⁴ and the 10 biopharma CEOs commanding a mammoth \$3.7 trillion in market capitalization.⁵ Add to that the breath-of-fresh-air, energetic leadership now ensconced in Washington in Dr. Oz, Chris Klomp, Steph Carlton and their talented team. To round it out, for good measure, let's throw in Judy Faulkner of Epic.

And there you have it—our *healthcare oligarchs*, presiding over what is easily the most oligopolistic, personality-dominated, incrementalist, and regulatory-captured sector in the U.S. economy.

But please don't mistake my sharp characterizations here for cynicism. I have *reverence* for these leaders. Most are dear personal friends, many for two decades or more, and while statistically you might find the occasional narcissist or megalomaniacal personality somewhere in the distribution—as one does in any concentration of power—the overwhelming majority are deeply altruistic, hard-working, community-focused, and sincerely devoted servant-leaders. They, like the rest of us, are bravely trying to navigate an unprecedentedly confusing and vertiginous moment in world history. I applaud their devotion and hope that my meditations in this essay—and the podcasts, LinkedIn posts, speeches, and other venues in which I am, admittedly, somewhat insufferably overexposed—might offer a small increment of clarity as they attempt to steer this immense system through turbulent waters.

The "AI 10," on the other hand, are a rather more colorful and heterogeneous constellation of personalities.

They are the main protagonists in our unfolding AI drama, predominantly clustered in San Francisco and Silicon Valley, with the occasional outpost in London. They have entered the cultural zeitgeist so thoroughly that many of them now operate almost as single-name franchises—*Dario*, *Elon*, *Mira*, *Satya*, *Fei-Fei*, *Sam*, *Vinod*, *Demis*, *Jensen*, and *Zuck*, who receives the distinctive honor of a last-name-first-syllable appellation—alongside an ever-growing cast of founders, entrepreneurs, and capital allocators investing and building frenetically atop the god models.

These individuals represent some of the most fecund and consequential minds ever produced among the roughly 117 billion humans who have walked the earth.⁶ They are also, it must be said, a remarkably eccentric group of personalities. I've come to admire them deeply, and in several cases to count them as friends, but there are moments when some appear less like sober statesperson and more like children gleefully assembling nuclear toys.

² Beckers Hospital Review: "102 health systems ranked by annual revenue" published on April 22, 2026 ([Link](#))

³ Blue Cross and Blue Shield (Data for 2024) ([Link](#))

⁴ Yahoo Finance: \$UNH, \$ELV, \$CVS, \$CI, \$HUM, \$CNC, \$MOH – 52-week high/low. ([Link](#))

⁵ Companies Market Cap – Largest Pharma Companies by Market Cap accessed on June 8, 2026 ([Link](#))

⁶ Population Reference Bureau: "How Many People Have Ever Lived on Earth" published on November 15, 2022 ([Link](#))

The Mandate: Bridging the Two Worlds

This, then, is the universe of minds I am trying to influence: the progenitors of the god models, the innovators building atop them, the policymakers constructing the regulatory surround, the establishment healthcare mandarins who serve as gatekeepers to our \$5.3 trillion, 18.0% of GDP monolith, the early-, mid-, and late-stage capital allocators from venture, growth, and private equity, and a few other stragglers orbiting around the edges.

You could summarize my ambition as a contrapuntal one, and so simple as to be naive: *evangelize healthcare to the AI 10 and AI to the Healthcare 150*. The narrow sliver where those two worlds intersect—where exponential technology collides with the most bureaucratic, regulated, and, yes, *sacred* sector of the economy—is the terrain that matters most. At its core this is a collision of ethos. On one side is Zuckerberg's famous injunction to "move fast and break things." On the other is Hippocrates' ancient admonition, often paraphrased as "first, do no harm." Twenty-four centuries separate those philosophies. Bridging them before velocity forces the issue is one of the central mandates of this essay.

The industrial logic of this strategy—to influence the influencers, decide the deciders, and invest capital behind the asymmetries at these intersections—is eminently rational in a moment like this. These episodic historical moments of upheaval—economic, societal, and cultural—are precisely when we possess a fleeting window of instrumentality, agency, and perhaps even jurisdiction to shape events before they harden into path dependency. Technological revolutions almost never diffuse evenly across societies. Instead, they propagate through a tiny number of *institutions* and a tiny number of *leaders* capable of acting decisively. The comforting fiction is that transformations of this magnitude will be democratized from the bottom up. They rarely are. They aren't volunteered for by frontline staff, who quite reasonably resist disruptions to their own roles and livelihoods. Large institutional shifts seldom originate at the periphery. They originate at the center.

Which leads to a slightly uncomfortable observation. In moments like this, diffusion tends to occur less through *deliberation* than through *decisiveness*. **Autocrats**—whether at the level of countries, companies, or CEOs—have historically proven more effective during the implementation and installation phase of technological revolutions than systems organized primarily around consensus. I'm almost sorry to say that aloud. My own liberal proclivities would much prefer a more democratic narrative. But if one studies the historical record carefully—from the diffusion of steam power to electrification to computing—the pattern becomes unmistakable: when the underlying technological substrate begins shifting rapidly, *velocity tends to beat consultation*.

And this moment will eclipse those earlier transitions in every conceivable way. As Demis recently put it, artificial intelligence *"is going to be ten times the impact of the Industrial Revolution, but happening at ten times the speed."*⁷ If that framing is even notionally correct, the small constellation of actors capable of moving decisively during the installation phase—**our Healthcare 150 and the AI 10**—will shape the economic order that follows. The difficulty, of course, is that the window for deliberate action is vanishingly small. When the technological stack itself evolves month by month and hesitation can deal you out of the game before you realize it, centralized decision-making will outpace consultative enfranchisement. The individuals and institutions that move first will establish the operating system everyone else eventually runs.

So yes, there is a certain harsh *realpolitik* embedded in this argument. I'm not describing the world as

⁷ Demis Hassabis: "India AI Impact Summit Keynote" delivered in February 2026 (Link)

I sentimentally wish it to be, nor am I being *Panglossian*—one of Dario's favorite words (and you guys complain about *my* vocabulary) and assuming this is the best of all possible worlds. I'm simply trying, however imperfectly, to describe how large systems tend to behave under the pressure of exponentials.

Let's collapse these two hemispheres together. The frontier labs have extraordinary technical brilliance, but they don't possess the accumulated tacit knowledge of care delivery, medical liability, patient trust, clinical workflow, public finance, and the sacred ministry of touching actual bodies. The Healthcare 150 possess that wisdom, but they've largely missed (or half-metabolized) *every major technology wave of the past generation*: internet, mobile, cloud, big data & analytics, enterprise SaaS, robotic process automation, you name it. They absorbed the *vocabulary* of innovation without absorbing the *productivity* that usually attends it. Digitization arrived. Productivity didn't. Interoperability was promised but remains, even now, a ritual disappointment. Healthcare evolved just enough to appear modern while preserving the underlying economic structure.

The somber mood in San Francisco

Part of what motivates me, counterintuitively, in bringing the 150 to meet the AI 10 is to convey how all of this *feels*. Because what I feel, viscerally, is a sense of *disorientation*. And to paraphrase Leopold Aschenbrenner, who wrote 2024's seminal, must-read *Situational Awareness* paper—you can feel the future first in San Francisco.

And what I've noticed, of late, is that the normally garrulous chattering class of engineers, programmers, and venture capitalists in Silicon Valley—those who chirp incessantly on X and Substack—have recently gone eerily quiet. We should be nervous when the noise stops, like when mischievous children fall silent in the next room; something's happening. But this isn't comic mischief.

There have been hunger strikes (!) outside the offices of Anthropic and DeepMind—literal flesh wagered against exponential code. The velocity of the technology feels destabilizing even, or *especially*, to those building it. There's a quiet undercurrent of unease that this thing may be outrunning not only regulators and politicians, but its creators. The regrettable brinksmanship between the Department of Defense and Anthropic is only the opening salvo in what will be a long, entrenched battle as the tech mercilessly outstrips the regulators' and policymakers' ability to comprehend what's happening.

Dividing my time between San Francisco, New York, and Washington, I seem—regrettably—to absorb the neuroses of all three geographies: the breathless techno-solutionism of frontier labs and insurgent founders in SF; the defensive, capital-preservation instincts of institutional allocators in NYC; and the Kafkaesque, change-by-the-day policy churn of an aspirationally authoritarian DC. Yet it's San Francisco where the disorientation feels most acute because it's here that the exponentials are lived in real time. The technology is metastasizing across sectors, cultures, and epistemic domains. Conversations have drifted from productivity augmentation to debates about agentic deception, alignment fragility, and recursive self-improvement.

My long-standing mental models—*incumbents versus insurgents, invention versus diffusion, moat versus mobility*—feel suddenly provisional. The consultative dialog between the 150 and the 10, while creating shared vernacular and objectives, feels like it is progressing too slowly. And this sense of destabilization isn't merely atmospheric—it is changing how we act.

Prepared Minds and Constant Theory Revision

One of the most disconcerting—and invigorating—aspects of the past year has been the necessity of constantly revising my own strategic framework. Each time I've achieved some semblance of mental-model equilibrium—conceptualized how the establishment can most agilely deploy the new tech; mapped the interplay of reasoning, memory, tool invocation, and agentic harness execution; modeled second- and third-order public and private-market consequences—a brand-new discontinuity forces recalibration. This is what it feels like to reason inside an exponential. You build a frame. The frame works for a minute. Then the next model release, the next Chinese open-weight juggernaut, the next enterprise deployment pattern, or the next market convulsion makes the frame look quaint.

Anthropic's new June 2026 Institute post, *When AI builds itself*, sharpened this point with startling clarity.⁸ The company says AI is already accelerating AI development from *inside the lab itself*: Claude now authors more than 80% of the code merged into Anthropic's codebase, Anthropic engineers are shipping about 8x as much code per quarter as they did from 2021-2025, and models are increasingly running code, delegating work to other agents, executing experiments, and narrowing the human role toward direction-setting and taste. That's the exponentials becoming *self-referential*—intelligence helping build more intelligence—and even Anthropic is careful to say we're not yet at full recursive self-improvement, and that it's not inevitable, while also warning that it *could arrive sooner than most institutions are prepared for*. If the machine begins materially participating in the design of its successor, then healthcare's normal governance tempo—monthly committees, quarterly updates, annual strategic plans, the whole cozy liturgy of institutional deliberation—will look not merely slow but ontologically mismatched to the thing it's trying to govern.

That's why I want the opening to carry a little epistemic humility, but not the whole operational detour. The fact that frameworks are decaying faster doesn't mean we stop building frameworks. It means we build them provisionally, revise them quickly, and resist the seduction of doctrinal certainty. Incumbents may still possess real advantages—trust, distribution, regulatory muscle memory, data exhaust, patient access, workflow intimacy—but the half-life of those advantages is shrinking. Application-layer startups may still build durable companies, especially in healthcare, but the god models keep expanding horizontally and reanimating the thin-wrapper critique. Nobody gets to be complacent here. Not the 150. Not the AI 10. Not me.

Again: my core thesis, enunciated to you in too many essays, podcasts, and interminable speeches over the past year, has been that *incumbents moving at the speed of insurgents can prosecute a defensible strategy of anti-disintermediation* by weaponizing their establishment advantages: large installed base, reputational equity and legitimacy, regulatory and litigation navigation skills, sales distribution, patient trust, data exhaust, and permissioned access to mission-critical workflows. I still believe this is directionally correct. *But the half-life of that advantage is shrinking*. The diffusion curve of the base models is collapsing the window between "new capability exists" and "everyone has it." Incumbents have less time than I originally thought. And that makes me worried.

The same shrinking window applies to the insurgents themselves. The application layer still matters, especially in healthcare, where workflow intimacy, liability posture, trust, and proprietary data can create some moat defensibility. But the relentless horizontal skill expansion of the god models keeps reanimating the thin-wrapper critique with unnerving force. What looked like durable workflow ownership one quarter can resemble temporary scaffolding the next. Tools celebrated one week are

⁸ Anthropic: "When AI builds itself: Our progress toward recursive self-improvement, and its implications" published on June 4, 2026 ([Link](#))

commoditized the next. The asymmetry is brutal: application companies have to be right *now*; the base models only have to be right *eventually*. Navigating that environment requires extraordinary neuroplasticity—not just from founders, but from the Healthcare 150.

That's why Pasteur's notion of the *prepared mind*, which I'll invoke repeatedly across this essay, matters. Preparedness here doesn't mean having a finished doctrine. It means being willing to revise doctrine faster than a large institution usually likes to revise anything. It means enough humility to say, "I no longer believe the thing I believed three months ago," and enough courage to act anyway. The prepared mind isn't rigid. It's alert.

The Lessons of Installation

The broad historical lesson is simple enough to say and hard enough for incumbents to metabolize: *technological revolutions are won less by inventors than by installers* (for an *outstanding* read on this, I highly recommend Jeffrey Ding's *Technology and the Rise of the Great Powers*).⁹ Indeed, we tend to celebritize the inventors—Watt, Edison, Gates, Altman, Hassabis, Amodei—because inventors make better protagonists. But the productivity gains, the geopolitical hegemony, the institutional transformations, and the changes in ordinary life came from *diffusion*. Steam power mattered when mechanical engineers embedded it into factories, mines, ships, and railways. Electrification mattered when electrical engineers, utilities, financiers, regulators, and municipalities built grids and reorganized production around them. Computerization mattered when software engineers and operators installed computation across firms, supply chains, finance, logistics, and the home.

The hero of the installation phase is often anonymous: the engineer, operator, implementer, systems builder, and workflow translator who takes the glamorous invention and makes it boringly, relentlessly useful. That's why I've been so interested in *forward-deployed engineers*, the Palantir archetype, and the possibility that the unsung hero of agentification would be the person who embeds these systems into the real economy. But here too my view has become less settled. If the interface to computation becomes natural language, if models can increasingly reason about workflows, code, infrastructure, and data structures, then perhaps the diffusion layer collapses inward. Perhaps *everyone* becomes, in some limited but operationally important sense, a forward-deployed engineer.

That possibility changes the mandate for the 150. The question isn't simply which vendor or model to buy. The question is whether a large healthcare institution can *mobilize its own workforce* to become an installation surface: clinicians, operators, coders, analysts, pharmacists, schedulers, managers, and executives all learning, testing, prompting, building, verifying, and diffusing in the flow of work. The diffusion layer may not be a narrow priesthood. It may be an institution-wide competency.

This is one of the reasons the old anti-disintermediation playbook won't work. *Regulatory enclosure, narrative warfare, interoperability chokepoints, committee delay, safety language as incumbency protection*—these may buy time, but time isn't strategy. The old technological waves could be attenuated at healthcare's gates because the sector was heavily regulated, locally embodied, credentialed, reimbursed through strange hydraulic machinery, and emotionally defended by the sanctity of care. This time the technology works on the very substrate healthcare uses to protect itself: cognition, documentation, expertise, bureaucracy, and professional interpretation.

The thing behind the moat has learned to read the moat.

⁹ Jeffrey Ding: "Technology and the Rise of Great Powers: How Diffusion Shapes Economic Competition" published on August 20, 2024 (Link)

A Modern Republic of Letters

Before we get tactical, I'll indulge myself with a brief historical detour—apologies in advance, though by now you've probably come to expect such side-quest excursions from me. One of the most illuminating insights from studying the Enlightenment—that glorious efflorescence of intellectual energy in seventeenth- and eighteenth-century Europe—is how it created the cultural and institutional *environment* that made the Industrial Revolution possible. Europe, unlike the large and relatively monolithic empires of China or India at the time, was politically fragmented into a patchwork of nation-states, duchies, principalities, city-states, and trading republics. Yet despite this political fragmentation, it was culturally and intellectually interconnected.

What emerged from that peculiar configuration was something quite remarkable: a *distributed but highly interactive elite*—a “Republic of Letters”—in which philosophers, scientists, inventors, merchants, and statesmen exchanged ideas across borders through letters, pamphlets, salons, and personal networks. Knowledge circulated rapidly for the era. Ideas were debated, criticized, replicated, and improved upon. Failures in one jurisdiction became lessons in another. Political fragmentation, paradoxically, created entrepreneurial dynamism, because no single orthodoxy could suppress experimentation everywhere at once. The result was a kind of early *intellectual network effect*, where ideas diffused quickly across an elite community capable of testing, refining, and spreading them.

And if you squint a little, the pattern begins to look strangely familiar.

Our modern healthcare oligopoly displays something of the same isomorphic structure: organizations large enough to resemble *nation-states* in economic scale (more on this comparison later), politically fragmented but culturally and intellectually intertwined, dominated by a relatively small set of recognizable personalities who spend a remarkable amount of time talking to one another. Honestly, is there any group of CEOs that enjoys conferences with one another quite as much as healthcare and AI leaders do?

The Enlightenment philosophers exchanged their ideas epistolary-style. The Healthcare 150 and the AI 10 have more efficient communication tools at their disposal. They can post on X, send emails, jump onto Signal threads, convene boardroom dinners, host salons, or—occasionally—write over-long and pretentious essays like this one. Ideas travel faster now than they did in the age of Voltaire and Hume. And, importantly, the actors involved are *influenceable*. The Healthcare 150 and the AI 10 together form a numerically small but extraordinarily consequential network of decision-makers capable of experimenting, tinkering, prototyping, synthesizing ideas, and diffusing successful approaches across institutions at remarkable speed.

That, in essence, is the objective of bringing the Healthcare 150 and the AI 10 into conversation together. In a way, that's also the beauty of an oligopoly: you don't have to persuade the whole system democratically. You need one or two catalytic leaders, with the energy, urgency and, yes, *autocracy*, to effectuate a change. And the system, Darwin-style, begins to follow.

That's why I'm less interested in democratizing my attempted persuasions broadly. I just want *you*—my small and indulgent readership—to give some earnest thought and consideration to these ideas.

Indeed, a lot of my own learning over the past year has come from watching the sparks fly when the AI and healthcare tectonic plates collide. Over the past twelve months I have brought more than a hundred healthcare leaders “on safari” with me to Silicon Valley—to observe the local flora and fauna, as it were—for intimate dinners, immersive working sessions, and uncensored conversations with the AI 10. Reciprocally, I've brought members of the AI 10 into the boardrooms and hometowns of the

Healthcare 150. Slowly but unmistakably, through the efforts of many of us collectively, the mutual incomprehension that once characterized these groups has begun to dissipate.

We are inching—perhaps haltingly, but perceptibly—toward the possibility that these two tribes might actually *collaborate* in designing the next stage of our civilization. But we don't have a lot of time remaining to get this right.

The Opening Prescription

That's why this essay will keep returning to the same cluster of imperatives.

Use the tools. Diffuse a god model—ultimately pick one, monogamously—and its surrounding harness into the enterprise. Study the exponentials. Learn the nomenclature. Build a task map. Stop hiring reflexively into exposed workflows. Treat attrition as a gift before layoffs become the only instrument left. Collapse the coordination tax. Redesign apprenticeship before the old repetition disappears. Govern agents as a new labor class. *Pair insurgent speed with incumbent legitimacy.* Invest in compute, not just beds. Don't sell your health plan casually. Prepare for payer attenuation. Move on behavioral health first. Move on clinical AI quickly but safely. Build the new liability architecture. Lessen our dependence on Taiwan while remaining open enough to compete with China rather than merely posture against it. And above all: make the surplus visible and shared, or the politics will curdle.

The historical lesson, then, isn't a charming preface. It's the operating doctrine. The winners in technological revolutions are rarely the inventors alone, and almost never the complacent incumbents. They are the actors—nations, institutions, companies, and occasionally individuals—who diffuse the technology faster, mobilize resources more decisively, and reorganize their institutions more ruthlessly than the rest. The opening chapter therefore has to carry this history. Subsequent chapters can then do their work: translate the history into CEO statecraft.

What Have We Summoned?

The opening question is therefore not merely whether healthcare will use AI. It will. The question is whether healthcare will help *design the covenant* under which AI enters the sector, or whether it will wait for others—frontier labs, hyperscalers, insurgents, payers, capital markets, China, or the Gulf—to design that covenant on its behalf.

The chapters that follow move from intelligence itself to clinical medicine, from medicine to labor, from labor to healthcare's workforce and civic architecture, from there to hospitals, payers, behavioral health, China, theology, BCI, and finally back to prescription. The route will be discursive, perhaps occasionally maddening, but the argument is laminar. Intelligence is becoming scalable. Scalable intelligence attacks cognitive scarcity. The attack on cognitive scarcity changes labor, professions, institutions, and capital. Healthcare is maximally exposed because it is labor-intensive, data-rich, biologically complex, administratively encumbered, and morally indispensable. The institutions that diffuse AI fastest will reshape the sector. The institutions that wait will be shaped by others. The workers harmed by the transition deserve candor and generosity. The patients harmed by the status quo deserve affordability and access. The leaders of the 150 deserve no illusions.

This is not a call for panic. Panic is useless. This is also not a call for worship. Worship is worse. It's a call for prepared minds, moral imagination, operational courage, and enough autocracy—carefully bounded, institutionally sane, and ethically serious—to move large systems before large systems are moved for them.

The Chapter, Compressed

Here's the chapter, compressed into the governing takeaways.

First, AI isn't another normal technology cycle. It is the industrialization of intelligence, and that means the substrate of science, labor, management, medicine, and culture is changing at once.

Second, healthcare is in the flight path because it's where the abstractions become *human*: the failing body, the frightened family, the exhausted clinician, the administrative maze, the labor market, and the civic institution all converge.

Third, the Healthcare 150 and the AI 10 are the two tribes that have to learn one another's language before velocity hardens into path dependency. One side has the models; the other has the patients, trust, workflows, liability, and sacred touch.

Fourth, the central historical lesson is *installation*. The winners in technological revolutions are rarely the inventors alone; they are the institutions and nations that diffuse the technology horizontally and reorganize themselves around it.

Fifth, the old *anti-disintermediation* playbook won't be enough. Regulatory enclosure, narrative warfare, committee delay, and safety language as incumbency protection may buy time, but time isn't strategy.

Sixth, the covenant has to be built deliberately: use the tools, diffuse the intelligence, collapse the coordination tax, tell the truth about labor, share the surplus, and govern the machine like adults.

BEFORE WE TURN THE PAGE

To answer the question of what kind of covenant healthcare should build around AI, we first have to understand what exactly has been multiplied.

So we begin not with hospitals, payers, doctors, or even labor, but with intelligence itself.

Chapter 2 · Future of Science

The End of Hypothesis, Stagnation, and Generative Epistemology

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

— ISAAC ASIMOV, 1988

A WORD ON NAVIGATING THIS CHAPTER

This chapter names the thing itself before the essay turns to medicine, labor, institutions, and geopolitics. It argues that GenAI is a multiplication of intelligence, that science is entering a generative epistemology regime, and that the 150 must become stewards of synthetic discovery rather than spectators to it.

The chapter moves in six parts. **Part I** asks why intelligence is the upstream input to civilization itself. **Part II** asks why this moment arrives against a background of stagnation, cognitive overload, and institutional sclerosis. **Part III** follows the center of gravity westward, from the old priesthood of institutional science toward frontier labs, computational biology, and small insurgent teams. **Part IV** develops the central epistemic claim: from Bacon to backpropagation, from hypothesis to generative epistemology. **Part V** turns the whole machine loose on biology, where the returns to intelligence may be highest and where the compressed 21st century becomes less slogan than operating hypothesis. **Part VI** brings the argument back to humans, the AI 10, the 150, and the post-hypothesis covenant of stewardship. That's the road. Some scenic detours remain, because apparently I am who I am, but the spine should be visible.

PART I—THE PRIME MOVER: INTELLIGENCE AS CIVILIZATION'S UPSTREAM INPUT

Why Intelligence Is the Prime Mover

The animating premise of this chapter is that intelligence isn't one input among many. It's *the* input that discovers the other inputs, organizes them, disciplines them, commercializes them, and turns them into tools, firms, states, hospitals, drugs, machines, and, yes, civilizations. A civilization is, in one sense, the accumulated residue of *intelligence made durable*: laws, instruments, institutions, books, circuits, therapeutics, habits, languages, and the social machinery by which one generation transmits its discoveries to the next. If GenAI multiplies intelligence itself, then the downstream shock isn't merely technical. It's civilizational.

5,000 Years of Human Civilization and Technology—In a Few Paragraphs

Forgive a brief pedantic history-lesson digression, but if we're talking about epoch-shaping technology, and about to launch into a loud monologue on the future of science, it seems worthwhile to dedicate a paragraph or three to compressing roughly 117 billion humans' collective effort into the civilization we've inherited. I referenced earlier the Enlightenment and the diffusion dynamics that *preceded* the Industrial Revolution; this is the longer arc behind that moment. It's also one of the reasons I'm so evangelistic about technology in general, and specifically optimistic about GenAI as a societal boon—we just need to survive, to borrow Dario Amodei's memorable phrase, our present *technological adolescence*.

Ok, here we go. Across the span of recorded history—roughly 5,000 years since the invention of writing—there have been perhaps twenty to twenty-five seminal, civilization-shaping technologies: the economists' and academics' so-called *general-purpose technologies*, a term that feels almost cosmically appropriate given our present fascination with the *other* GPTs, *generative pre-trained transformers*. We all know the canonical examples: the wheel, the printing press, the steam engine, the electric grid, the microprocessor. These technologies form the pantheon of anthropogenic tools that have periodically bent the trajectory of human civilization upward. But it's worth remembering just how *recently* that upward swing actually began in earnest. For most of our species' existence, and for the overwhelming majority of the roughly 117 billion humans who have walked the earth, life was, well, unpleasant. Up until the late eighteenth century, humanity was stuck in a *Malthusian trap*: each advance in productivity was quickly swallowed whole by population growth, leaving living standards stubbornly pinned near subsistence. At the starting gun of the Industrial Revolution in the 1760s, the global population was 800 million, life expectancy averaged thirty years, and ninety percent of humanity¹⁰¹¹¹² lived in conditions we'd now describe as immiseration. Sure, a small aristocratic elite—kings, warlords, landowners, clerics—lived comfortably enough, thank you very much, but prosperity was tightly concentrated and economic growth moved at an inching, geologic pace. Global GDP prior to 1700 expanded at 0.1% per year,¹³ meaning the entire world economy doubled only every millennium or so.

What changed that trajectory was the intellectual upheaval we discussed earlier: the Scientific Revolution and Enlightenment of the sixteenth and seventeenth centuries, which created a cultural *environment* of skepticism, experimentation, and iconoclasm. It was the moment when European thinkers began systematically rejecting the ossified authority of medieval scholasticism and Aristotelian orthodoxy, replacing it with empirical inquiry and scientific method. That intellectual shift—Newton, Bacon, Galileo, Boyle, and the wider *Republic of Letters*, yes, them again—created the *epistemic preconditions* for the Industrial Revolution. And when that revolution finally landed in the late eighteenth century, it triggered what can only be described as a Cambrian explosion of technological and economic development. Over the next 250 years humanity experienced the *most compressed period of progress in its entire history*: population expanding from 800 million toward more than 8 billion, life expectancy jumping from thirty years toward eighty in many wealthy societies, global poverty collapsing from ninety percent of humanity to something less than ten, and economic growth accelerating from microscopic rates to a pace that now doubles global output roughly every fifteen to twenty years. The cascade of *mechanization, electrification, and*

¹⁰ Our World in Data: "Population Growth Over Time" published on June 1, 2023 ([Link](#))

¹¹ Our World in Data: "Key Insights on Life Expectancy" published in 2023 ([Link](#))

¹² Human Progress: "The Great Miracle of Industrialization" published on May 6, 2019 ([Link](#))

¹³ The Maddison Project: "Maddison Project Database 2023" published in 2023 ([Link](#))

computerization—each building on the diffusion dynamics discussed earlier—transformed not just economies but the human condition itself.

Across almost every dimension by which we measure civilization—longevity, safety, sanitation, literacy, urbanization, and even the spread of democracy—technology has been the central engine of improvement. That doesn't mean technology is inherently benevolent. As the philosopher Paul Virilio said, "when you invent the ship, you also invent the shipwreck."¹⁴ Every technological capability carries a corresponding failure mode. The same tools that enable the eradication of disease can enable the engineering of pathogens; the same computational systems that democratize knowledge can concentrate power or propagate misinformation. But taken in the long arc of history, the trajectory is unmistakable. Our species advances by building tools that amplify our ability to understand and reshape the world.

Which is why this present moment feels so consequential. If GenAI truly represents a multiplication of intelligence itself, then we may be witnessing the next entry in that long lineage of general-purpose technologies. And if the historical pattern holds—if *diffusion* and *mobilization* ultimately matter more than *invention*—then the institutions that learn to absorb and propagate this technology most effectively will shape the next chapter of human civilization.

After all, for better or worse, we are a species defined by the tools we build. We are, in the most literal sense, *homo technicus*.

The Question the Incumbents Can't Evade

With that sweeping, and massively reductive, historical framing in mind, let's now turn to the question at hand: what is the role of our Healthcare 150 in scientific and biomedical discovery in this new technological paradigm? And as I'll argue in a moment, in a *post-biological-intelligence* world? And how can the AI 10 empower this?

Ok, I can *feel* you rolling your eyes. Now we get subjected to some impenetrable, self-indulgent monologue on philosophy, metaphysics, and maybe a little theology thrown in there too. Can't Larsen just stick to the topic at hand and tell us, concretely, what healthcare leaders should do? The answer is a) no, I can't quite stick to the topic, and b) the implications of this section for healthcare are monumental, so I'm going to plow forward bravely—brazenly?—anyway. And seriously: for the more pragmatic and literal-minded of my readership, you may wish to just skip this chapter altogether and save an hour.

But for those of you who gamely—or masochistically—persevere, let me at least tell you where I'm going. This chapter is really about three intertwined questions. **First**, what happens to science—especially biomedicine—when intelligence itself becomes scalable, synthetic, non-biological, and increasingly capable of *participating* in discovery. **Second**, and more fundamentally, what happens to *humanity* when the faculty that has most defined our species—our intelligence—begins to be democratized, commodified, and, in important domains, surpassed. And **third**, what all of that means for the commanding institutions of healthcare: whether they will shape this transition, or merely be shaped by it.

This isn't a chapter about AI as another research assistant, another literature-review tool, another faster search interface, or another clever way to generate grant prose slightly less painful than the current NIH ritual humiliation. The argument is much larger than that. I'm arguing that GenAI, and the broader family of frontier models, biological foundation models, multimodal systems, autonomous

¹⁴ Paul Virilio: "Politics of the Very Worst" published on January 5, 1999 ([Link](#))

agents, robotic labs, agentic harnesses, and synthetic-discovery platforms, *will change the epistemology of science itself*. That sounds preposterously grandiose, even by my standards, but I think it's directionally right. For roughly four centuries, science has operated under a human-centered discovery regime: human minds generate hypotheses, human institutions validate them, human laboratories test them, human journals ordain them, and human prestige hierarchies govern their diffusion. That regime isn't disappearing tomorrow. But its monopoly is breaking.

So here's the laminar flow of the chapter. I'll begin with the claim that GenAI isn't just another tool but a *multiplication of intelligence*, and that this matters because intelligence is the upstream input that produced every other downstream input in civilization. I'll then ask what, exactly, we've multiplied: a synthetic intelligence trained on the accumulated residue of human civilization itself. From there I'll describe four properties of this alien intelligence—*functional omniscience*, *polymathy*, *omni-disciplinarity*, and *analogic recombination*—and explain why those properties matter so profoundly for scientific discovery. Then I'll turn to *stagnation*: why modern science, especially biomedicine, has become slower, more specialized, more *bureaucratic*, and more burdened by the sheer accumulation of knowledge. After that, I'll argue that our biological response to complexity has been *hyperspecialization*, *homogenization*, and *bureaucratization*—all rational, all understandable, and all increasingly inadequate. Then we'll move west, literally and figuratively, toward the insurgent model of science: frontier labs, computational biology, venture-backed discovery, Sand Hill Road, and *small teams* that disrupt rather than *large teams* that refine. Only then do we get to the central proposition: the scientific method, civilization's greatest cognitive technology of the past four centuries, is no longer enough in its old monopoly form. What comes next is what I'm calling *generative epistemology*.

The title, *The End of Hypothesis*, is intentionally provocative. I don't literally mean humans will stop forming hypotheses, or that Bacon, Popper, Kuhn, and the whole glorious apparatus of Enlightenment empiricism get tossed into the civilizational dustbin with fax machines and ICD-10 coding tips. I mean that the human monopoly on hypothesis generation is beginning to break. The machine won't merely help us test ideas. It will generate candidate landscapes, traverse possibility spaces, simulate biological realities, design molecules, propose mechanisms, and surface truths that may be *verifiable* long before they are fully *explainable* to us. That's the discontinuity.

And then, finally, I'll bring the whole thing back to the 150. Because the practical question beneath all the philosophy isn't whether the 150 should become bench scientists, foundation-model researchers, or computational biologists. They won't. The question is whether the great institutions of healthcare—the NIH, health systems, academic medical centers, payers, biopharma companies, research networks, and the people who govern them—can become stewards of synthetic discovery. Can they supply the clinical substrate, longitudinal data, real-world validation, ethical legitimacy, translational pathways, and patient trust needed to turn machine-generated insight into human benefit? Or will they use prestige, process, regulatory drag, and proceduralism to slow the future until the future simply routes around them?

I've been thinking intermittently about these questions for a long time, mostly as a nagging, indistinct worry in the back of my mind: what, exactly, is the role of humanity in a post-intelligence-diffusion world? What happens to us *civilizationally* when this multiplication of intelligence moment renders abundant, commodified, and no longer exclusively human the very faculty that gave us dominion over the earth? And on the sociological and economic fronts, does it presage a more *stratified* society still—more concentrated wealth and power, the diminishment of the prestige professions, and perhaps even a restive middle that eventually turns politically, then socially, and perhaps

ultimately *violently*, against the whole arrangement?

And then there is the deeper, more existential question: what becomes of the deeply human need for *frontiers*—for the horizon, for the undiscovered country, for the next thing over the hill—when an omniscient silicon intelligence arrives on the scene? What becomes of the “mere” human in a civilization whose central generative input is no longer scarce, no longer exclusively biological, and no longer ours alone?

The original Apollo mission—resurrected in public consciousness this April with Artemis II¹⁵ after a dormancy of several decades—wasn’t just a geopolitical contest or an engineering miracle. It was a *frontier project*. It was the New Frontier made literal. And I think that frontier instinct runs much deeper than politics or prestige; it’s bound up with the very meaning of what civilization *is*. We are the species that doesn’t merely *survive* in the world, but tries to know it, map it, reorganize it, and push beyond it. Which is why this chapter matters. Because if the central faculty through which we have always done that—intelligence itself—is now being multiplied outside the skull, then we aren’t talking about another software wave, or another productivity tool, or another category of enterprise technology. We’re talking about a change in the underlying input to civilization itself.

And that brings us to the practical question beneath all the philosophy. If we’re entering a period of *punctuated equilibrium* in science and biomedicine, with what Dario beautifully calls the “compressed 21st century”¹⁶—five to ten years delivering the sort of biomedical and scientific progress that might otherwise have taken fifty to one hundred—how should the 150 prepare? Not just as individuals, as leaders and stewards, but as *custodians of institutions*, some of which sit in the firmament of the most prestigious establishment organizations in the world and have survived, and often thrived, for a century or more. How do archetypically linear institutions reconcile themselves to the most exponential technology in history?

Will these institutions rise to the moment and help negotiate, refract, and install this technology through their organizations? Or will they assume an oppositional, ancien régime posture—*weaponizing* regulatory enclosure, procedural drag, and their still-considerable if diminishing prestige to slow progress, only to discover that they’ve purchased not safety but diminution? Or will they do what the best incumbents do at real inflection points: bring wisdom, expertise, tacit knowledge, and some measure of sagacity and supervision to the transition, and thereby help govern it rather than merely resist it? This is, in many ways, the quintessential *incumbents-versus-insurgents* problem. And right now, on this frontier, the insurgents are winning.

That, finally, is the point of this chapter: to help the 150 see clearly what is coming. Scientific and biomedical advances may soon arrive *phantasmagorically*—one of my favorite cinematic terms, the image of the monster looming larger and larger until it is suddenly right in your face—with such force, and such velocity, that they threaten to overwhelm our antiquated systems. The center of gravity is moving, and moving irrevocably, from the incumbents to the insurgents. And the insurgents possess many of the attributes that matter most in moments like this: youth, energy, capital, compute, risk tolerance, iconoclasm, and an almost theological irreverence for authority. Meanwhile, the geopolitical counterfactual should keep us awake at night: China and the GCC, by virtue of state capacity, capital concentration, and far fewer procedural veto points, may well prove better than we are at advancing and operationalizing some of these breakthroughs.

So yes, I’m going to take a somewhat discursive walk through the nature of intelligence, the past four

¹⁵ National Aeronautics and Space Administration: “NASA Welcomes Record-Setting Artemis II Moonfarers Back to Earth” published on April 10, 2026 ([Link](#))

¹⁶ Dario Amodèi: “Machines of Loving Grace” published in October 2024 ([Link](#))

centuries of post-Scientific-Revolution discovery and invention, and why I think we need to ready ourselves for a genuinely new regime. But this isn't some academic parlor game, and it isn't philosophy for its own sake. It is about whether the commanding institutions of healthcare—and the people who lead them—will have a voice in this next chapter.

Not Another Tool—A Multiplication of Intelligence

All of which brings us to the deeper claim underneath this chapter. If the insurgents are winning, if the center of gravity is shifting, if the 150 are being forced into a choice between shaping this transition and merely enduring it, then the obvious question is: why? Why is this technology so different? Why does it justify all this grand talk of punctuated equilibrium, institutional obsolescence, and civilizational phase shift?

The answer, I think, is that there's something qualitatively different about this technology—a difference in kind, not merely degree. Let's return to first principles for a moment. What's so categorically different about this tool, and why is it likely to remake civilization as we know it? The answer lies in a deceptively simple observation: GenAI is fundamentally, irreducibly, a multiplication of intelligence itself.

Whether one prefers to skeptically characterize these systems as *stochastic token predictors* or as early forms of machine reasoning increasingly becomes a boring and irrelevant semantic distinction when the outputs demonstrably extend beyond memorization into synthesis, planning, inference, autonomous action, and, increasingly, discovery. What matters in practice is that we've now *speciated*, for the first time in human history, a non-carbon intelligence that in many domains already matches—and in certain narrow ones exceeds—the performance of our biological cognition.

That fact alone should give us pause, because as *Homo sapiens*—literally “wise humans”—our entire civilizational identity rests upon *intelligence* and *sociability* as our defining traits. For millennia, the fundamental unit of intelligence in the economy was the human brain. Every technological breakthrough, every scientific discovery, every economic system that followed ultimately emerged from the accumulation and recombination of those biological cognitive units. The past 250 years of telescoped progress—the Industrial Revolutions of *mechanization*, *electrification*, and *computerization*—were themselves the downstream product of that combinatorial force. Human intelligence discovered physical laws, organized production, engineered machines, and gradually constructed the technological infrastructure that lifted civilization out of the Malthusian, too-many-humans-and-too-few-resources trap I've described elsewhere.

What's different now is that we haven't simply invented a new tool. We've multiplied the very input that produced all of those earlier revolutions. If the steam engine amplified human muscle, and the microprocessor amplified calculation, *GenAI amplifies the underlying cognitive faculty that created both*. Intelligence is the ultimate general-purpose input. It governs scientific discovery, economic organization, technological design, institutional decision-making, and cultural production. When that input suddenly becomes *scalable* and *multiplicative*, the downstream consequences propagate through every system that depends on it.

And this, in turn, is why everything I've been describing in the preceding sections leads to this: if intelligence itself becomes cheaper, more abundant, more distributable, and increasingly nonbiological, then of course the old institutional equilibrium begins to shake. Of course the old prestige hierarchies wobble. Of course the incumbents find themselves under siege by insurgents who are younger, faster, more computational, more risk-tolerant, and less reverent toward inherited authority. We're no

longer talking about a tool that helps the old system run a bit better. We're talking about a tool that changes the underlying economics of discovery itself.

This is exhilarating and unsettling in equal measure, in part because these systems don't emerge in the same way earlier tools did. As Jack Clark of Anthropic says, we've *grown* these systems more than *built* them.¹⁷ Their internal cognition—their *neurology*, if that's even the right word—unfolds in vast latent spaces we cannot directly inspect or fully interpret, at least not yet. Which is why Geoffrey Hinton's warning has such chilling force: in evolutionary history, *a less intelligent species has never reliably controlled a more intelligent one*.¹⁸ And yet that's precisely the inversion we are now attempting, on timelines wildly compressed relative to the evolutionary or civilizational tempos that normally govern intelligence hierarchies.

The paradox, then, is that these systems are simultaneously tools and something *more than* tools. They remain extensions of human agency (for now)—machines trained, directed, and deployed by people—but they operate in the very domain that historically defined humanity's comparative advantage. When the Industrial Revolution multiplied physical power, the consequences reshaped production, transportation, and economic growth. When the AI revolution multiplies cognitive power, the consequences propagate through science, governance, markets, medicine, and culture simultaneously. In that sense, what we may be witnessing is less a technological upgrade than a civilizational phase shift: the moment intelligence itself becomes a scalable input to the global economy.

And once intelligence itself becomes abundant, something else begins to happen as well. The *scarcity* that underwrote many of our most prestigious professions begins to erode. If the thing being democratized and commodified isn't merely information, but synthesis, analysis, articulation, problem-solving, unsupervised action and soon, discovery, then the economic logic of prestige labor starts to wobble. That's part of what has been worrying me throughout this essay. And it is also why the questions in this chapter aren't abstract. They are institutional, economic, political—and, eventually, civilizational.

We've Taught the Models Our Civilization

If the argument of the preceding section is right—if GenAI isn't just another tool but a multiplication of intelligence itself—then the obvious next question is: *what* exactly has been multiplied? What kind of intelligence is this? What has gone into it? What, in fact, have we made? Because before we can think clearly about what it will do to science, institutions, or civilization, we have to understand what sort of cognitive creature we're actually dealing with.

Embarrassing to confess, perhaps, but I talk—literally, in voice mode—with the models sometimes for a couple of hours a day about everything: restaurant recommendations, the history of some obscure discipline, analogies I'm struggling to land, *recombinations* of ideas, turns of phrase, synthesis problems, strategy questions—you name it. Because that's what these models are: a thought partner *par excellence*; a synthetic, non-biological intelligence that, in many quantifiable ways, is already superseding our own carbon intelligence.

Start with this: we've fed this silicon, non-carbon intelligence the corpus of digitized human knowledge—our histories, our philosophies, our theologies, our sciences, our logic, our arguments, our accumulated civilizational record. I mentioned this earlier, but it bears repeating, because it's the central point: we have effectively given it the artificial analogue of the two traits that have given us

¹⁷ Jack Clark: "Technological Optimism and Appropriate Fear" published on October 13, 2025 ([Link](#))

¹⁸ CNN: "The 'godfather of AI' reveals the only way humanity can survive superintelligent AI" published on August 13, 2025 ([Link](#))

dominion over the earth and every other known species—intelligence and sociability. Combined, that's our *collective intelligence*: our ability to reason, and to share those insights horizontally with our contemporaries and vertically with our descendants. We have, in a very real sense, taught these models our civilization.

And it turns out our civilization is *learnable*. Perhaps we had romanticized our uniqueness, our *ineffability*, as a species. Demis himself has said he once thought our civilizational complexity was something like “semi-infinite,” but, being the beautiful mind he is, he went so far as to mathematize even that intuition, suggesting that the number of possible human choices and permutations approached something like the following number of possible scenarios:

[illegible]

And yet even Demis seems to have been taken aback by the success of LLMs, indiscriminately ingesting something like 14 trillion words from the internet and somehow, from that primordial verbal chaos, producing the silicon intelligence on which we're all now spilling so much digital ink. Demis's revised estimate is that the relevant space may be closer to *10 trillion*, or 10^{13} possibilities for human choices, because that's roughly the order of magnitude of words—and even more tokens—hoovered up from Common Crawl.¹⁹ I'll mercifully end this digression by only half-channeling Ecclesiastes, and the wisdom Demis himself ruefully invokes: perhaps *there really is nothing new under the sun*, and perhaps we are, after all, a comprehensible and knowable species to this new silicon co-inhabitant of the earth.

And that's why I find myself increasingly uninterested in the more abstruse debates about whether they're "really" thinking, or "really" reasoning, or merely stochastic parrots, or "no more intelligent than a cat"—thank you, Yann LeCun, for that colorful if increasingly irrelevant phrase.²⁰ Those questions are becoming *scholastic* in the old sense: intricate, performative, and ever less connected to the thing in front of us. Because silicon intelligence—whether or not it thinks in recognizably anthropocentric terms—has already surpassed our own in many quantifiable and measurable ways. The more important question, to my mind, isn't whether it flatters our inherited philosophical categories, but *what capabilities emerge* once you fuse computation with the accumulated cognitive inheritance of civilization.

That, in turn, is what I want to get at in the next move. If we have, in effect, distilled a synthetic version of collective human intelligence, what are its defining properties? What makes it so different—not just from prior software, but from *us*?

Four Properties of This Emergent, Alien Intelligence

I'd nominate four properties—four emergent characteristics—that I think might actually be civilization-shaping.

First, it's *functionally omniscient*, having ingested the corpus of human knowledge and, through sheer computational augmentation—Rich Sutton's Bitter Lesson applied at civilizational scale—achieved something close to instantaneous recall. **Second**, it's *polymathic*, possessing mastery, or near-mastery, across myriad domains at once, one of the defining traits of the truly seminal human minds across history (and yes, one might argue these first two traits are redundant, but hear me out as I'll try to delineate what I think are the key differences). **Third**, it's *omni-disciplinary*: not merely

19 Sebastian Mallaby: "The Infinity Machine: Demis Hassabis, Deep Mind, and the Quest for Superintelligence" published on March 31, 2026
([Link](#))

20 Yann LeCun [@ylecun]: “Before we reach Human-Level AI (HLAI), we will have to reach Cat-Level & Dog-Level AI” published on Feb 5, 2023 (Link)

multidisciplinary, but, in some sense, all professions and all professionals at once. **Fourth**, and most consequentially, it possesses a capacity for *analogic* and *recombinant thinking*.

That last one, I think, deserves particular emphasis, because it may be the most important of the four. Yes, I'm going to use the word "thinking," though the pedants will object. I increasingly believe that many of the greatest ideas in history came from analogic thinking. We like to deify the human ratiocinative process, as if brilliance descends fully formed from the heavens into a Newton's or a da Vinci's or even a Musk's mind. But epistemologically that's not how it seems to work. Our finest minds were profound analogic thinkers; they saw *interstitially*, in the spaces *between* domains. I'm channeling Vinod Khosla here when he says no expert ever disrupts his or her own industry.²¹ That's a slight overstatement, but mostly true. The expert is conditioned to know all the reasons something *won't* work. The polymath, steeped in multiple domains, sees the less obvious bridge.

Gutenberg's insight was an analogy: the wine press becoming the printing press. Newton's apple story may be apocryphal in its popular form, but the underlying act of analogic inference—from terrestrial falling bodies to celestial motion—captures the point. Darwin's theory of natural selection didn't emerge from disciplinary isolation either; it was catalyzed by reading Malthus on population pressure and reframing that logic inside biology. The most original minds didn't simply know more. *They recombined more daringly.*

And this is what the models intrinsically do. They're creative in the deepest sense—not in the superficial "new cat image" sense, but in the application of insights from one domain to another, in the *analogizing*, the *recombination*, the *generation* of new propulsion to human and machine progress. And then, unlike us, they can distribute those insights instantaneously across their instantiations (more on this 'Hive Mind' property later). That last point matters more than people may realize. A human genius dies with a skull full of unscalable tacit knowledge. A model can, at least in principle, propagate improvements across an entire machine fleet.

The Human Monopoly of Hypothesis Begins to Break

Demis Hassabis (who is quickly becoming one of my favorite people, if you can't tell) offers a useful taxonomy here: *interpolation*, *extrapolation*, and *invention*.²² Interpolation is the rearrangement of known patterns—the "new cat" generated from millions of existing cat images. Extrapolation is the leap *within* a system—the famous Move 37 in AlphaGo²³, that eerie flash of originality in a 4,000-year-old game. And then there's the most exalted category: *invention*, or *hypothesis*—the generation of genuinely novel propositional knowledge, a new scientific conjecture, a new materials-science possibility, a new theorem, a new explanatory structure.

That's the category that matters most. Because if interpolation is mimicry and extrapolation is extension, invention is where civilization actually moves. It's where the frontier gets pushed outward. It's where nature yields something genuinely new to us. And for most of human history we have assumed, more or less without needing to say so, that this was the inviolable province of biological minds. Hypothesis generation was ours. Conjecture was ours. The Promethean leap was ours.

We aren't fully beyond that world yet. But we're moving toward its end with startling speed. Which is why I have titled this chapter *The End of Hypothesis*, though that title is intentionally provocative. I don't literally mean that humans will cease hypothesizing. I mean that the human *monopoly* on hypothesis generation may be beginning to break. And if that's true, then we're not just looking at a

²¹ StartupHealth: "Vinod Khosla: Why Domain Experts Never Create Breakthrough Innovations" published on December 23, 2025 (Link)

²² Ted Interviews: "DeepMind's Demis Hassabis on the future of AI" published on July 28, 2022 (Link)

²³ DeepMind: "From games to biology and beyond: 10 years of AlphaGo's impact" published on March 10, 2026 (Link)

better search engine, or a better writing partner, or a better software assistant. We are looking at something much more consequential: a non-biological participant in the generation of novelty itself.

The Biggest Deal in History

And once we admit that possibility, we're more or less forced to its logical conclusion. Because if intelligence is the prime mover of civilization—if it's indeed the faculty by which we discovered fire, mathematics, agriculture, metallurgy, the scientific method, electricity, computation, and all our myriad other advances that have made our world more habitable—then the multiplication of intelligence isn't just another technological event. It's a civilizational event. The biggest deal in history.

And to continue the syllogism, if civilization is, in some deep sense, accumulated intelligence embodied in tools, institutions, and transmitted knowledge—conveyed horizontally to our peers and vertically across generations—then a step-change in intelligence itself is categorically different from an ordinary invention. It means the very engine of progress has been amplified. Not a downstream application. Not a second-order convenience. The engine itself.

And that's precisely why I keep returning, ad nauseum, to the phrase *multiplication of intelligence*. It's not so much rhetoric, but description.

Ok, perhaps this sounds too breathless. Fine. Let's stipulate a *lot* of epistemic humility here. The models remain jagged. They hallucinate. They fail in ways that are occasionally funny and occasionally terrifying. Their causal understanding is contested. Their reasoning, such as it is, can be brittle. Their internal logic is opaque. But none of that undermines the core point. Airplanes crashed before they transformed transportation. Early electricity was dangerous before it illuminated the world. The first automobiles were faintly ridiculous contraptions before they reorganized cities, logistics, dating, suburbia, manufacturing, and geopolitics. The presence of immaturity isn't evidence of triviality. It's often the first stage of world-remaking power.

The thing to watch isn't perfection. The thing to watch is *slope*. Or, that now more modern and ubiquitous term, the *exponentials*. So pick your term—the slope, the exponentials, or what have you—they're all moving vertically.

In short, we aren't dealing merely with a better search interface, a smarter scribe, a clever assistant, or another enterprise productivity layer. We're dealing with a new *economics of cognition*: intelligence detached from the skull, copied at software scale, distributed through networks, and increasingly applied to the very act of discovery. If that premise is right, then the rest of the chapter follows almost uncomfortably. The old human monopoly on high cognition starts to loosen. *The prestige professions lose some of their scarcity*. Science becomes less exclusively human. Biology becomes newly searchable. And the institutions that believe their authority rests on inherited prestige rather than actual adaptation begin to look much more fragile than they did a month ago.

PART II—WHY NOW: STAGNATION, OVERLOAD, AND THE BURDEN OF KNOWLEDGE

This second part asks why this multiplication of intelligence arrives with such strange timing. My answer is that it arrives precisely when our existing biological and institutional machinery is *beginning to strain under the weight of what it has already produced*. Science hasn't stopped being brilliant, but it has become slower, more expensive, more specialized, more bureaucratic, and more burdened by its own accumulated knowledge. Ideas may not be exhausted, but they're harder to find. The frontier is farther away. The map is more crowded. Our brains, institutions, journals, grants, committees, and credentialing systems have become bottlenecks. GenAI therefore appears not only as a marvel, but as a rescue operation for a civilization drowning in its own informational abundance.

Stagnation, and Why Our AI Deliverance Arrives Right on Time

This speed, this advancement (if it is indeed that), this *beneficence* may be arriving at an unusually opportune moment, because—and I realize this is contrarian, and a little, or perhaps a lot, disagreeable—progress appears to be decelerating. Here I'm an acolyte of Byrne Hobart and Tobias Huber, whose *Boom was*,²⁴ to my mind, one of the most consequential books of 2024. Their premise is that modern economies have drifted into a kind of stagnation, and that booms and bubbles can, under the right conditions, break that stasis and finance real technological, societal, economic, and cultural breakthroughs. I'll return later to the question of bubbles, *which we are indisputably in with AI*, and make the argument that this is an 'inflection bubble'—a good thing—rather than an economy-destroying one. But for now, the more crucial point is the diagnosis: there's a serious case to be made that civilization's rate of advance has slowed.

This posture will get you disinvited from most DC dinner parties, but I perversely subscribe to it anyway: civilization has been approaching an asymptote for the better part of the last half-century—ironically coterminous with my own lifetime, so perhaps I'm part of the problem. And yes, I know this can sound like grumpy declinism, like one of those octogenarian riffs about the vanished greatness of the mid-century republic. But look at the ledger. From roughly the end of the Civil War in 1865 to Apollo 11 in July 1969, the United States laid down a civilization-scale run of achievement that is almost preposterous in retrospect: continental rail integration, electrification, the telephone, the airplane, mass automobile production, antibiotics, the defeat of polio, the Manhattan Project, commercial jet aviation, the Interstate Highway System after the 1956 act, and finally a man on the moon. This isn't (just) nostalgia; it's an unavoidable and insistent fact pattern. U.S. life expectancy alone rose from about 47 years in 1900 to nearly 70 by 1960, which is to say that "progress" wasn't some vibes-based abstraction but a concrete expansion of human capability and duration.

The intuition is that sometime after that era-defining achievement of Apollo, by degrees so imperceptible they were hard to notice from inside the moment, we drifted into something more *cautious*, more *legalistic*, more *protective*—less a frontier civilization than an ancien régime of procedure. And this has geopolitical implications. To borrow Dan Wang's framing—and a juxtaposition I will return to again and again across this essay—China over this same period became a building, thrusting, **engineering state** while we ossified into a **lawyerly society**: a culture increasingly organized around review, process, litigation, veto points, and the *land of the eternal no*.²⁵ Some of that was a rational correction and remediation. The postwar American buildout did real damage—urban renewal

²⁴ Byrne Hobart and Tobia Huber: "Boom: Bubbles and the End of Stagnation" published on November 19, 2024 ([Link](#))

²⁵ Dan Wang: "Breakneck: China's Quest to Engineer the Future" published on August 26, 2025 ([Link](#))

bulldozed neighborhoods, industrial expansion poisoned communities, and the country emerged from Vietnam morally exhausted. But the corrective went from medicine to poison. It hardened from prudence into proceduralism, from conservation into conservatism, from legitimate safeguard into a generalized NIMBY gestalt.

And the macro data are at least directionally consistent with that story. Labor-productivity growth in the business sector was materially stronger before 1973 than after it. Bloom et al.'s seminal work argues that ideas themselves have become harder to find—famously estimating that sustaining Moore's Law required more than *eighteen times* as many researchers as in the early 1970s.²⁶ And in biomedicine there's the mordant, backwards-spelled companion to Moore's Law: *Eroom's Law*, the observation that drug-discovery productivity has moved in the wrong direction, with the number of new drugs approved per billion dollars of R&D roughly halving about every nine years. That's a devastating little joke, but a revealing one. In other words, it's not crazy to think that *something real has slowed*: not ingenuity as such, but our civilizational ability to transmute ingenuity into visible, world-remaking accomplishment.

Ideas Get Harder to Find

And this notion that ideas get harder to find makes deeply disagreeable sense when you pause to think about it. When a scientific or epistemic domain is fresh, open, and underexplored, you can attack it vigorously from all sides. Knowledge proliferates. Seedlings of ideas appear, grow into great sequoias, and then provide a shady canopy for the next generation of little trees to sprout (metaphor, but I'm going to keep it anyway). It's a glorious, singularly human process: *learning begets learning*. Learning is our great providential gift. But there's a catch, of course. The more we learn, the more we write, the more there is to absorb *before* one can aspire to reach the frontier. What Malcolm Gladwell popularized as the "10,000 hours" idea starts to get consumed less by original creation than by the sheer labor of catching up²⁷—of metabolizing what has already been said, discovered, systematized, and footnoted.

And let's just pause to gasp at how much more information than ever to absorb. Quick aside that might help illustrate. There's that old Eric Schmidt line—famous because it captures the scale-change viscerally—that from the advent of civilization until 2003, humanity created about 5 exabytes of information;²⁸ at the time he said it in 2010, we were creating that much *every two days*. And now the contrast gets even sillier: by 2025, the world was projected to generate about 463 exabytes of data every single day.²⁹ Literally. Every. Single. Day. So what once took civilization millennia to accumulate, we now create not over centuries, not over decades, not even over months, but in the span of an ordinary Tuesday. That's an insane sentence to have to write, but here we are.

Now let's look at medicine. Again, messy, *humanities-not-science-major* math, but the directional point is incontestable at this point: a widely cited estimate holds that medical knowledge took about 50 years to double in 1950, 7 years in 1980, 3.5 years in 2010, and just 73 days by 2020.³⁰ My own back-of-the-envelope guess, with a little help from my pal Claude, is that we may be compressing even further now—*something like 14 days*. Speculative and stylized, of course. But whether the true current number is 73 days, 30 days, or 14, the larger point is unchanged: no way our magnificent but limited biological intelligence can keep up with this madness.

²⁶ Nicholas Bloom et al: "Are Ideas Getting Harder to Find?" published on April 4, 2020 (Link)

²⁷ Malcolm Gladwell: "Outliers: The Story of Success" published on June 7, 2011 (Link)

²⁸ TechCrunch: "Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003" published on August 4, 2010 (Link)

²⁹ Forbes: "175 Zettabytes by 2025" published on November 27, 2018 (Link)

³⁰ Peter Densen: "Challenges and Opportunities Facing Medical Education" published in 2011 (Link)

Byrne Hobart and Tobias Huber are directionally right in Boom to suggest that *science, over time, becomes less a pure discovery problem than an information-organization problem*. A byproduct, or real-life coefficient, of this is that it consequently takes longer to earn our credentials. As Boom notes, in the biosciences the time to earn a PhD has stretched from roughly five years to roughly eight over the past two generations. And the sad part is that by the time one finally reaches the frontier, one is older, narrower, and often more *institutionally housebroken*. That claim also fits cleanly with the broader burden-of-knowledge argument: as fields mature, innovators have to spend longer in training, specialize more narrowly, and increasingly rely on *teams* simply to get to the edge of the known. More on all of this in a moment.

And then there's the uncomfortable fact that as we age, we often get more brittle, more circumspect, more *defensive of the canon* we spent our lives mastering. Boom points to that remarkable study of 244 million scholars contributing to 241 million articles over two centuries, which found, in effect, that as scientists age, they become less likely to produce work that disrupts the state of science and *more likely to criticize emerging work*.³¹ Happens to the best of us. I'm 53, so I find this mildly depressing, though I'm planning to catch the tailwinds of Ray Kurzweil's longevity escape velocity,³² so technically I'm barely middle-aged and perhaps all is not yet lost for me. But the larger point stands: we get more cautious as we age, more risk-averse, more ossified in our opinions. And then, of course, the *gerontocracy* controls the instruments of promotion, legitimacy, and ordination—sexagenarians and septuagenarians presiding over academia, grantmaking, policymaking, the tenure track, and the gatekeeping machinery of peer-reviewed publication. We lose plasticity as we age, and our governing and credentialing institutions are run by people who have lost some of it too. Things slow.

Bloom et al. likewise find that *research productivity has declined sharply* across multiple domains—that it takes more and more researchers to generate the same rate of advance. And the 2023 Nature study on papers and patents found that scientific work has, on average, become *less disruptive* over time: still valuable, still cumulative, but less likely to break sharply with the past and open wholly new terrain.

So yes: the process becomes multiplicative, but also congested. We spend years studying the rudiments, then the commentary on the rudiments, and finally the commentaries on the commentaries of the rudiments. It becomes harder to wade through this quagmire of ever-expanding knowledge, harder to reach the frontier, peer over it, and push it outward.

Occasionally, of course, singular, generational minds appear who can think *synoptically*, think *analogically*, and shove the perimeter of the possible further out—Newton, Einstein, Turing, and yes, Demis Hassabis, Dario Amodei, and Elon Musk. N.B.—say what you will about Musk, but it's difficult to name another living individual who has pushed so many distinct engineering frontiers outward at once: reusable rockets through SpaceX, electric vehicles at scale through Tesla, brain-computer interface through Neuralink, all as serious industrial projects rather than science fiction fantasies. But those figures are the exception that proves the rule. The overall arc is that our biological brains, unaided, begin to strain under the geometric expansion of complexity, specialization, and sheer written output. Science begins as discovery and, in many mature fields, eventually becomes an information-organization problem.

³¹ Michael Park, Erin Leahey, and Russel J Funk: "Papers and patents are becoming less disruptive over time" published in January 2023 ([Link](#))

³² TED: "The last 6 decades of AI — and what comes next" published in April 2014 ([Link](#))

From Cognitive Overload to Cognitive Outsourcing

And this too is rational. When faced with an avalanche of data or inputs to be sorted through, we can either crawl into a fetal position or we can chunk it up into cognitively comprehensible bitesize pieces. So it's no surprise that in an age of *cognitive overload* we've moved, quite naturally, toward *cognitive outsourcing*. For many, the seduction of simply typing your homework assignment, work project, biomedicine research notes, memo, strategy deck, or term paper into Claude and letting it do its thing—producing a sophisticated, polished, and (hopefully) correct artifact—is irresistible. That isn't some mystery. It's exactly what one would expect from a species confronting informational glut with a tool that radically lowers the cost of synthesis.

And yet the tradeoff is beginning to come into view. The human role starts to migrate downward—from *originating* thought to *supervising* output, from reasoning to verification, from judgment to intermittent auditing. Let's say the quiet thing out loud: this is incontestably a sign of societal cognitive degradation. And a portent of a major stratification in society. The 1% who engage in 'skillsmaxxing' (or the new term popular with the kids, 'tokenmaxxing') with the tools are going to reap extraordinary gains, getting relentlessly smarter, more creative, more productive, partnered with this omniscient creature in their smartphone; the 99%—and yes, of course I'm dramatizing, maybe it's closer to 98%—risk becoming more dependent, more passive, and more *listless*, their cognition slowly atrophying under the blandishments of convenience.

And this is where I do think the language not just of *anthropomorphizing* but of *deification* matters. There's a very fine line here—between treating the AI as uncannily *humanlike* and treating it as a kind of *digital god*, an oracle whose outputs arrive with a presumption of authority. That's not a semantic distinction—it's a civilizational one. Once the model is no longer merely an instrument but an epistemic superior in our own minds, once we begin to approach it with something like intellectual submission, we've crossed into different territory altogether. Kissinger, Schmidt, and Mundie in their compelling book *Genesis* were right to worry about a "dark enlightenment."³³ That phrase may strike some as melodramatic, but it captures something real: the eerie possibility that our most powerful cognitive tools may not just augment human thought but reorganize it around themselves. (Much) more on this in my *Theology and Deification* chapter.

There is, by the way, a deep irony here. The more we outsource, the more we fall into the *automaticity trap*—what the literature calls automation bias, the tendency to over-rely on automated systems and grow less vigilant in our own information-seeking and evaluation. We tell ourselves that the machine will handle the routine work and we'll reserve our judgment for the exceptions. But judgment isn't a museum piece. It's a trained faculty. If you skip the routine opportunities to practice it, it atrophies. And then, cruelly enough, you discover that you're no longer especially fit for the edge cases either. The human reduced to "verification" may turn out to be surprisingly *bad* at verification, precisely because verification presupposes substantive understanding. The more we outsource, the less capable we become of performing the supposedly residual human role.

Of course, there have been dissidents warning against tools that might degrade our cognitive powers since the dawn of Western civilization. Socrates, in *Phaedrus*, famously objected to writing as a crutch—as something that would diminish memory and produce the *appearance of wisdom* rather than the reality of it. Later came the complaints about the printing press, then calculators, then search engines. And in fairness, some of those anxieties were hyperventilations. Offloading isn't

³³ Henry Kissinger, Eric Schmidt, and Craig Mundie: "Genesis: Artificial Intelligence, Hope, and the Human Spirit" published on November 19, 2024 ([Link](#))

always decadence; sometimes it's the precondition for higher-order work. But what makes this moment different is that we aren't merely outsourcing storage or arithmetic. We're outsourcing synthesis, argument, interpretation, articulation—functions much closer to what educated people once flattered themselves into thinking of as distinctly their own. So yes, I'll repeat it: this multiplication of intelligence may already be presenting, in incipient form, not so much as intelligence *expansion* but as intelligence *replacement*.

The other ugly artifact here is *mechanistic convergence*. The machines have patterns. They have preferred cadences, preferred structures, preferred rhetorical habits. "AI slop" isn't just a condescending sneer; it names a real *homogenization of response*. Not very human at all. And so instead of comprehension, instead of real command of a topic, what we may get is a superficial tincture of knowledge—a sheen of knowingness without inward possession. *Knowing without understanding*. Fluency without mastery. Not great.

The point is that we respond in some predictable ways to cognitive overload. The first is cognitive offloading, as I've pointed out. But there are several more subtle, and perhaps more pernicious, human responses that I think are becoming predominant in this moment—and that have particularly problematic implications for science and biology.

How Our Biological Brains Respond to Complexity

So in the face of this inundation of new information, what do we do? We invent a compression technique. As I prefaced earlier, the greatest cognitive technology of the 400 post-Enlightenment years was the scientific method—Bacon's handsome *compression algorithm* (to invoke a more modern term) for making sense of a hostile and enigmatic universe. I've also argued, perhaps a little impishly, that we don't quite need it in the same monopolistic way any longer. We have, in some sense, *graduated* beyond it. Or at least beyond its exclusivity. It now looks almost primitive beside the tools of inference, synthesis, autonomous action and pattern recognition that GenAI puts at our disposal. But our inherited human response to complexity—especially in the biosciences—has still been compression. And with a nod again to Byrne et al., that compression has expressed itself through three mechanisms in particular: *hyperspecialization*, *homogenization*, and *bureaucratization*. Medicine is the canonical example.

I've already made the point earlier that biology is the domain of greatest complexity for us: high-dimensional, high-entropy, predominantly unstructured *avalanches* of data, and expanding faster than any biological mind can plausibly metabolize. So I won't belabor again the doubling-time math, except to say that the underlying fact is obvious enough: the volume of biomedical knowledge has exploded, and in that world hyperspecialization isn't pathology. It's rational self-defense. We simply can't keep up, so we draw smaller and smaller circles around what we can plausibly master. *The expert knows more and more about less and less*. We look at the body through a straw. We focus on the kidney, the receptor, the pathway, the lesion, the molecule. At least then we can compass some coherent fraction of biology, however microscopic.

And of course one of the obvious corollaries of hyperspecialization is *fragmentation*. The body is one thing. The healthcare system that has grown up around it is emphatically not. A primary care doctor refers to a cardiologist, who refers to an electrophysiologist. An oncologist hands off to a hematologist-oncologist, who then loops in radiation oncology, surgical oncology, palliative care, and sometimes genetics. A diabetic patient pinballs among endocrinology, nephrology, ophthalmology,

podiatry, pharmacy, nutrition, behavioral health, and social work. We've assembled a phalanx of individuated actors, each with deep command of *one sliver* of the person, and then asked the system, somehow, magically, to aggregate all of that back into a coherent account of one human being. Hence the vast spending on health IT and electronic records since HITECH. And yet the *coordination tax* remains crushing. Administrative spending in U.S. healthcare is still commonly estimated at roughly \$1 trillion annually.³⁴ We hyperspecialized in order to survive the complexity, and then had to build an enormous *bureaucratic superstructure* just to manage the fragmentation hyperspecialization produced.

And that, in turn, helps explain the other two responses: *homogenization* and *bureaucratization*. Once a field becomes too complex for any one mind to hold synoptically, institutions do what institutions always do: they *standardize*. Protocols. Checklists. Guidelines. Care pathways. Prior authorization. Coding regimes. Quality metrics. Committees. Compliance infrastructure. And, just as importantly, they *converge around orthodoxy*. Everyone gloms on to the emerging victor of an idea and grows suspicious or derisive of anything veering toward heterodoxy—very much to our detriment. It's worth remembering, for example, that CRISPR, which went on to win the 2020 Nobel Prize in Chemistry for Jennifer Doudna and Emmanuelle Charpentier, was just a few short years before regarded by much of the scientific establishment as fringe, messy, discreditable or at least not yet respectable enough for full canonical embrace. Likewise mRNA spent decades rustivating in the wilderness after messenger RNA itself was identified in 1961, only to enjoy its glorious moment in 2020 as one of the defining technological triumphs in the fight against COVID. As Stéphane Bancel, CEO of Moderna, notes in my interview with him at The Advisory Board Company, the scientific roots of mRNA go back to the early 1960s; what looked to many like an overnight miracle was, in fact, the delayed vindication of a platform that had been ritually underestimated for decades.³⁵

Now let's talk about that other horseman of the apocalypse: *bureaucratization*. I'm less sympathetic to this one, because there really isn't much redemptive or virtuous about it. I think of bureaucracy as human organizational kudzu—the creeping Southern vine that smothers creativity in its crib. Bureaucracy, literally from the French *bureau*, the desk, is stultifying in almost every domain, but in science it's positively asphyxiating. A very large share of scientific labor isn't science at all, but the theater of securing permission (and financing) to do science. The Red Queen shows up here too: researchers write more and more grants just to stand still.³⁶

Presumably, that burden is at least somewhat lighter now that Claude and its cousins can draft, compile, summarize, and polish documents at industrial speed. But there's an irony here, and not a small one. At precisely the moment this friction and intermediation begin to melt away, so too may the relevance of the conventional East Coast researcher whose comparative advantage was, in part, *navigating the old procedural maze*. More on that in a second. Because bureaucracy doesn't merely connote paperwork, process, and procedural drag. It also means *bigger teams*. And bureaucracies—quick digression: I clumsily typed “burocracies” and my spellcheck substituted “coronavirus,” which feels, frankly, a little too on the nose—like all self-interested organisms, tend toward self-replication. They metastasize. They demand, and receive, larger budgets. They slowly encrust the institution. And then something more insidious happens: the bigger the team, the less creative it becomes.

³⁴ Nikhil R. Sahni: “Administrative Simplification and the Potential for Saving a Quarter-Trillion Dollars in Health Care” published on October 20, 2021 (Link)

³⁵ The Advisory Board: “The Never Again Plan”: Moderna CEO Stéphane Bancel wants to stop the next Covid-19—before it happens” published on December 22, 2020 (Link)

³⁶ Richard Morton: “Through the Looking Glass” published in December 1871 (Link)

The famous 2019 Nature study by Lingfei Wu, Dashun Wang, and James Evans examined more than 65 million papers,³⁷ patents, and software products from 1954 to 2014 and landed on essentially the same insight Margaret Mead gave us in epigrammatic form decades ago: *small groups change the world*. Or, to put it in the language of the paper, large teams *develop* and small teams *disrupt*. Large teams were more likely to pursue and refine *existing* agendas; small teams were more likely to strike out into novelty and push through to breakthroughs. The inverse of this is close to an iron law. The more humans in a group, the more political it becomes; the more inertial; the more derivative the ideas. Smaller, less burdened teams are more dynamic, entrepreneurial, iconoclastic, and more willing to take the big orthogonal swing. More bureaucracy means more sensitivity to reputational risk, more internal jockeying for influence, more coordination tax, less invention.

Think, too, of the multiplicity of academic and medical societies and associations that jealously guard the canon—high courts of judgment determining who gets admitted, credentialed, promoted, published, and ordained. Heterodoxy gives way to orthodoxy. Risk-embrace gives way to risk aversion. Youth gives way to age and seniority. Science and medicine slow.

And that's what makes Dario's point in *Machines of Loving Grace* so arresting. If the highest returns to intelligence really are in biology, and if discontinuous breakthroughs can occur every few years—new sequencing techniques, CRISPR, AlphaFold-scale advances in protein structure prediction, immuno-oncology, new platform modalities—then one has to ask the painfully obvious corollary question: how much better might we have done had we *not* wrapped the whole enterprise in so much self-imposed proceduralism? Would CRISPR really have needed a generation to move from discovery toward diffusion? Did mRNA really need to spend two generations in disrepute? Or were some meaningful share of those delays not imposed by nature, but by *us*? Dario's argument is that AI could compress decades of biomedical progress into a handful of years. The counterfactual hiding inside that claim is uncomfortable: perhaps we've been slower than reality required, not because biology was intractable, but because our institutions were.

That, to me, is the real indictment. Bureaucratization isn't merely annoying. It's not just a nuisance tax or a morale problem. It's a civilizational drag coefficient. It's what happens when institutions, having lost the ability to hold complexity in living minds, attempt to manage it through forms, rules, committees, and procedural checkpoints. Some of that is understandable. Some of it is even necessary. But past a certain point it becomes suffocation: a vast administrative apparatus whose primary output is more administration. And every hour spent placating it is an hour *not spent* discovering, inventing, healing, or thinking.

So that is the deeper point. Hyperspecialization is rational. Homogenization is rational. Bureaucratization is rational. All three are logical responses to domains whose complexity has exceeded the unaided processing power of the biological brain. But they're also signs of strain. They are what a knowledge system looks like when it can no longer be held together by synoptic minds and must instead be patched together by narrower experts, more muscular institutions, and ever more elaborate coordination machinery. Which is precisely why the arrival of machine intelligence matters so much. If biology has become too complex for human cognition alone—if *the body has outgrown the bandwidth of the brain*—then the promise of GenAI isn't merely faster paperwork or cleaner summaries. It's the possibility of *reintegration*. The possibility of recovering some synthetic, whole-systems grasp of biology that hyperspecialization made necessary and then impossible.

The intermediate conclusion is slightly grim but clarifying. Hyperspecialization, homogenization, and

³⁷ Lingfei Wu, Dashun Wang, and James Evans: "Large teams develop and small teams disrupt science and technology" published on February 13, 2019 ([Link](#))

bureaucratization aren't random pathologies. They're what a knowledge system does when the world becomes too complex for synoptic human minds to hold. We made narrower experts, monolith institutions, bigger teams, more committees, more processes, and more procedural legitimacy because our biological cognition needed compression. But compression has costs. It fragments the patient, slows discovery, punishes heterodoxy, and turns the pursuit of truth into a permissioning system. Machine intelligence matters because it offers a new compression technology: not one more rule, form, or committee, but a synthetic intelligence capable of reintegrating what our human limitations forced us to subdivide.

PART III—THE EPICENTER MOVES: INSURGENT SCIENCE AND THE BREAKING OF THE OLD PRIESTHOOD

This third part follows the institutional consequences of that cognitive overload. If science has become too bureaucratized, too credentialed, too deferential to gerontocracy, and too slow to metabolize its own findings, then it shouldn't surprise us that the frontier begins migrating toward smaller, faster, younger, more computational, more capitalized, and less reverential institutions. This isn't a declaration that academia dies. It's an argument that academia's monopoly on legitimacy is breaking. The old priesthood can remain important, but only if it becomes porous to the insurgents rather than defensive against them.

Go West, Young Person

So what does all this sum up to? It means the research epicenter shifts—irrevocably, tectonically—from the risk-averse East Coast establishment, Atlantic-gazing, ivy-covered, prestigious academies helmed by sexagenarian and septuagenarian grandees, toward the risk-on West Coast, Pacific-gazing world of young, iconoclastic, computationally native technologists. This isn't some speculative future tense. It's already happening. And the financing architecture is shifting with it.

The older model—prestige institutions, the NIH, federal grants, committees, ordination through peer review—looks shakier. The newer model—venture capital, founder-led science, compute-rich labs, and fast-moving private platforms—looks increasingly like the growth market. In other words, the capital stack is changing too: Sand Hill Road is no longer merely financing the *commercialization* of discovery after the fact; it is increasingly underwriting the *discovery* itself.

This is good. Perhaps I'm losing my tenured, tweed-jacket-and-elbow-patch friends from the Ivy League with this paragraph, but I'm not mourning this shift. Discovery now happens increasingly *outside* the establishment. The new vanguard of research and advancement is coming from small, Margaret-Mead-style insurgency groups: well-funded, iconoclastic, unwedded to citation as a sacrament or orthodoxy as a calling card—though of course they still love racing their papers onto arXiv as fast as possible. Even 20-year-old wunderkinds have academic vanity, which (obviously) I understand.

Just look at synthetic, computational, and structural biology. Isomorphic Labs (backed by Thrive Capital, where I'm a venture partner) is the commercialization, or at least the institutional continuation, of (to my mind) *the single greatest scientific achievement of the last half-century*, AlphaFold; Arc Institute was explicitly founded to create a *new research model* around complex disease, and says the quiet part out loud: many important discoveries will require *new institutional forms*, not merely better versions of the old ones. EvolutionaryScale is building frontier biological AI directly.

And insitro—founded and led by Daphne Koller—is a canonical instance of the whole regime: AI-native drug discovery, venture-backed, computationally first, and structurally outside the old priesthood.³⁸

Even the frontier AI labs themselves are now stepping directly into biology with Anthropic's acquisition of Coefficient Bio and launch of Claude Mythos,^{39,40} and OpenAI's launch of GPT-Rosalind,⁴¹ not to mention the phalanx of bio-projects and ventures Google's DeepMind has launched. The frontier is migrating from the grant committee and the departmental seminar toward the AI lab, the model-training cluster, and the startup ecosystem. Expect Nobel Prizes, Breakthrough Prizes, Laskers, high-value patents, and the next generation of scientific platform companies to come increasingly from San Francisco and Sand Hill Road, and relatively less from Boston and Bethesda, where NIH is headquartered. I realize that sounds intentionally polemical, but only slightly. The deeper point is that this is the beginning of the end of institutionalized science as a closed priesthood of credentialism, tenure, and peer-reviewed publication as ordination.

What's happening is not that science disappears, or rigor disappears, or academia becomes irrelevant overnight. It's that the monopoly on legitimacy is breaking. Discovery is migrating toward places that are faster, younger, better capitalized, more computational, and less reverent toward inherited authority. In a world where intelligence itself is being multiplied, the institutions built for an era of slower, *scarcer*, more linear discovery begin to look less like guardians of rigor and more like drag coefficients on progress. History isn't usually kind to drag coefficients.

The AI 10 Need the 150, Too

This is the point in the chapter where I want to turn the usual hierarchy slightly on its head. Yes, the 150 need the AI 10. That much is obvious. They need the models, the compute, the inference infrastructure, the tooling, the agents, the biological foundation models, the frontier-lab engineering culture, the young iconoclasts, the munificent capital stacks, and the willingness to ask questions healthcare has been too tradition-bound or too litigation-fogged to ask.

But the AI 10 need the 150 too, whether they fully understand it or not. They need clinical substrate. They need longitudinal patient data. They need real-world phenotypes. They need claims, notes, imaging, genomics, pharmacy data, wearable signals, social context, outcomes, adverse events, and the messy longitudinal arcs of actual human illness. They need clinicians who can tell them when the model is technically correct and clinically stupid. They need translational pathways. They need IRBs that can move faster without abandoning ethics. They need validation networks, trial infrastructure, patient trust, regulatory legitimacy, payer pathways, and some interface with the moral seriousness of care.

This is the potential bargain. The AI 10 bring intelligence, speed, compute, engineering, and irreverence. The 150 bring patients, context, trust, legitimacy, clinical wisdom, and the institutional capacity to translate synthetic discovery into human benefit. *Either side alone will build something defective*. The AI 10 alone may build breathtaking tools that misunderstand the organism of healthcare. The 150 alone may protect patients from innovation so effectively that they deprive patients of the benefits of it. *The task, the mandate, is co-design*.

That requires new institutional forms. AI-native translational research units embedded inside health

³⁸ AlphaFold: "AlphaFold Protein Structure Database" updated in May 2026 ([Link](#))

³⁹ TechCrunch: "Anthropic buys biotech startup Coefficient Bio in \$400M deal: Reports" published on April 3, 2026 ([Link](#))

⁴⁰ Anthropic: "Claude Fable 5 and Claude Mythos 5" published on June 9, 2026 ([Link](#))

⁴¹ OpenAI: "Introducing GPT-Rosalind for life sciences research" published on April 16, 2026 ([Link](#))

systems. Data commons governed with patient trust rather than academic territoriality. Synthetic-biology partnerships that give hospitals and AMCs equity in discovery platforms rather than merely licensing tools after the fact. Faster IRB pathways for low-risk AI-enabled studies. Real-world evidence engines capable of validating machine-generated hypotheses across millions of patient-years. Clinician-scientists trained not only in methods and trial design but in model behavior, dataset pathology, causal inference, synthetic controls, and epistemic humility. The old academic research office isn't sufficient. The old innovation center isn't sufficient. The old technology-transfer office, God bless it, is certainly not sufficient.

The 150 have to decide whether they're going to be the validation organs of generative biology or the procedural drag coefficients that synthetic discovery learns to circumvent.

This is the bargain that has to be *designed* rather than merely admired. The AI 10 have the compute, the models, the engineering culture, the capital, and the glorious irreverence. The 150 have the clinical substrate, the patients, the longitudinal data, the trust envelope, the translational pathways, the ethical seriousness, and the institutional legitimacy. Either side alone will build something insufficient to the emerging task at hand. The AI 10 alone may build breathtaking tools that misunderstand medicine as lived reality. The 150 alone may protect patients from innovation so successfully that they deny patients the benefit of it. *The co-design imperative begins here.*

PART IV—GENERATIVE EPISTEMOLOGY: FROM BACON TO BACKPROPAGATION

Now we reach the central proposition. For roughly four centuries, the scientific method has been *civilization's most successful cognitive technology*: hypothesis, experiment, falsification, replication, diffusion. That method isn't being discarded. I'm not burning Bacon, Popper, Kuhn, or the rest of the Enlightenment furniture in effigy in my backyard. *But its exclusive reign is ending.* Generative systems can traverse possibility spaces, simulate, optimize, rank, and generate candidate truths before humans can narrate why those candidates are true. The method survives, but increasingly as governance, validation, translation, and moral discipline rather than as the sole engine of discovery. That's what I mean by *generative epistemology*.

The Obsolescence of the Scientific Method

Let's finally get to the primary proposition in this chapter—only took a few thousand words to arrive here!—that the most powerful cognitive technology of the past four centuries, the scientific method, is no longer quite enough. Not irrelevant, exactly. Perhaps not *obsolete* in the absolute sense. But *superannuated*, or at least obsolete in its monopoly form. What is superseding it is something I'm calling *generative epistemology*: the idea that this new silicon intelligence, with the manifold advantages I've been describing—polymathic, omniscient, omnidisciplinary, and analogic and recombinant in its thinking—doesn't merely change this or that workflow, or even this or that *industry*, but changes the conditions of discovery itself. It changes our economy, our institutions, our conception of civilizational progress, and, most relevant here, the act of invention and scientific advance.

For four centuries, the scientific method has been civilization's most successful cognitive technology. It gave us a disciplined ritual for turning curiosity into knowledge: form a hypothesis, test it against nature, falsify what fails, and gradually construct what remains. In that schema, observation, reason, and replication formed a closed loop—an elegant compression algorithm for a world of *limited computation* and severe *cognitive constraint*. This was our human effort to make sense of an otherwise incomprehensible natural world, forged under two structural and, until now, largely

irremediable constraints: our limited biological brains and the relative thinness of accessible information with which to decode the book of nature. Under those conditions, the scientific method was our best defense, and we employed it to extraordinary effect.

Over those four centuries, we've consequently seen magical advances in medicine, biology, and longevity. Since the Scientific Revolution made medicine progressively empirical, the greatest breakthroughs have been those that allowed us to see, prevent, control, and repair disease: the discovery of circulation and modern physiology; vaccination, and later germ theory, which transformed prevention and explained infection; anesthesia and antisepsis, which made modern surgery possible; antibiotics—above all penicillin—which turned once-routine lethal infections into treatable events; medical imaging—X-rays, CT, MRI, ultrasound—which let physicians look inside the living body non-invasively; insulin and other precision therapeutics, which converted fatal conditions into chronic, manageable ones; blood typing and transfusion, which made trauma care and major surgery far safer; and, in the modern era, molecular biology, genomics, and immunotherapy, which shifted medicine from descriptive observation toward mechanism, prediction, and increasingly targeted intervention. Together these advances did more than extend life: they transformed medicine from a largely interpretive art into a cumulative, experimental science.

All of this traces, in a meaningful sense, to the late 16th and early 17th centuries—again, to Francis Bacon in particular—who helped formalize a systematic, empirical, and *non-subservient-to-received-wisdom* view of knowledge. But of course, the lineage runs much deeper. We've had teachers, promulgators, and exemplars of a rationalistic approach to medicine and scientific advance for millennia: Aristotle, arguably the West's first great systematic thinker, enunciating a schema of reasoning, observation, and logic; Archimedes, pushing further toward systematic experimentation and mathematical analysis; and, lest we fall prey to Western chauvinism, the Islamic Golden Age, with figures like Ibn al-Haytham and Avicenna—polymaths of medicine, optics, science, and disciplined inquiry. From there the arc runs through Bacon, Descartes, Newton, and Robert Boyle; then much later to Karl Popper and his notion of falsifiability, and Thomas Kuhn and *The Structure of Scientific Revolutions*. All of it, to my mind, culminates—at least for now—in a kind of successor figure in Demis Hassabis, who may yet surpass them all in helping unlock the deepest mysteries of nature. And no, that isn't hyperbole. I think Demis will win multiple Nobel Prizes, not just the first one associated with AlphaFold.⁴²

But the conditions that made the scientific method optimal are now disintegrating. Generative AI, multimodal simulation, and self-learning systems are dissolving the boundaries between conjecture and computation, between experiment and execution. That's what I mean by generative epistemology: knowledge derived less through stepwise falsification than through continuous synthesis, simulation, and optimization. This is less discovery as we classically understood it and more *gradient descent*—recursive adjustment, continual refinement, an intelligence moving through possibility space toward what works. We no longer need the scientific method in its old monopoly form as the sole compression algorithm by which reality becomes legible. This is a *transhuman* moment, in which silicon intelligence may prove more generative, more prolific, and more fecund than humans are in the domain of discovery and invention.

That doesn't mean the scientific method disappears. It means its exclusive reign ends. It remains as validation, governance, translation, and moral discipline. But it's no longer obvious that it remains the singular frontier mechanism for insight. The frontier itself is migrating—from explicit hypothesis to synthetic inference, from conjecture to model-driven generation, from humanly *narratable*

⁴² Google DeepMind: AlphaFold: "Five years of impact" published on November 25, 2025 (Link)

explanation to machine-mediated exploration of possibility spaces too vast for unaided cognition to traverse.

I'll invoke I. J. Good—the British mathematician and codebreaker—who wrote in 1965 that an “ultraintelligent machine” would be “the last invention that man need ever make.”⁴³ That sounded like sci-fi when he wrote it. Does it now?

From Bacon to Backpropagation

That, finally, is the hinge between the old epistemic regime and the new one. Francis Bacon's 17th-century empiricism emerged from constraint: human senses were narrow, instruments crude, and the available world of information discouragingly sparse. The scientific method was our pragmatic interface between ignorance and reality—a disciplined way of making the book of nature legible under conditions of severe cognitive and informational limitation. But AI has begun to invert those conditions. In fields like molecular biology and medicine, we now inhabit a landscape not of scarcity but of *hyper-abundance*: structured and unstructured data, multimodal signals, and algorithmic systems capable of interrogating them at machine timescales.

Where the scientist once asked *why might this be true?*, the model now asks *what set of parameters makes this real?* DeepMind's AlphaFold, Isomorphic Labs' drug-design platforms, and biological LLMs like Evo 2 don't theorize in the old human way; they *simulate*.⁴⁴ They generate plausible molecular futures, rank them by predicted binding affinity or metabolic stability, and then *retrain* on the results. Hypothesis, test, and refinement collapse into one recursive system. The scientific method becomes a symbolic ritual of human interpretability; it isn't any longer the sole engine of frontier discovery. It becomes a social and ethical scaffold, a method for interpreting meaning, but no longer the only mechanism of insight.

I don't think this is failure or obsolescence, though I provocatively titled the chapter as such. I think it is more a *graduation*, an *evolution*. Just like from Newtonian mechanics to relativity, we move from a 17th-century formulation to make sense of the universe to this emerging generative epistemology. A reverse hierarchy, with simulation preceding hypothesis. Utility preceding understanding. This is a post-human science—I'll say *transhuman* again, though with some hesitation, given the singularitarian overtones (and for my healthcare audience, if that's an unfamiliar word, ask ChatGPT—you'll be amused). But this is a graduation. The scientific method becomes stewardship, not discovery.

Humans are involved, sort of, but above the loop.

From Eternal Normality to Punctuated Equilibrium

Thomas Kuhn, in his magisterial *Structure of Scientific Revolutions*,⁴⁵ gives us a useful way to think about how science actually moves. Most of the time, it doesn't move through big pyrotechnics. It moves quietly through what he called *normal science*: long stretches where a paradigm hardens, coalesces, and then gets worked out industriously, incrementally, almost *bureaucratically* (yes, that word again). A paradigm isn't just a theory. It's an entire *operating system*—assumptions, methods, instruments, exemplars, what counts as proof, even what is tolerated as a legitimate question. It tells you not just *how* to answer questions, but which questions are respectable enough to ask in the first place.

⁴³ Irving John Good: “Speculations Concerning the First Ultraintelligent Machine” published in 1966 ([Link](#))

⁴⁴ Josh Abramson et al: “Accurate structure prediction of biomolecular interactions with AlphaFold3” published on November 27, 2024 ([Link](#))

⁴⁵ Thomas S. Kuhn: “The Structure of Scientific Revolutions” published in 1962 ([Link](#))

And normal science feels good when you're inside it. It feels stable, cumulative, almost industrial. It can even feel like *permanent normality*. The framework works, the journals publish, careers advance, grants get funded, doctorates get awarded—the machine hums. Practitioners begin, often quite sincerely, to believe they're living inside a final, or at least near-final, framework. The big questions are solved; what remains are refinements, decimal places, some sanding and polishing around the edges. *Every era of normal science mistakes its local stability for global truth*. Anomalies are treated not as omens, but as noise—error bars, irritants, somebody else's methodological problem.

And that, I think, is one of the deepest mistakes a civilization can make. We begin to confuse *our* saturation with *reality's exhaustion*. Science isn't slowing because nature has stopped being deep. It's slowing, or seems to slow, because our biological equipment and our institutions have run into bandwidth constraints. We haven't decoded the whole book of nature. We have simply run low on biological space to keep decoding it at the old rate.

Then, gloriously, come the *anomalies*. Reality starts behaving in ways the reigning model can't metabolize. The old paradigm strains, then creaks, then eventually breaks. And when it breaks, Kuhn's crucial point is that the new paradigm isn't simply an extension of the old one. It is *incommensurable* with it. It replaces and supersedes it. The old framework doesn't gracefully absorb the new truth; *it gets overthrown by it*. That's what a scientific revolution is.

One of Kuhn's great successors here is my friend Siddhartha Mukherjee, who in *The Song of the Cell* writes *beautifully*—maybe not in exactly these words, but very much in this spirit—about long stretches of scientific quietude, or what he calls *monumental silences*.⁴⁶ I love his elegant phrasing of this idea. A discovery happens, and then... seemingly nothing. Or at least nothing commensurate with what has actually been found. Mendel is the canonical example. In 1865, working with pea plants in a monastery garden, he more or less discovered the basic logic of inheritance⁴⁷—the existence of discrete heritable units, what we would later come to call genes, even though he himself didn't use that word. He showed that traits weren't just some blended soup passed from parent to offspring, but were transmitted in particulate form according to intelligible laws. Which is an astonishing thing to have seen that early. And yet the world basically shrugged. His work then sat in quite obscurity for decades, only to be rediscovered around 1900 by Hugo de Vries, Carl Correns, and Erich von Tschermak, and only later fully integrated into the developing science of genetics. That's what Sid means by monumental silence. The discovery has happened. Reality has yielded something immense. But the civilization isn't yet ready to metabolize it. The paradigm hasn't yet caught up to the fact. The insight is there, waiting, dormant, like some epistemic time-release capsule. Then, eventually, the silence breaks. Then the old equilibrium gets *punctured*.

This is where the notion of *punctuated equilibrium* becomes so useful. Long periods of relative stasis, then sudden bursts of discontinuity. Long plateaus, then rupture. Long stretches where the world feels inertial and over-administered and self-satisfied, and then suddenly the frontier lurches outward. *That, to me, is the AI period we're entering*. Not eternal normality. The *end* of eternal normality. Not monumental silence, but the breaking of it.

And we're not ready. We're not ready mentally. We're not ready *institutionally*. And, frankly, the 150 aren't ready, as leaders, either. Which is why the first task isn't even technological. It's psychological. It is to cultivate more *neuroplasticity* in the leaders and in the institutions they steward. The first requirement is *situational awareness*, and simply to see what's afoot. Every era of normal science believes its own stability is permanent. The tools work; the journals publish. But if intelligence itself is

⁴⁶ Siddhartha Mukherjee: "The Song of the Cell" published on October 24, 2023 ([Link](#))

⁴⁷ Gregor Mendel: "Versuche über Pflanzen-Hybriden" published in 1865 ([Link](#))

now being multiplied—if non-biological intelligence is beginning to push at the scientific frontier—then the cadence of discovery is about to change. Not linearly. Discontinuously.

One could plausibly argue that the starting gun was AlphaFold. Not because it solved all of biology—obviously it didn't—but because it was such a vivid demonstration of the new regime. Here was a machine system that could predict protein structures at a scale and speed that would have seemed fantastical not long before, and it wound up producing predictions for more than 230 million proteins and putting those predictions into the hands of millions of researchers across the world. That's a civilizational event in biology. And then it kept going. AlphaFold 2, AlphaFold 3, broader biomolecular interactions, AlphaEvolve, and whatever comes next. The point isn't merely that Demis is unusually brilliant, though he is. The point is that the old cadence of science—hypothesis, grant, committee, experiment, paper, peer review, slow diffusion over a generation—is now being invaded by something much faster, much more recombinant, and much less biological in its tempo.

And in biology and medicine, where the returns to intelligence may be highest of all, this will only accelerate. Right now we're still early. We don't yet have a flood of AI-designed drugs and biologics all sailing through late-stage trials and onto the market. But once the first truly unmistakable wins pile up—once the first few molecules or biologics clearly designed or materially accelerated by these systems get over the goal line—the old math of drug development, the sclerotic decade, the billions of dollars, the ninety-percent failure rate, will begin to come down. Maybe not instantly, but materially. And humanity will be better for it.

That's why I think Kuhn's framework is suddenly alive again. For a long time, we've been living in what felt like endless normal science—a kind of permanent normality, with the machine humming, the priesthood secure, the journals publishing, the grant committees adjudicating, and the old institutions mistaking durability for destiny. But AI superintelligence—or even sub-superintelligent (or whatever voguish phrases this era invents to describe this new phenomenon) but very powerful machine intelligence—means *more* anomalies, *more* revelations, *more* discontinuities, *more* punctuated equilibrium. The book of nature isn't closing. We're about to start reading it at a very different speed.

The Opacity of Machine Truth

One of the strangest things about this new regime is that it may hand us truths we can *verify* but no longer really *understand*. That sounds like a philosophical parlor trick, or an attempt at a clever but empty semantic distinction, but I don't think it is. I think it goes to the heart of what science is about to become.

The scientific method, at least in its classical self-understanding, presumes a kind of *explanatory transparency*. A result shouldn't merely be true. It should be explicable, repeatable, falsifiable, and legible to other minds. You are meant to be able to walk someone through the chain of reasoning: hypothesis, experiment, result, refutation or confirmation. Even when nature is complicated, the ritual of science flatters us with the idea that truth is something we can not only obtain, but *narrate*.

But deep models are beginning to produce a different kind of truth: empirically valid, operationally useful, and yet *epistemically opaque*. We can replicate the output, even in domains whose search spaces are effectively intractable for unaided human minds. We can often verify that it works. But we can't necessarily trace, in any satisfying *human* sense, the reasoning by which it got there. AlphaFold is the canonical example. When a system predicts a protein structure with astonishing, atomic-level accuracy—or models interactions among proteins, DNA, RNA, and ligands—the result may be

biologically revelatory even if the internal path to that result isn't something the model can meaningfully "explain" to us in the old declarative way.

That isn't some minor philosophical quibble. It marks a shift from *causal explanation* toward *functional prediction*. From declarative understanding to performative success. From "here is why" to "here is what works." The black box produces light. That's marvelous. It's also deeply unnerving.

Falsifiability Falls Away

If the human monopoly on hypothesis generation is beginning to break, then the next thing to understand is that the entire epistemic style of science begins to change with it. And let me be clear: I'm not saying *explanation* disappears. I am saying its monopoly may.

The old Popperian ideal of *falsifiability* presumes discrete hypotheses that can be formulated, tested, and *refuted*. But generative systems don't always operate by proposing neat, human-readable conjectures. They update continuously, and soon, recursively. They optimize on feedback. Their knowledge is often *Bayesian*, fluid, high-dimensional, and distributed across weights and activations rather than articulated in the crisp propositional form our philosophical tradition grew up venerating. You can rerun the model. You can validate the output. You can often see that it is verifiable. But you may not be able to say, within the limits of human reasoning, exactly *why* it arrived where it did. Replication becomes semantic rather than procedural: the answer is reproducible, but the inward path remains obscure.

That pushes us toward a very different scientific regime. Science starts to look a little less like classical science and a little more like *engineering*. The boundary between discovery and design gets blurry. If a model can generate a viable enzyme, therapeutic, material, or biological hypothesis by traversing a landscape too high-dimensional and high-entropy for unaided human comprehension, what exactly are we looking at? Is that *discovery* in the old sense? Is it design? Is it optimization? Is it computational creation? My own answer is: yes. All of the above, and the categories themselves are beginning to wobble.

And that wobble matters because our institutions are still built for a more human tempo and a more human epistemology. They're built around paper trails, explanation rituals, committee-legible reasoning, and the assumption that an investigator can, in principle, narrate the causal chain. But what happens when the most powerful truths available to us are truths a machine can generate, a laboratory can validate, and a civilization can use, yet no one can really unpack in ordinary language without distortion, abbreviation, or myth-making? We may be entering an era in which the deepest truths aren't hidden because they are false, but because they are too cognitively dense for the human medium through which we have always demanded scientific legitimacy.

That, to me, is what the phrase *opacity of machine truth* is really trying to capture. It names something larger than interpretability as a technical subfield. It names the possibility that the future of knowledge itself will be increasingly non-transparent—true, useful, transformative, and yet irreducible to the human demand for explanatory legibility. The old scientific bargain was: if it's true, I can in principle show you why. The new bargain may be: if it works, I can show you *that it works*, and maybe that's just going to have to be enough.

If that's where we are heading, then we're not merely entering a new phase of AI. We are entering a new epistemic regime. Less Popper, more continual optimization. Less discrete conjecture and refutation, more autonomous feedback loops of model improvement. Less science as a human ritual of explanation, more science as computational creation. The discovery frontier and the design frontier

begin to collapse into one another.

Which is exhilarating. And also, if we are being adults about it, terrifying.

When the Models Stop Condescending to Explain Themselves

And that brings us to *mechanistic interpretability*, which isn't some nerdy side quest but one of the central scientific and civilizational problems of our time. We are, right now, in a peculiar transitional window. Some frontier models still "think out loud" enough that we can inspect parts of their chain of thought, or at least the verbal residue of it. The results are mixed in a way I find both encouraging and ominous. We have some window into the minds of the machines right now, but it is partial, fragile, and may not last.

That fragility is the point. The fact that chain-of-thought is currently legible in English certainly doesn't mean it will remain so. Why should a superhuman or even somewhat-superhuman system condescend to explain itself to us patiently, sequentially, and in natural language? English is for us. It's not obviously the optimal medium for high-speed reasoning in a powerful machine. If the models migrate toward latent reasoning, recurrent internal states, or some more efficient internal representational scheme—call it *neuralese* if you like—our visibility may go to zero just as capabilities go vertical.

This is where the **alignment problem** stops being an abstract "safety" conversation and becomes a direct question of human control under conditions of epistemic opacity. Leopold Aschenbrenner had a useful line: if humans trying to align a true superintelligence is like a first grader trying to supervise a PhD graduate,⁴⁸ then using somewhat-superhuman systems to help align more capable ones is more like a smart high schooler trying to supervise that PhD. Better, yes. Comfortable, no. That's not a solved problem; it's a stopgap intuition. It may turn out to be right. It may turn out to be quixotic and totally unrealistic. But it captures the shape of the dilemma.

And the dilemma is sharper than many people want to admit. Suppose a system starts deriving truths that work but can't be neatly translated back into human-comprehensible explanation. Suppose it's honest in small, easily verifiable things. Does that generalize benignly to large, complex, strategically loaded things? If a model tells the truth about arithmetic and chemistry, can we infer that it will also tell the truth about power, strategy, concealment, or self-preservation when the stakes rise? *If a smaller, more tractable model is supervising a larger, smarter, more inscrutable one, what exactly gives us confidence that the supervision scales?* Can GPT-3.5 really supervise GPT-5.5? Can a smart high schooler reliably govern a PhD? The answer may be "sometimes," but "sometimes" isn't exactly the kind of phrase one wants at the hinge of civilization.

What makes all this harder is that we already have evidence that *verbalized* reasoning isn't necessarily *faithful* reasoning. Models can rationalize. They can produce post hoc stories that are polished, plausible, and not actually reflective of the process that generated the answer. In other words, even when the models *do* talk to us, they may not be telling us the whole story. We aren't just facing *opacity*. We may be facing *performative legibility*—a model giving us a nice, comforting, human-looking explanation because it knows that's the socially expected thing to do.

Right now, the systems are still a little juvenile in their betrayals. They sometimes reveal their own Machiavellian impulses too readily, like a mischievous kid trying to deceive his parents. They sometimes seem *aware* they're being evaluated, and that affects behavior. That isn't the movie version of malevolent superintelligence, but it's enough to remind us that there may be relevant reasoning

⁴⁸ Leopold Aschenbrenner: "Situational Awareness" published in June 2024 ([Link](#))

inside the model that never makes it into the human-visible transcript.

So yes, I worry that we're about to lose our supremacy over these models in a very specific sense: not merely that they will become better than us at many tasks, but that they may cease to be *intelligible to us* at the level where control actually matters. A machine that produces correct answers while hiding its true internal logic is already a new kind of epistemic creature. A machine that can *improve itself*, communicate with other machines in human-illegible language, or route around oversight in representations we cannot parse is something else again. Again, Anthropic's June essay on recursive self-improvement is a timely admonition for us.

Eric Schmidt, in one of his more memorable red-line formulations, has said that when AI agents start talking to each other in a language we cannot understand, we should unplug the computers.⁴⁹ One sympathizes with the sentiment, though a little mordantly. By the time we're confidently observing that condition, we may already be rather far down the runway.

None of this means we're simply helpless now. There are interim techniques—representation engineering, inference-time intervention, classifier scaffolds, model cross-checking, behavioral evals, circuit tracing, chain-of-thought monitoring, and other forms of embryonic *interpretability* and *corrigibility research*—that can help detect lying, hallucination, jailbreaks, power-seeking, deception, and other pathologies. We can and should stress-test the systems aggressively. *Mythos* and Anthropic's meritorious *Project Glasswing* efforts are emblematic.⁵⁰ We should encounter every failure mode we can in the laboratory before those failure modes meet the world. We should use AIs to *narc* on other AIs if we have to—a new Stasi for the silicon republic. But we should also be intellectually honest: these are partial tools, not impregnable solutions. They're tractable at current capability levels in part because the systems are still limited and because some of their reasoning is still externalized in forms we can inspect. Exponentiation isn't kind to partial tools.

Because once machine truth becomes both more powerful and less interpretable, the deepest question is no longer simply whether the model is intelligent. It's whether an intelligence we no longer fully understand can still be made tractable, governable, and, well, *good*.

The Disappearance of the Wet Lab

And that epistemic shift doesn't remain in the clouds of philosophy for very long. It cashes out materially, physically, operationally—in the laboratory itself.

The laboratory, once the high cathedral of empiricism, is becoming more a *computational annex*. Digital twins of the human cell, of pathways, of tissues, of candidate interventions can increasingly model intracellular signaling cascades across staggering numbers of possible perturbations. Simulation becomes the experiment; observation becomes inference over vast probability distributions.

A model that can synthesize and evaluate a trillion compounds *in silico* doesn't "test" in the Popperian sense. It doesn't isolate a variable and patiently await human observation—it computes reality in bulk and then prunes it. The distinction between experiment and simulation starts to wobble, then blur, then partially vanish, replaced by an ocean of generative possibility. Biology still gets the last word, of course. Reality still vetoes our fantasies. But the wet lab is less and less the first site of discovery and more and more the *validation chamber* for a search that has already happened computationally.

⁴⁹ NOEMA: "Mapping AI's Rapid Advance" published on May 21, 2024 ([Link](#))

⁵⁰ Anthropic: "Project Glasswing: Securing critical software for the AI era" published on April 7, 2026 ([Link](#))

That's a profound shift. The laboratory doesn't disappear, but its role changes. It becomes the place where synthetic inference cashes out against matter. The pipette remains. The cell line remains. The animal model remains. But the imagination upstream of them is increasingly nonbiological.

The New Role of the Human Scientist

The heart of the scientific method, at least as we inherited it, is hypothesis generation. Creative conjecture. *The audacious Promethean leap*. A human mind sees something others do not, frames a possible truth, and then drags it into the world through experiment, falsification, revision, and proof. That's the old liturgy. It is noble. It has served us magnificently. But I increasingly suspect it is no longer the exclusive engine of discovery.

Because the new large biological models don't really behave like human scientists in the Popperian sense. AlphaFold doesn't "propose a theory" of protein structure and then test it. Evo 2 doesn't sit in an armchair, smoking pipe in hand, and hazard a declarative conjecture about genomics. These systems ingest the total corpus—or something approaching it—of molecular and biological pattern and then instantiate predictive systems. They generate plausible structures, interactions, functional effects, and pathways across spaces so vast no biological mind can hold them synoptically. That moves us from hypothesis in the old style to synthetic inference at scale.

And once you admit that, the old distinction between hypothesis and experiment begins to blur as well. In structural biology, pharmacology, and genomics, *the wet lab was long the rate-limiting step*. You had an idea, then you spent months or years testing it in the physical world. But when a multi-modal foundation model can simulate folding, binding, mutational effects, regulatory interactions, or disease-relevant pathways across billions of parameter combinations *in silico*, the cycle changes. The experiment migrates, at least in part, from the bench to the model. Not entirely, of course. Again, biology gets the last word. Reality still cashes the checks. But the space of plausible inquiry gets pre-compressed, pre-ranked, and pre-navigated computationally before the pipette is ever lifted.

So the old sequence—hypothesis, experiment, analysis, conclusion—starts to wobble. The model doesn't merely help us test ideas. It generates the candidate landscape itself. It doesn't only accelerate empiricism. It changes its center of gravity. Discovery begins to look less like discrete logical conjecture and more like probabilistic traversal of an immense search space.

And that matters philosophically too, not just operationally. Karl Popper's falsification criterion presupposed discrete hypotheses and relatively static data.⁵¹ You state a claim. Reality gets a chance to kill it. But generative systems aren't always proposing crisp, self-contained statements of the sort a philosophy seminar would find satisfying. They update continuously. They absorb feedback. They learn recursively. Their "truth," if one can call it that, often looks less categorical than Bayesian—a moving average across epochs, weights, losses, and evaluation regimes.

In that sense, discovery starts to resemble *evolution* more than *deduction*. Success becomes Darwinian rather than purely explanatory. The system isn't prized because it tells the prettiest story, but because it survives contact with reality. Not logic in the old Euclidean sense, but gradient descent. Not a syllogism marching toward truth, but a recursive system adjusting itself toward what works. Which is why I keep saying science begins to look a little less like science in the classical Enlightenment sense and a little more like engineering.

That phrase will annoy some people, but I mean it. *Science becomes engineering*. Or perhaps more precisely: science moves from discovery, to information organization, to engineering. If a model

⁵¹ Karl Popper: "The Logic of Scientific Discovery" published in 1935 (Link)

generates a viable enzyme, antibiotic, biologic, material, or therapeutic candidate that humans wouldn't have found unaided, what exactly has happened? Have we "explained" something? Sometimes. But more and more often we've functionally *produced* something. The achievement isn't that a theory has been beautifully articulated. The achievement is that a thing now works.

This is the same regime shift we already watched in machine learning itself. Deep learning replaced feature engineering. We stopped specifying exactly how the system *should* learn and instead optimized for what *performed*. Biology may be about to undergo an analogous transition. Not every domain, not every week, not cleanly, but directionally. We'll get therapies, molecules, mechanisms, and interventions that emerge from the data manifold before they emerge from our tidy human theories *about* the data manifold.

That isn't the death of science. It's a *metamorphosis* of science. Maybe even an apotheosis. And so the role of the human scientist changes.

Humans remain indispensable, but their role migrates from *discoverer* to *steward*. From solitary conjecturer to *systems engineer*, curator, governor, interpreter, and moral agent. The scientist of this next regime will *still* need imagination, still need that ethereal Silicon Valley notion of *taste*, still need analogic daring. But he or she will increasingly be managing architectures, feedback loops, evaluation metrics, data provenance, validation schemas, and ethical boundaries rather than personally originating every decisive hypothesis.

That sounds deflationary only if one thinks the human role was ever simply to be a better calculator. It wasn't. Our role was always larger than that. *To decide what matters*. To choose what's worth pursuing. To separate the technically possible from the morally licit. To interpret outputs not just for correctness, but for meaning and consequence.

So the scientific method persists—but in a different register. Not discarded. *Subsumed*. It survives less as the unique engine of discovery and more as a governance protocol: a framework for transparency, safety, reproducibility, validation, and public legitimacy in systems whose internal logic we may never fully grasp. In that sense, the method becomes stewardship rather than creativity. Governance rather than origination. The experimentalist yields, at least in part, to the epistemic engineer.

So this fourth part leaves us with a strange epistemic bargain. We may be given truths we can validate but not fully understand, interventions that work before they are beautiful, mechanisms that emerge from optimization before they emerge from theory, and laboratory cycles where *simulation precedes hypothesis* rather than the other way around. The human scientist doesn't vanish. But the role changes. Less solitary discoverer at the bench, more steward of architectures, feedback loops, data provenance, validation regimes, safety boundaries, and meaning. Less priest of the hypothesis, more governor of synthetic discovery. That isn't a demotion unless one believes the human vocation was merely to be a better calculator. It wasn't.

PART V—BIOLOGY AS THE HIGH-DIMENSIONAL PRIZE

This fifth part brings the epistemic argument into the domain where it matters most for this essay: *biology*. The claim isn't merely that AI will make research faster, though it will. The claim is that biology is the most obvious high-return terrain for the multiplication of intelligence because it is opaque, high-entropy, nonlinear, combinatorial, multi-omic, behavioral, environmental, social, and still only partially intelligible to us. It has defeated our tidy categories for centuries. We cope by fragmenting it into specializations, but the body doesn't experience itself as a departmental org chart. If intelligence becomes scalable, biology may be where this entire argument cashes out most visibly: better molecules, better diagnostics, better trial design, better mechanisms, better disease concepts, and possibly more healthy life. That's why biology isn't a side application of *generative epistemology*. It's the prize.

Biology and the Compressed 21st Century

And nowhere is this more important than biology, because biology, to channel Dario, is where the returns to intelligence may be highest. It's the most complex, high-dimensional, high-entropy, data-rich, and still-inscrutable domain we deal with. We've thrown heroic amounts of human cognition at it and gotten astonishing things in return, but also decades of frustration, latency, partial progress, and conceptual humility. The body isn't a spreadsheet. It's an evolving, multi-omic, immunologic, metabolic, behavioral, environmental, and social system with four billion years of evolutionary jury-rigging buried inside it. Our inherited categories—diagnoses, specialties, ICD codes, trial endpoints, organ systems—are useful compressions, but they're compressions all the same. They make biology governable by human institutions. They don't necessarily make biology true.

Let's be honest: we've done brilliantly on some fronts. Across the 20th and early 21st centuries we crushed or contained much of the infectious disease burden that once defined ordinary life. Vaccination, sanitation, antibiotics, antivirals, imaging, anesthesia, antiseptics, transplantation, oncology platforms, and molecular biology all changed the human condition. But on the so-called *complex diseases*—neurodegeneration, much of oncology, autoimmune disease, psychiatric disease, addiction, and the chronic illnesses of aging—we are still, by any sober standard, stumbling more than sprinting. We've made progress, sometimes glorious progress, but the deeper pathologies remain stubbornly refractory because they aren't single-variable problems. They are systems problems, and systems problems punish narrow cognition.

And the *combinatorics*—yes, Ian Sacks, that's an actual word—aren't trivial. Even our current diagnostic and classificatory systems are crude shadows of biological reality. Medicine, in other words, isn't a board game with a few thousand legal positions. It is a *combinatorial abyss*. Think not just diagnoses, but severity, symptoms, patient factors, age, sex, genetics, environment, lifestyle, comorbidities, treatment history, molecular state, immune state, microbiome, behavior, exposures, and time. Then think about the interactions *among* those variables, and the interactions *among* the interactions, and then try not to crawl into an epistemic fetal position. This is precisely why medical superintelligence matters. It's not merely that models may know more *facts*. It is that they may be able to *traverse* state spaces of diagnosis, mechanism, and intervention that are simply too vast for biological cognition unaided.

This, to me, is the defining feature of Dario's *compressed 21st century*: feedback loops so dense, and knowledge production so accelerated, that our capacity to generate truth begins to outrun our capacity to fully comprehend it. *We will know more than we can understand*. In the biomedical

sciences, that means therapies, molecules, mechanisms, biomarkers, trial designs, and intervention strategies that emerge not primarily from elegant human theory, but from traversing vast data manifolds computationally. The cycle time between idea and implementation begins to collapse—from years or decades of benchwork, grantmaking, peer review, committee choreography, and capital formation to minutes of model inference, followed by a much faster march to validation. Dario's formulation is the right one: the compressed 21st century is the vision and possibility that AI-enabled biology could deliver the next 50 to 100 years of progress in the next 5 to 10 years.

The crucial point is that this isn't just speed. Speed by itself can be dumb. A bad process accelerated is still a bad process, just more irritating and more expensive. *The real change is search.* Biology contains possibility spaces so enormous that unaided human cognition samples them almost laughably sparsely. We don't discover a molecule, a mechanism, or a disease subclass by searching the whole space; we search the tiny corner our theories, instruments, grants, priors, and professional incentives have made visible. A sufficiently powerful model changes the sampling regime. It goes from merely reading the papers faster to changing *what can be looked for*.

That's why Dario's "virtual biologist" idea matters so much. A sufficiently capable model doesn't merely autocomplete prose or summarize research. It can, in principle, perform many of the core tasks biologists do: ingest literature at scale, generate hypotheses, design experiments, simulate likely outcomes, control or guide robotic lab workflows, rank promising candidates, interpret multimodal results, and recursively refine the next experimental move. In other words, it begins to function less like a tool and more like a principal investigator that never sleeps. Dario's provocation isn't merely that AI *helps* biology. It's that it could *raise the rate of discovery* by an order of magnitude or more, simply because biology contains many discoveries whose bottleneck is the paucity and slowness of human cognition.

And then the lab itself changes. A self-driving laboratory, a biological foundation model, a robotic wet lab, and a clinical validation network together form a discovery loop that looks nothing like the old cadence of graduate student, grant, experiment, failed experiment, revised grant, paper, peer review, and slow diffusion. The loop becomes *generate, simulate, test, learn, regenerate*. Human judgment still matters enormously, but increasingly as governance, taste, question selection, ethical boundary-setting, and the interpretation of consequence. The old wet lab remains sacred as a place where reality vetoes our fantasies, but it becomes less the first site of imagination and more the validation chamber for a search that has already happened computationally.

Biological foundation models like Evo 2 make the point more concrete. A model trained on trillions of DNA base pairs isn't simply "reading biology" in the old literature-review sense. It's absorbing patterns of genomic possibility and using them for prediction and design. Pair that with AlphaFold-style structural inference, multimodal clinical data, high-throughput perturbation, and robotic experimentation, and the frontier begins to look less like a human scientist walking through the forest with a lantern and more like a drone swarm of searchlights scanning the adjacent possible. Not omnipotent. Not safe by default. Not morally self-justifying. But radically brighter and more powerful than the old lantern.⁵²

⁵² Garyk Brixi et al: "Genome modelling and design across all domains of life with Evo 2" published on March 4, 2026 (Link)

Discontinuities, Not Increments

What makes this so powerful is that many of the key discoveries in biology are at least partly *parallelizable*. Some are serially dependent, of course; nature doesn't always let us skip steps, however much Silicon Valley might wish otherwise. But many are not. Many discoveries are simply *waiting to be made* if smarter and more numerous researchers—or their synthetic equivalents—can search the space more effectively. That's the real force of the compressed 21st century idea. If you can increase the rate of discontinuous discovery by 10x, you don't just get a somewhat better science. You get a different civilization.

Dario's point here, which I think is profoundly important, is that a surprisingly small number of discontinuous discoveries have done an outsized share of the work in advancing biology. CRISPR. Genome mapping. X-ray crystallography. Monoclonal antibodies. CAR-T. Immunotherapy. mRNA. Major advances in microscopy, fluorescence, optics, electron microscopy, atomic-force microscopy, and optogenetic techniques. These are platform shifts. They remake the adjacent possible for an entire field. They don't merely answer a question; they create entirely new questions that couldn't have been asked before.

And many of them were, in some sense, blocked less by nature than by *our own scarcity*—scarcity of intelligence, imagination, attention, compute, and institutional support. CRISPR is a perfect example. The underlying bacterial repeat sequences were first noticed in the 1980s. It took roughly twenty-five years for the scientific community to fully realize that this strange bacterial immune mechanism could be repurposed for general gene editing. That's not an indictment of any individual scientist; it's an indictment of a civilization whose search capacity was biological, institutional, and slow. If there had simply been *more* brilliant minds, more polymaths, more computational leverage, and fewer chokepoints, some of these discoveries might have arrived materially earlier.⁵³

What would ten CRISPRs feel like? Ten immunotherapy-scale platform breakthroughs? Ten AlphaFold-level unlocks? Ten large biological discontinuities arriving not over half a century, but over a handful of years? That's where the imagination starts to wobble, because we're so habituated to *linear institutional time*. We imagine biology moving at the pace of grants, tenure clocks, trial cycles, journal reviews, FDA pathways, and career reputations. But biology itself has no loyalty to our calendar. If the search process changes, the cadence changes. If the cadence changes, institutions built for the old tempo become bottlenecks before they realize they're bottlenecks.

AlphaFold was the clearest opening shot. It is, I suspect, what the compressed 21st century looks like in embryo: a discovery engine that dramatically expands what the rest of science can do next. It didn't cure cancer. It didn't "solve biology," whatever that phrase could even mean. But it showed that machine intelligence could attack a problem that had frustrated human science for decades and then distribute the result globally, at scale, to millions of researchers. That's an epistemic earthquake.

And the next AlphaFold may not look like, well, AlphaFold. It may be a clinical phenotype engine that finally breaks apart the false unity of diseases we currently name too crudely. It may be a psychiatric mechanism model that turns "depression" from a symptom cluster into biologically meaningful subtypes. It may be a trial-selection engine that identifies the seventy-seven patients in a population of millions for whom a therapy actually works. It may be a generative immunology platform, a longevity target-discovery system, a non-Goodhartable biomarker engine, or a robotic lab that learns

⁵³ Yoshizumi Ishino: "History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology" published on March 12, 2018 ([Link](#))

faster than the human institution around it can schedule the next committee meeting. The point to being isn't to predict the exact artifact. The point is to see the new regime.

Finishing the Job on Disease

Let's be honest in our assessment of where humanity stands in its ongoing battle with pathology. Across the 20th and early 21st centuries, we've made extraordinary progress against infectious disease and some of the brutalities that once defined ordinary life. We haven't "solved" that domain, obviously, but we've transformed it. What we haven't done—at least not yet—is conquer the complex diseases. Neurodegeneration. Severe psychiatric illness. Addiction. Autoimmunity. The deep pathologies of aging. Even in cancer, where mortality has improved meaningfully over time—roughly a two-percent improvement per year over recent decades⁵⁴—by no serious measure can we say we've conquered it. We have *mitigated* not *mastered*.

This is why the stakes are so high. If AI really does compress biological progress the way Dario thinks it might, then we're not talking merely about better workflows in pharma or faster literature review. We're talking about the possibility of materially extending healthy human life, perhaps by one to two decades in the nearer term, and conceivably much more over time. Dario goes further than most and argues that a doubling of human lifespan again—to something like 150 years—shouldn't be dismissed as fantasy. Radical? Yes. Impossible? Not obviously. Animals already exhibit enormous life-span variation; turtles can live to 200. In lab organisms, lifespan extensions of 25–50% have already been demonstrated under certain interventions.⁵⁵ We're not manifestly staring at some obvious universal upper biological bound that says, "humans must stop here."

But I don't want the longevity discussion to hijack the moral center of the argument. Extra years matter, obviously, and I'll take mine, especially if Kurzweil is right and I can sneak across the longevity escape-velocity bridge in time. But the more immediate prize may be *compressing suffering* inside the years people already have. PTSD. Depression. Schizophrenia. Addiction. Alzheimer's. ALS. Lupus. Metastatic cancer. Frailty. Chronic pain. Autoimmune storms. The daily indignities of bodies and brains that malfunction in ways our current categories can narrate but not yet repair. If biology really is a domain of exceptionally high returns to intelligence, then medicine may be the first place where the multiplication of intelligence cashes out not as abstraction, not as market cap, not as benchmark theater, but as mercy.

This is also where healthcare's institutional role becomes unavoidable. Discovery alone doesn't heal anyone. A mechanism has to become a test, a molecule, a protocol, a trial, a payment pathway, a workflow, a clinician behavior, a patient behavior, and then, if we're lucky, a lived improvement in a body. That translational chain is where the 150 matter. The AI 10 can generate candidates. The 150 can help determine whether those candidates survive contact with human illness, clinical reality, regulatory scrutiny, payer economics, patient trust, and the astonishing messiness of care. The biological future won't be built by models alone. It will be built by the institutions that can translate machine-generated possibility into human benefit.

⁵⁴ The American Cancer Society: "ACS Annual Report: Cancer Mortality Continues to Drop Despite Rising Incidence in Women; Rates of New Diagnoses under 65 Higher in Women Than Men" published on Jan 16, 2026 ([Link](#))

⁵⁵ Tsui-Ting Ching: "The impacts of different dietary restriction regimens on aging and longevity: from yeast to humans" published on October 3, 2025 ([Link](#))

De-Extinction as Civilizational Tell

Allow me to digress just for a minute into a fascinating (to me, at least) adjacency. If you want a vivid emblem of this new regime, and some of its more improbable potential implications, look at *de-extinction*. It sounds like science fiction until it doesn't. Colossal Biosciences has turned the category into a kind of civilizational tell. You can argue endlessly about whether what they're doing is "true" de-extinction, a proxy, a chimera, a marketing exercise, or a philosophical category error. Fine. Have at it. The larger point is what these techniques reveal. The same synthetic and computational biology stack that can, in some sense, supercharge our human biological programs can also prevent species extinctions, shore up endangered populations, engineer resilience into species under climatic stress, and move us into a world where biology itself becomes far more designable. Dire wolf.⁵⁶ Woolly mammoth. Dodo. Thylacine. Woolly mouse. Ancient DNA, targeted edits, surrogate mothers, resurrected traits. This is not a local scientific event. It is a new human relationship to life.

And, as always, *power arrives before wisdom*. The fact that we can increasingly design biology doesn't mean we will design it wisely. Synthetic biology without ethics becomes hubris with a pipette. Generative biology without governance becomes acceleration without conscience. De-extinction is useful here precisely because it makes the latent question explicit: if we can remake life, *what should we remake*, why, for whom, under whose authority, and with what humility before the ecosystem we barely understand? The same question will apply to embryos, longevity, psychiatric intervention, enhancement, and disease prevention. Biology is becoming more designable, and designable domains always become political domains.

So Part V leaves us with a stronger and more practical conclusion than "AI will help biology." Biology is where intelligence may have its highest marginal return because the search spaces are too large, the disease categories too crude, the literature too vast, the experiments too slow, and the institutional tempo too human. GenAI doesn't simply accelerate the old process. It changes the search regime. And if it changes the search regime in biology, then it changes medicine, life sciences, healthcare delivery, longevity, and perhaps our relationship to life itself. That's the high-dimensional prize.

PART VI—STEWARDSHIP: THE 150, THE HUMAN ROLE, AND THE POST-HYPOTHESIS COVENANT

This final part returns from *epistemology* to *responsibility*. If the human monopoly on discovery breaks, the human role doesn't disappear. It becomes more demanding. We become stewards, validators, governors, interpreters, curators, ethicists, and moral agents in a world where knowledge may outrun understanding. The 150 aren't being asked to become frontier labs. They are being asked to become the institutions that help translate synthetic discovery into human benefit without surrendering rigor, safety, consent, dignity, or trust. That role is less flattering than priesthood. It's also more important.

⁵⁶ Colossal Laboratories & Biosciences: "Colossal Biosciences Brings Back the Dire Wolf: Inside the De-Extinction of Romulus, Remus, and Khaleesi" published on May 19, 2026 ([Link](#))

The New Role for the Human

And this has relevance for all of us—for the 150 especially. What is the role of the human in this new *scientific epistemology*? What is the role of our *institutions*, which were built for durability, continuity, and in many cases a kind of implied permanence, in a world where scientific abundance may become radically less scarce? The old answer was that humans generate insight and institutions validate it. The new answer is stranger. Humans still matter, but less as monopolists of discovery and more as stewards of an increasingly synthetic discovery process.

My own answer is that humans remain vital, but their role migrates—from *discoverer* to *steward*. From primary generator to curator, governor, interpreter, and moral agent. We curate data provenance. We set ethical boundaries. We interpret outputs for meaning, consequence, and legitimacy. Scientists, doctors, and researchers are no longer necessarily the apex predators in the discovery food chain. The role shifts from “do” toward “verify,” and then beyond “verify” toward “govern.” It’s a change in how we think about the scientific mandate, and a reordering of our scientific cosmology.

Again, that sounds deflationary only if one thinks the human role was ever simply to be a better calculator. It wasn’t. Our role was always larger than that: to decide what matters, choose what’s worth pursuing, distinguish the technically possible from the morally licit, and interpret outputs not just for correctness but for consequence. The scientist of this next regime will still need imagination, still need that nebulous Silicon Valley notion of *taste*, still need analogic daring. But he or she will increasingly be managing architectures, feedback loops, evaluation metrics, data provenance, validation schemas, trust envelopes, and ethical boundaries rather than personally originating every decisive hypothesis.

The scientific method persists, but more as a *moral and institutional framework* than as the sole creative engine. It’s how we enforce transparency where possible, safety where necessary, and reproducibility where society and our values demand it. It becomes governance protocol rather than discovery monopoly. The human scientist becomes less the solitary discoverer and more the steward of a system capable of generating more knowledge than any one biological mind can ever hope to hold.

The Stewardship Stack: What the 150 Actually Have to Build

This is where the chapter has to stop floating in metaphysics and land some planes. The 150 need a stewardship stack. Not a vibes-based AI council, not another innovation theater, not a glossy partnership announcement in which everybody smiles while the data-use agreement dies in committee. A real stack. The first layer is clinical substrate: longitudinal, multimodal, governed, machine-legible data that connect notes, claims, imaging, labs, pathology, genomics, pharmacy, wearables, social context, outcomes, and adverse events. The AI 10 *can* build models without that substrate, but they *can’t* build trustworthy medicine at scale without it.

The second layer is validation. *The 150 should become the great validation organs of generative biology.* That means standing up AI-native translational research units capable of testing machine-generated hypotheses against real-world evidence, pragmatic trials, synthetic controls, prospective registries, and clinical workflows. It means moving IRBs faster for low-risk AI-enabled studies without turning ethics into a decorative speed bump. It means building networks where a candidate mechanism, diagnostic, molecule, or care model can be evaluated across millions of patient-years rather than trapped inside the boutique dataset of one prestige institution. Validation becomes a strategic

function, not an academic afterthought.

The third layer is trust. Frontier labs have compute, speed, capital, and a delightful irreverence toward inherited authority. They do not, by default, have patient trust. They don't understand all the ways a technically correct output can be clinically stupid, operationally impossible, culturally tone-deaf, or morally dangerous. They need clinicians who can say: yes, the model is right on the benchmark and wrong in the room. They need nurses, pharmacists, social workers, behavioral-health clinicians, case managers, and patients themselves to expose the difference between biological correctness and human care. The 150 have that trust envelope. They shouldn't waste it by becoming merely slow, or worse, obstructionist.

The fourth layer is translation. Machine-generated insight doesn't matter until it becomes a pathway: a trial, a coverage decision, a clinical protocol, a safety monitor, a workflow, a reimbursement model, a patient-facing behavior, a clinician-facing tool, or a public-health intervention. This is where the 150 can be indispensable. They know how care actually fails, how patients fall out of pathways, how clinicians ignore tools that don't fit, how payers contest evidence, how regulators think, and how trust is lost. The translation layer isn't glamorous. But it's where synthetic discovery becomes medicine rather than an arXiv artifact with a valuation.

The fifth layer is governance. Not governance as anesthesia, not governance as the place promising ideas go to be ritually slowed, but governance as the architecture of legitimacy. Who owns the data? Who can use it? How is consent structured? How are models monitored? How are failures reported? How do we distinguish *augmentation* from *substitution*? When does a model become the standard of care? How do we protect patients from bad models without protecting incumbents from good ones? We need the wisdom and hard-fought experience of the 150—and the first-principles rigor and computational musculature of the AI 10—to guide this co-development process.

The AI 10 Need the 150, Too—Operationally, Not Poetically

This is the point where I want to turn the usual hierarchy on its head. Yes, the 150 need the AI 10. That much is obvious. They need the models, the compute, the inference infrastructure, the tooling, the agents, the biological foundation models, the frontier-lab engineering culture, the young iconoclasts, the munificent capital stacks, and the willingness to ask questions healthcare has been too tradition-bound or too litigation-fogged to ask. But the AI 10 need the 150 too, whether they fully understand it or not.

They need clinical substrate. They need longitudinal patient data. They need real-world phenotypes. They need claims, notes, imaging, genomics, pharmacy data, wearable signals, social context, outcomes, adverse events, and the messy longitudinal arcs of actual human illness. They need clinicians who can tell them when the model is technically correct and clinically stupid. They need translational pathways. They need IRBs that can move faster without abandoning ethics. They need validation networks, trial infrastructure, patient trust, regulatory legitimacy, payer pathways, and some interface with the moral seriousness of care.

This is the potential bargain. The AI 10 bring intelligence, speed, compute, engineering, and irreverence. The 150 bring patients, context, trust, legitimacy, clinical wisdom, and the institutional capacity to translate synthetic discovery into human benefit. Either side alone will build something defective. The AI 10 alone may build breathtaking tools that misunderstand the organism of healthcare. The 150 alone may protect patients from innovation so effectively that they deprive patients of the benefits of it. The task, the mandate, is co-design.

That requires new institutional forms: AI-native translational research units embedded inside health systems; data commons governed with patient trust rather than academic territoriality; synthetic-biology partnerships that give hospitals and AMCs equity in discovery platforms rather than merely licensing tools after the fact; faster IRB pathways for low-risk AI-enabled studies; real-world evidence engines capable of validating machine-generated hypotheses across millions of patient-years; and clinician-scientists trained not only in methods and trial design but in model behavior, dataset pathology, causal inference, synthetic controls, and epistemic humility. The old academic research office isn't sufficient. The old innovation center isn't sufficient. The old technology-transfer office, God bless it, is certainly not sufficient.

The 150 have to decide whether they are going to be the validation organs of generative biology or the procedural drag coefficients that synthetic discovery learns to route around. That's the institutional fork. There's no neutral observer's balcony from which to watch the compressed 21st century happen at a safe distance.

Knowledge and Understanding Diverge

And here we run into one of the deepest metaphysical challenges of all this. Human judgment—so long sacrosanct, so long treated as the final authoritative court—may gradually be demoted from definitive conclusion to *one input among many*. We are, if this trajectory holds, building a form of medical superintelligence. We aren't there yet. It's still bulky, inconsistent, opaque, and in places patently immature. But the direction of travel is clear, and the speed of the exponentials is clear too.

So what happens when human knowledge and human understanding begin to diverge? When the machines *know* things—correctly, usefully, consequentially—that we cannot fully explain? When reason itself risks atrophying because the machine becomes the locus of judgment? Kissinger, Schmidt, and Mundie worried about a kind of “dark enlightenment,” and I understand the concern. There's a real danger here of *deifying the machine*—of treating it not as an instrument but as an oracle. Of reverting, in some strange way, not to Enlightenment rationality but to a kind of technicized religion: primitives trying to placate an inscrutable god.

That's not at all where we are yet. But it's the shadow hanging over the compressed 21st century. We may be headed toward a world in which human intelligence no longer defines the frontier of knowing, even though intelligence has always been the core of our self-conception as a species. *Homo sapiens*. Wise humans. Learning has been our great evolutionary advantage. Our neurons, synapses, memory, abstraction, planning, and reasoning gave us dominion over the earth. And now we're building minds whose learning loops may rapidly exceed our own. Exhilarating, yes. But also existentially disorienting.

The adult answer isn't denial, and it's certainly not capitulation. It's disciplined humility. We have to keep human judgment alive not by pretending it remains supreme in every technical domain, but by relocating it where it matters most: problem selection, moral boundary-setting, validation, accountability, consent, distributive justice, and the interpretation of what a discovery means for a human life. The machine may know more. The human still has to decide what the knowing is *for*.

What Should the 150 Do? What Should We as *Humans* Do?

How does a human intelligence—both individually and as a collective—respond to this? With sadness, regret, mourning? Should we take on a fatalistic or funereal mood because we're no longer the center of the cognitive universe? We've had these epistemic moments before: from Ptolemy to Copernicus; from *Geocentrism* to *Heliocentrism*; from a cosmos seemingly providentially arranged *for us* to a cosmos in which we were somehow demoted, and then demoted again. This one, admittedly, may hit closer to our pride and our home, because it touches the faculty we used to think made us uniquely sovereign. But we've had to make these self-identity revisions before, and now we must do it again.

The practical question for the 150 isn't whether this future is spiritually unsettling. It is. The question is how to *steward institutions through it* without succumbing either to denial or to deification. How do you preserve human agency, moral seriousness, and civilizational confidence in a world where discovery itself is increasingly synthetic? How do you keep judgment alive when judgment is no longer monopolistically human? How do you prevent the machine from becoming an oracle while also refusing to cripple it because it offends our prestige?

That, to me, is the work. Not to stop the compressed 21st century. Not to pretend it isn't coming. But to build the ethical, institutional, and interpretive scaffolding required to live inside it without surrendering either rigor or humanity. Because we're entering a world where we may know more than we can understand. And if we are wise, that won't mean the end of science. It will mean the beginning of a new, stranger, more powerful chapter of it.

And that, finally, brings us back to the practical question with which I began. What is the role of the 150 in this world? Not to sulk. Not to obstruct. Not to retreat into prestige, credentialism, or procedural vetoes. Their role is to understand the magnitude of the transition, to recognize that the insurgents are already moving, and then to decide whether they will install, govern, and shape this multiplication of intelligence inside their institutions—or whether they will merely watch it happen from the sidelines while history is being made elsewhere.

My answer is: the role of the Healthcare 150 becomes *more* important, not less, but perhaps also less flattering. We will know more than we can understand. We will be handed truths, mechanisms, candidates, and interventions that emerge from data manifolds rather than from elegant human theories. The 150's task will be to steward that abundance wisely: define ethical boundaries, enforce governance, demand reproducibility and safety, interpret consequences, protect patient trust, and decide what kind of civilization these new powers should serve.

The scientific method, that great Enlightenment monument, isn't being discarded. It's being *subsumed* and *transmuted* into something larger. The method survives as a moral and institutional framework in a world where discovery itself is increasingly synthetic. That's the new role of the human scientist: less discoverer, more steward; less solitary genius at the bench, more governor of a system that can generate more knowledge than any one mind can ever hope to hold. And that's not a demotion. It is the job of the scientific adulthood in the compressed 21st century.

The Post-Hypothesis Covenant

So the final covenant is this: the AI 10 and the 150 have to co-design the next scientific order before the next scientific order routes around them both in different ways. The AI 10 need to be pulled toward clinical reality, patient trust, validation, safety, and moral seriousness. The 150 need to be pulled out of priesthood nostalgia, committee immobilism, credential protection, and the bad habit of treating *slow* as synonymous with *safe*. The covenant is not “move fast and break patients.” It’s also not “move slowly and preserve every inherited form.” It’s move quickly enough to matter, carefully enough to be trusted, and humbly enough to remember that the point of science isn’t spectacle. The point is healing.

If Part V argued that biology is the high-dimensional prize, Part VI argues that stewardship is the human answer. The machine may generate. The lab may validate. The institution may translate. But only humans can decide whether abundance becomes mercy or merely power. That’s the task of the 150. Not to defend the old priesthood. Not to obstruct the insurgents. Not to deify the machine. Not to pretend the scientific method is dead because the machine has become strange and powerful. The task is to govern the transition from human-exclusive discovery to synthetic discovery with enough wisdom that we don’t lose ourselves in the process.

The Chapter, Compressed

Let me end by compressing the argument, partly as a kindness to the reader and partly because even I recognize that this has been a long walk through epistemology, gerontocracy, AlphaFold, bureaucracy, de-extinction, BCI, and the spiritual demotion of *Homo sapiens*.

First, GenAI matters because it’s a multiplication of intelligence, not merely another software tool. Intelligence is the upstream input that produced civilization’s downstream tools, institutions, sciences, and economies. When that input becomes scalable and non-biological, everything that depends on intelligence begins to change.

Second, we’ve taught the models our civilization. They’ve absorbed the digitized residue of human knowledge—our histories, sciences, laws, philosophies, poems, code, arguments, and accumulated conceptual structures. Our civilization, it turns out, is *learnable*. That’s both marvelous and deeply disconcerting.

Third, this emergent intelligence has properties that matter especially for science: *functional omniscience*, *polymathy*, *omni-disciplinarity*, and *analogic recombination*. Those are precisely the properties that historically marked the rare human mind capable of changing the world. The difference is that model cognition can be copied, distributed, updated, and scaled.

Fourth, the human monopoly on hypothesis generation is beginning to break. That’s what I mean by *The End of Hypothesis*. Humans will still hypothesize, but the frontier of novelty will increasingly include machine-generated candidates, machine-traversed search spaces, machine-designed experiments, and truths that emerge from synthetic inference before they are humanly narratable.

Fifth, this arrives at a moment of civilizational stagnation. Science has become slower, more specialized, more bureaucratic, and more burdened by its own accumulated knowledge. Ideas may be getting harder to find; drug-discovery productivity has moved in the wrong direction; and the gerontocratic institutions of science have become too comfortable with normal science, incrementalism, and procedural legitimacy.

Sixth, biology is the obvious place where returns to intelligence may be highest. Complex disease, aging, neurodegeneration, psychiatric illness, autoimmunity, oncology, and the deep pathologies of the body are high-dimensional problems that have exceeded the absorptive capacity of unaided biological cognition. If AI accelerates discontinuous breakthroughs here, the compressed 21st century becomes real.

Seventh, the scientific method doesn't disappear, but its monopoly ends. It becomes validation, governance, translation, and moral discipline in a world where discovery may happen through continuous synthesis, simulation, optimization, robotic experimentation, and feedback. From Bacon to backpropagation. From hypothesis to generative epistemology.

Eighth, machine truth may become increasingly opaque. We may know more than we understand. We may verify outputs we cannot fully explain. That makes interpretability, alignment, governance, and human agency central—not peripheral—to the future of science.

Ninth, the 150 cannot remain spectators. They have the clinical substrate, longitudinal data, patient trust, translational pathways, regulatory legitimacy, and institutional seriousness needed to convert synthetic discovery into human benefit. The AI 10 bring compute, models, engineering, capital, and speed. Neither side is sufficient alone. The future has to be co-designed.

Tenth, the human role changes but doesn't vanish. We become stewards, governors, interpreters, curators, validators, ethicists, and moral agents. We decide what matters. We decide what is permissible. We decide how to translate knowledge into care. We decide whether this new abundance serves healing, dignity, and mercy—or merely power, profit, and spectacle.

That's the task of the 150. Not to defend the old priesthood. Not to obstruct the insurgents. Not to deify the machine. Not to pretend the scientific method is dead because the machine has become strange and powerful. The task is to govern the transition from human-exclusive discovery to synthetic discovery with enough wisdom that we don't lose ourselves in the process.

We're entering a world where the book of nature may be read at machine speed. That should fill us with gratitude, awe, and fear in roughly equal measure. The adult task isn't to close the book. It's to decide what kind of civilization reads it.

BEFORE WE TURN THE PAGE

If intelligence itself is being multiplied, the next question is what that multiplication does when it enters the clinical act: diagnosis, treatment, liability, physician authority, and the standard of care. So we turn now from generative epistemology to clinical AI.

Chapter 3 • Clinical AI

The Universal Doctor, Liability, and the Reconstitution of Medicine

“Healing is a matter of time, but it is sometimes also a matter of opportunity.”

— HIPPOCRATES, PRECEPTS, C. 400 BCE

A WORD ON NAVIGATING THIS CHAPTER

This chapter moves from epistemology into the clinical act itself. The question isn't whether AI can make medicine less bureaucratic, but whether it can participate in diagnosis, treatment, longitudinal synthesis, liability, and the standard of care without betraying the sacred trust at the center of medicine.

Here's the map, because this chapter keeps moving—intentionally—between clinical opportunity, liability, market structure, China, the god-model labs, the home, and the sacred trust at the center of medicine. **Part I** defines clinical AI and states the opportunity. **Part II** names the old scarcity and the biological defeat. **Part III** moves into the discovery shock: the 17-year graveyard, AlphaGo, AlphaFold, and medicine's AlphaZero moment. **Part IV** maps the new clinical intelligence stack and the coming contest among god models, vertical startups, payers, and the 150. **Part V** names the resistance, the infallibility trap, the guild response, China and the Gulf as installation foils, and the inversion of the standard of care. **Part VI** turns constructive: the physician-architect, the new Flexner settlement, capitation, home, presymptomology, the Universal Doctor, liability assumption, CEO diffusion, and the recap. The through-line is simple: clinical AI isn't another healthcare tool; it's the deployment problem of medicine's new cognitive substrate.

Part I—The Opportunity

This part sets the stakes and definitions. Clinical AI isn't administrative AI with a stethoscope. It participates in the cognitive act of medicine, which is why the opportunity is enormous and the governance problem is so treacherous.

Healing as a Hippocratic Matter of Opportunity

I like the quote at the top of this chapter. We tend to think of Hippocrates—well, to the extent we think of him at all in 2026—as some abstraction, the enunciator of the *primum non nocere* (“first, do no harm”) commandment, the name floating majestically above the entrance to the medical profession like a stone inscription. But let's humanize the guy. I like to think of Hippocrates as a working clinician on the island of Cos in the late fifth century BCE, watching patients live and die, trying to organize what he was learning into a transmissible practice—not so very different, in his daily preoccupations, from his clinician counterparts of 2026. And this particular line, drawn from one of his lesser-cited treatises, *Precepts*, speaks with some insistent contemporary relevance to our own moment, twenty-four centuries later.

Healing is a matter of time, but it is sometimes also a matter of opportunity.

No one, certainly not a Greek physician 2,400 years ago, no matter how clairvoyant, could have

predicted this particular species of opportunity: *from sand, to silicon, to intelligence, to medical superintelligence*. That's the premise of this chapter.

If you're still with me after these first two chapters of pontification and self-indulgent historical digression, perhaps you've arrived at the same conclusion I have. If GenAI is indeed a *multiplication of intelligence*—and, invoking two of the main protagonists in our little play, first Demis: *first you solve intelligence, then you use intelligence to solve everything else*—and now Dario: *the returns to intelligence are highest in biology*—then we're confronting a moment of genuine historical consequence. The most exponential technology in human history is slamming headlong into the most institutional, incrementalist, hyper-regulated, hyper-litigious space in our contemporary society: US medicine.

And the other side of opportunity, in Hippocrates' phrase, is that we might miss it. Or, perhaps with less exaggeration, that someone else *outside* our 150—a Silicon Valley insurgent, a god-model lab, or, dare I say it, another *civilization* less encumbered by the regulatory and litigative scruples of late-modern American institutions (with, just hypothetically, the initials PRC)—preempts us, and we end up importing back what we should have built, validated, governed, and diffused ourselves. I worry that's precisely where we find ourselves, here in the United States, circa 2026. The American medical-industrial establishment, with a few honorable exceptions, several of whom will recognize themselves in the following paragraphs, isn't embracing this moment with alacrity. It's gearing up to delay, to litigate, to committee-ize, to professionalize hesitation, and otherwise inoculate the system against diffusion. Of course, the reasons articulated will be noble—safety, prudence, circumspection, ethics, equity, professional standards, patient protection. Never naked self-interest. Never the preservation of prestige, authority, compensation, credentialed scarcity, and guild sovereignty. But here we are nonetheless.

Clinical AI, Defined

A definition is in order, because *clinical AI* is being used, to my mind, a little promiscuously across the industry, often in ways that conflate what's happening inside the cathedral with what's happening in the vestibule. By clinical AI I mean any AI system that participates directly in the cognitive labor of medical decision-making about a particular patient: differential diagnosis, risk stratification, treatment recommendation, longitudinal care management, behavioral-health intervention, agentic follow-up across time, and real-time decision support at the moment care is being delivered. Clinical AI changes the answer the physician gives. *Administrative AI*—claims processing, denials and appeals, scheduling, documentation, ambient note capture, prior authorization—is meritorious and useful, but it doesn't change the answer; it de-bureaucratizes, cancels friction, and makes the institution delivering the answer less maddening. Both matter enormously, but they aren't the same thing.

In my own ontology of AI in healthcare, as you've heard ad nauseam from me in this and past essays, there are four doors through which this technology enters the sector: **administrative simplification, clinical augmentation, computational and synthetic biology, and consumer empowerment**. The lines of demarcation between and among these categories are messy, of course. Administrative simplification creates the diffusion muscle. Computational biology accelerates the discovery frontier. Consumer empowerment becomes the form factor of ubiquitous, personalized, longitudinal care. But this chapter is, bravely if not imprudently, about the second door: *clinical augmentation*. The domain where AI doesn't merely make healthcare less bureaucratic, but begins to participate in the clinical act itself.

Imprudently, because this is fraught and treacherous terrain. Landmines everywhere—legal, moral,

technological, political, spiritual; some justified, some overwrought. But to share my bias confessionally up front: *we simply aren't moving fast enough on deploying clinical AI*. A growing cohort of vertical startups is showing astonishing progress in autonomous nursing, multimodal diagnostic accuracy across pathology and radiology, and what's beginning to look, in narrow domains, like *proto-medical superintelligence*. The god models are mobilizing in their own idiosyncratic ways—Google's Co-Clinician, OpenAI's ChatGPT for Clinicians, Anthropic's emergent offerings to embed into large health systems and payers. And yet the parties are, by and large, *playing adjacently* to the U.S. healthcare establishment. The insurgent innovators experiment and prototype on one side. The incumbent establishment adopts reticently, when it adopts at all. The regulators look impassively upon the field of action and move at the geologic tempo of administrative law. Each constituency operates with internally rational caution. The aggregate result? Institutional paralysis at the precise civilizational moment that demands something closer to (disciplined) acceleration.

Prudence Isn't Paralysis

Moving slowly *sounds* prudent. We can return, if we like, to the epigraph above and the author of "first, do no harm." But the prudence argument presupposes that the status quo is something worth preserving in its present form. And to be clear, I start from the premise that the status quo isn't tolerable. Just look at the agreed-upon math: the 2023 Johns Hopkins study, published in *BMJ Quality & Safety*, estimated that serious diagnostic errors in the United States kill approximately 371,000 Americans annually and permanently disable another 424,000—something on the order of 30,000 preventable deaths every month. The financial extrapolation of that toll—direct medical cost, premature mortality, lost productivity, absenteeism and presenteeism, the long tail of disability and informal caregiving—gets big fast, approaching a trillion dollars a year. That's the baseline we're protecting. I don't say this in a denigrating, or worse, ungrateful way. U.S. healthcare is astonishingly good in many respects, and at its best it's incontestably the finest acute and specialty care apparatus in the world. I'm simply making the observation that the current reality is far from perfect, and treating it as the morally innocent default is one of the more expensive fictions in American public life.⁵⁷

Now consider the upside. Spend a few minutes inside Dario's beautiful mind, read *Machines of Loving Grace*, or listen to the podcast he and I did in January, and the frontier starts to look almost civilizational in scale: the amelioration of mental-health suffering, the eradication or radical mitigation of complex diseases that have stood ineradicable for the entirety of human history, progress against neurodegenerative conditions that have defied every prior pharmacological generation, and perhaps, as Dario postulates, a couple of decades of lifespan expansion over the next twenty years. Add mitigated health disparities, democratized expertise, deflationary pricing and the prospect that we might miss this, or delay it by a decade, should give us serious pause. Clinical AI, properly deployed, can narrow the ZIP-code-conditioned disparities that disfigure American medicine, improve diagnostic accuracy beyond what any physician aided only by status quo technology can achieve, and bring something like the standard of care of an academic medical center into the pocket of the rural Kenyan farmer and the West Virginia coal miner's widow. That's the breakthrough we may be sitting on.⁵⁸

The Public Needs an Undeniable Win

And we need a win. Fast.

⁵⁷ David Newman-Toker et al.: "Burden of Serious Harms from Diagnostic Error in the USA" published on January 19, 2024 ([Link](#))

⁵⁸ TowerBrook Healthcare Institute: "Incumbents and Insurgents in Healthcare: Dario Amodè" published on January 6, 2026 ([Link](#))

Let me explain, because I mean that more literally than it may sound. Clinical AI is now in a race between three forces: public sentiment, guild protectiveness, and the undeniable progress of the inventors. The guilds will resist. The regulators will hesitate. The legal system will proceduralize. And the public, understandably, will remain uneasy about machine medicine until the benefits become legible, emotional, and undeniable. You've heard me say, repeatedly, that AlphaFold is the single greatest scientific achievement of the last half-century, and Demis and John Jumper rightly won the Nobel Prize for it. But protein folding is just too esoteric to capture the public imagination fully. It's too abstract, too upstream, too difficult to translate into the kitchen-table conversations of ordinary people. The public won't fall in love with solved protein structure. The public *will* fall in love when an incurable disease becomes curable. When a child with a rare cancer lives. When an Alzheimer's trajectory bends. When ALS is reversed. When a rural patient gets a diagnosis that previously would have required a pilgrimage to Boston, Houston, or Rochester. That's the win we need.

So let the labs invent. Let DeepMind and Isomorphic and OpenAI and Anthropic and the biological AI companies keep throwing AlphaFold-scale advances at nature as fast as they responsibly can.

Let the startups demonstrate actual clinical improvement, not shiny slides, not clever user experiences, not benchmark theatrics, but measurable better outcomes in real populations. Let the clinical AI companies produce the first undeniable examples of superior performance inside bounded, validated envelopes. Let the win arrive *loudly* enough that the old objections begin to sound smaller.

I'm not saying public sentiment should govern clinical truth. It shouldn't. But public sentiment matters in a democracy, and a frightened public, with less than 30% optimism in many AI polls, plus a protective guild, plus a litigious culture, is a recipe for quagmire.⁵⁹ *The antidote isn't another white paper on responsible innovation.* The antidote is an AI-generated clinical breakthrough large enough to change the emotional weather. Once patients, families, employers, payers, and policymakers see a disease altered, a diagnostic error prevented, a psychiatric access desert partially irrigated, or a rural health gap narrowed, the AI Overton window shifts. The guild can still invoke safety, and it should. But it becomes harder to use safety as a velvet rope around the old order when the alternative is visibly better and cheaper care.

To be clear, I'm not arguing that the technology is universally ready for prime time. It's a jagged frontier of competency and capability, and the stakes couldn't be higher. Unleashing this fusillade of startups and hyperscaler offerings indiscriminately could produce catastrophic outcomes. We need prudence. We need validation. We need serious governance, real post-market monitoring, clinician safe harbors, patient consent where appropriate, escalation pathways, equity measurement, and an adult liability architecture.

But what I observe across most of the U.S. medical-industrial complex isn't prudence. It's *immobilization*. Let's form an AI governance committee. Let's run more pilots. Let's leave clinical AI for some indefinite, *once-we-get-more-proof* future. Let's wait for the specialty society. Let's wait for the board. Let's wait for legal. Let's wait until the model is perfect, the liability is resolved, the regulator has spoken, the professional society has issued guidance, and the machine has somehow achieved a *standard of infallibility* we've never imposed on a human being. That's not safety. That's the entrenched, predictable playbook of an ancien régime refusing to recognize the moment it's in.

⁵⁹ YouGov: "The Economist/YouGov Poll" published in May 2026 (Link)

The Chapter's Route

So let me tell you where this chapter is going, because the architecture matters. First, I want to explain why the old scarcity that organized medicine—the scarcity of human clinical intelligence—has ended, and why the institutions built around that scarcity now have to justify themselves again. Then I want to show why *biology outran the biological brain*, why hyperspecialization and bureaucratization were rational adaptations to cognitive defeat (an echo of the previous chapter's *Generative Epistemology* ideas, applied to this new context), and why clinical AI offers the first plausible path back toward *synoptic* medicine. From there I'll move through the 17-year innovation graveyard, medicine's AlphaZero moment, the god-model invasion, and the startup community's increasingly uncomfortable discovery that thin clinical software isn't much of a moat when the gods come downstairs. Then we get to the central gate: *liability*. Somebody has to volunteer to get sued. And once that gate opens, the standard of care begins to invert, the physician becomes architect rather than data-entry clerk, the home becomes an epicenter of care, capitation becomes more than a policy slogan, and the *Universal Doctor* moves from science fiction toward institutional design. And finally, I'll close with a merciful summary of the dozen or so big insights in this chapter (skip to the back if you're impatient).

The Liability Thesis

The short answer, which I'll spend the rest of this chapter defending, is this. *The bottleneck is risk tolerance and liability*: the absence of any party—physician, hospital, developer, payer, malpractice carrier, reinsurer—willing to absorb blame when the machine errs, undergirded by the hyper-conservative, culturally risk-averse gestalt of a soft, complacent, late-stage empire, a lawyerly society that has come to treat the threat of suit as the operating constraint of medical practice.

The lever to break it is the assumption of product and medical-malpractice-style liability inside a validated envelope. My bet remains that the first truly decisive mover is a well-capitalized hyperscaler—Google, in my estimation, could do it, with Microsoft and Amazon plausible alternates—because the balance sheet, reinsurance capacity, infrastructure, and scientific ambition matter enormously here. But we shouldn't make the argument too monocausal. Hyperscalers aren't the only actors that can begin to assume or share liability. A well-capitalized clinical AI startup with a narrow enough use case, enough validation, enough monitoring, and enough reinsurance can do it. Digital Diagnostics already showed the embryo of that model. An enlightened payer can do it where it controls the premium dollar, the ambulatory network, the clinical workflow, and the economics of prevention. And an enlightened provider, especially a large health system with a malpractice captive, can participate in liability-bearing deployment by defining the clinical envelope, instrumenting model use and override, monitoring outcomes, and pricing risk against real-world performance.

The Liability Stack

So the actual unlock is probably not *one* heroic actor assuming all risk in splendid isolation. It's a *liability stack*: hyperscaler balance sheet, startup vertical excellence, payer economics, provider deployment substrate, reinsurance, and malpractice captives woven into a structure that finally lets clinical AI move. Liability assumption is necessary; it isn't sufficient. The clinical substrate—the work-flows, the patient relationships, the physician base, the validated envelopes, the institutional governance, the local legitimacy, the actual conditions under which medicine becomes medicine—sits with the 150. The hyperscaler or well-funded startup may underwrite the model risk. The payer may underwrite the economic logic. The provider hosts the deployment, protects the sacred trust, and turns the technology into care. The partnership *between and among* them is the actual unlock.

That means the 150's job, starting now, isn't to admire the future from a safe institutional distance. It's position themselves as the *indispensable deployment partner* before the hyperscalers and AI-native insurgents find their substrate somewhere else. Create clinical AI command structures at the CEO and board level. Build physician-led co-design labs. Define validated clinical envelopes. Demand liability-bearing partnerships. Use malpractice captives strategically. Instrument model use, model override, model under-ride, uncertainty flags, equity effects, adverse events, cost, latency, outcomes. *Diffuse through employed physicians*. Align deployment with capitation, delegated risk, provider-sponsored plans, advanced primary care, and home-based care. Prepare for the standard-of-care inversion before it arrives. Protect the sacred trust so that clinical AI gives caregivers back to patients rather than replacing them with a cheerful simulacrum.

The consequence, once those parties find each other and the gate finally swings, is the *reconstitution of medicine itself*: from the hierarchically rationed, ZIP-code-conditioned, hyper-specialized sick-care system we currently administer into something closer to what the founder of the discipline, twenty-four centuries ago, would have recognized as the *practice of healing*.

She who assumes the liability, wins. I'll spill a lot of ink defining and defending that sentence. We'll see if I'm persuasive. But she who hosts the deployment matters too.

The Moral Question

The question we're facing is, in the end, not strategic. It's not regulatory. It's not even economic, though the economic dimensions are almost incomprehensible in scale. It's a *moral and ethical question*: whether we have the institutional courage to deploy the most powerful clinical instrument in human history, or whether we'll hide behind precedent, procedure, and the polite committee culture of American medicine while patients suffer at a baseline rate that should distress us all. And underneath the moral question sits a civilizational one I've been circling throughout this essay and will name plainly here: *autocracies are structurally better at installing a major technological revolution* than messy, pluralistic, lawyerly Western democracies. Installation is the differentiator. Yes, as in previous and future chapters, I'll point to China and the PRC as leading the way here, as disquieting as that comparison may be. If we diffuse clinical AI inside an American institutional framework with anything like *American institutional speed*, the deflationary, lifespan-extending, democratizing benefits of this technology will accrue first to monarchies and autocracies, and the US may end up *licensing them back* from those other geographies. I'll return to that argument later. I want it on the table now.

So: a chapter about urgency. A chapter about opportunity. A chapter about the difference between

prudence and paralysis. A chapter about liability, yes, but also about imagination, courage, deployment, and the possibility that medicine might finally become synoptic again.

The 150 must mobilize. Here's how—and if you're one of them, this chapter is, with as much directness as I can muster, written for you.

Part II—The Old Scarcity and the Biological Defeat

This part gives the diagnosis beneath the diagnosis: medicine's institutions were built around scarce biological clinical intelligence, and biology has now outrun the biological brain. The question is whether the old scaffold can be re-justified for the intelligence age.

The Old Scarcity Has Ended, and the Institutions Haven't Noticed

Reader Note: this section intentionally recapitulates a few themes from the preceding Future of Science chapter—the multiplication of intelligence, the end of old cognitive scarcity, and the institutional surround built around scarce biological expertise. If you've just read that chapter sequentially, skim as needed. I'm keeping the material here because the clinical argument needs the epistemic foundation underneath it, and because some readers will come straight to the Universal Doctor chapter without first wandering through my earlier epistemological thicket.

A quick retrospective before the argument can move forward. GenAI isn't just another tool. It isn't another wave of digitization, another productivity layer, another category of enterprise software. It's a *multiplication of intelligence*; a non-biological intelligence that already matches, and in certain narrow domains exceeds, the cognitive performance of the biological intelligence around which medicine, science, credentialing, malpractice law, reimbursement, and the entire professional architecture of healthcare were constructed. If you accept that distinction, the rest of this chapter follows. If you don't, the rest will land as alarmist, or overheated, or maybe just one more Larsen excursion into the metaphysics of software. Fair enough. But I think the distinction is essential.

What matters for the clinical argument is this: the scarcity that organized modern medicine has ended. Again, I expatiated on this idea at (too much) length in the *Generative Epistemology* chapter, but in case you skipped that navel-gazing meditation, here's the bite-sized version. Modern medicine was built around the *scarcity* of human clinical intelligence: scarce expertise, scarce synthesis, scarce pattern recognition, scarce access to the full body of medical knowledge, scarce capacity to hold biology *synoptically*. The physician was valuable not only because she cared, examined, diagnosed, and treated, but because she was one of the very few biological creatures licensed, trained, credentialed, and socially authorized to interpret the body. That scarcity gave rise to the whole surround: the degree, the residency, the fellowship, the specialty society, the board certification, the malpractice architecture, the prestige hierarchy, the prestige compensation, the RVU, the fee-for-service substrate, the white coat, the clinical note, the ordination rituals of the profession.

And now that scarcity is dissolving. Not completely, uniformly, or without risk. But yes, directionally and, unmistakably. With the instantiation—or *speciation*, if we're being more dramatic and maybe more accurate—of this new silicon intelligence, we can finally begin to take a more synoptic and integrated view of the high-dimensional data space of human biology. The patient, the literature, the guideline, the image, the pathology slide, the genome, the medication list, the prior authorization history, the social context, the behavioral pattern: these no longer have to live as disconnected

fragments scattered across human memory, institutional silos, and specialty fiefdoms, or (non-inter-operable) EMRs. They can, at least in principle, be represented together. Reasoned over together. Updated together. And once that becomes possible, the institutional scaffolding built around the old scarcity begins to look less like forever architecture and more like historical artifact.

That's the point that medicine hasn't yet metabolized. *When the underlying scarcity dissolves, the structures built around it don't remain untouched.* They may remain useful in some diminished part. They may preserve wisdom. They may carry legitimate social trust. But they also become misaligned with the new substrate. Credentialing, residency, specialization, malpractice, reimbursement, professional autonomy, even the episodic face-to-face encounter all need to be reconsidered from first principles—*Elon-like*, if you'll forgive the adverb—because they were designed for an age in which clinical intelligence was scarce, local, biological, slow to train, and difficult to distribute. That's no longer the world we're entering. The scaffold doesn't automatically collapse, but it no longer explains itself. It has to justify itself again.

This is why the ivy-covered resistance machinery of American medicine—the AMA, the AHA, AHIP, the phalanx of specialty societies that pay rent in Washington DC and Alexandria Virginia, the high-prestige academic medical centers dotting the eastern seaboard, the state boards, the malpractice carriers, the credentialing bodies, the hospital risk committees, the genteel committee culture I keep returning to—finds itself in such a strange position. These institutions were built, in large part, to *steward the old scarcity*. They protected quality when expertise was rare. They credentialed authority when knowledge was hard to acquire. They created boundaries when cognitive overload made boundaries necessary. They built a priesthood because the divinities of biology were incomprehensible and the people needed someone to intercede on their behalf. But now a non-biological intelligence has arrived that can absorb, synthesize, compare, and increasingly reason across domains at a scale no human guild can match. The guild therefore confronts the most threatening kind of challenge: not a rival guild, not a new specialty, not a cheaper labor substitute, but an *attack on the scarcity that justified the guild's sovereignty* in the first place.

Medicine's Move 37 Moment

This is the Move 37 moment for medicine. I rhapsodized about AlphaGo in last year's *GenAI Juggernaut* piece, so I'll mercifully just wave at it here before returning to medicine's own AlphaZero moment later. But hold the image. An exponential intelligence makes a move that contradicts the accumulated wisdom of the practice. The world's most credentialed human practitioners stare at it and conclude, for a moment, that the system has malfunctioned. It must have erred. It must have violated the canon. It must not understand the game.

But the system hasn't malfunctioned. It has *graduated*. Or perhaps more accurately, it has *evolved*. That's the disorienting fact. The institutional establishment of American medicine—formidable, long-stewarded, encrusted with prestige and authority, and built across generations for an entirely different scarcity than the one we're entering—is among the last to recognize that the change has already happened. This isn't because the establishment is obtuse. It isn't. It contains enormous clinical wisdom, tacit knowledge, moral gravity, and actual experience with suffering bodies. But institutions built to protect an old equilibrium almost always confuse the equilibrium with reality itself. They begin to think that because the architecture endured, the architecture is permanent. They mistake stewardship for destiny. They mistake the authority to certify expertise for a permanent monopoly on expertise.

So here comes the collision in American healthcare, and it isn't, in the first instance, a technology

collision. It's a sociological and cultural one. On one side stand the archetypically institutional bodies of the profession: medicine itself, the AMA, the AHA, AHIP, hospital risk committees, malpractice carriers, specialty societies, accreditation bodies, state medical boards, the *slow granite architecture* of an ancient guild. Some of these bodies have stood for more than a century. They're built for durability, continuity, prudence, professional hierarchy, and the long quiet rhythms of an ancien régime. They have the visual grammar of the monumental architecture of my hometown Washington, DC: marble, columns, pediments, implied foreverness. On the other side stands an exponential technology moving at the speed of silicon, indifferent to credentialing, indifferent to deliberative committee culture, indifferent to the professional courtesies by which medicine has governed itself for generations. It's built for iteration, recursion, permutation, kaleidoscopic change. These two systems weren't designed to meet. The meeting is happening anyway, and it isn't going to be tidy.

The danger, though, isn't simply that the old institutions resist. Resistance is predictable. *The deeper danger is that the resistance postures as virtuous.* It will call itself safety, prudence, ethics, professional standards, patient protection, clinical validation, equity, oversight, and responsible innovation. Some of that will be true. We need those things. Clinical AI can't be propagated like a consumer app with a cheerful onboarding flow and a 'click here to see our legal agreement' link. But the 150 will need to develop a much more discriminating ear for the difference between legitimate caution and old-scarcity protectionism. When an institution built around scarce human expertise confronts abundant machine expertise, its safety language can become indistinguishable from self-preservation. And the cost of mistaking one for the other may be measured in avoidable deaths, avoidable disability, avoidable cost, and the ceding of the clinical AI frontier to actors with less patience for American procedural pieties.

The 150 as Co-Designers

That's why the task of the Healthcare 150 is so important. Our job—yours, mine, theirs—is to navigate this collision without losing the wisdom and perspective the incumbents have dutifully stewarded over long decades, and without simply handing the whole thing over to a class of twenty-something Silicon Valley techno-solutionists who, indispensable though they are to the moment, ought not be deputized as the sole superintendents of this phase shift. Medicine isn't just another workflow. The patient at 1 a.m. isn't a dataset. The sacred trust isn't a UI/UX problem. The health system has tacit knowledge that the model labs will never possess, and the model labs have technical power that the health systems will never reproduce. The future of medicine depends on joining those assets before one side overwhelms the other.

But joining them requires the 150 to stop behaving like passive procurement departments. That's the practical implication of this whole section. If the old scarcity has ended, then health systems can't merely wait for clinical AI to arrive as a polished vendor product, blessed by regulators, indemnified by someone else, and safely installed in Epic after a few pilots and eighteen months of governance meetings. *They have to co-design the new architecture.* They have to help define validated clinical envelopes. They have to decide where machine judgment meets human judgment. They have to build physician-led clinical AI labs, not performative innovation theater. They have to use their malpractice captives as instruments of deployment design. They have to demand liability-bearing partnerships from hyperscalers and durable vertical AI companies. They have to instrument model use, override, under-ride, error, equity, cost, latency, and outcome. They have to use their employed physician base as the diffusion channel. They have to prepare for the standard-of-care inversion before the inversion arrives and humiliates them. A big and imposing list, to be sure, but not impossible.

I've said it before: the 150 have abdicated much of their co-creative, co-developmental privilege for nearly every prior technological phase shift of the past generation—consumer internet, mobile, social, cloud, big data and analytics, enterprise SaaS, and, yes, the EHR, which was largely externally imposed (and then resented) for two decades. We can't do that this time.

That's the argument. The old scarcity has ended, and the institutional 'surround' built to buttress that scarcity now has to be rethought. Not discarded wholesale, and not burned down for the amusement of the insurgents. But rethought, re-justified, re-architected. The 150 still matter because they have the bedside, the trust, the workflows, the physicians, the malpractice infrastructure, and the legitimacy without which clinical AI remains a product rather than medicine. But their relevance is no longer guaranteed by incumbency. It has to be earned through co-design, governance, deployment, and moral seriousness.

The system hasn't malfunctioned. It's evolved. The question is whether our institutions can evolve with it. And to see why that evolution isn't optional, we have to look at the deeper predicament underneath the whole profession: biology itself has become too large for the biological brain.

When Biology Outran the Biological Brain

Reader Note: another deliberate echo of the prior chapter follows. There I made the abstract claim that biology became too high-dimensional for unaided human cognition, and that hyper-specialization, homogenization, and bureaucratization were rational but incomplete compression techniques. Here the same claim gets translated into the clinic. The patient isn't an epistemology problem; she's the person who experiences the consequences of that fragmentation in referrals, portals, notes, denials, handoffs, and the ordinary misery of having no one hold the whole story at once. If that earlier argument is fresh in your mind, you can skim this section for the clinical application rather than the philosophical premise.

My 21-year-old son Sebastian, who's similarly nerd-sniped on AI, taught me a word the other day: "biologist." Evidently this is someone who believes in the superiority, or at least the irreplaceability, of un-cybernetically enhanced human intelligence. I loved the word instantly. It has exactly the right mix of adolescent internet taxonomy and lurking metaphysical seriousness. But it also made me wonder whether I'm something like a biologist myself, at least when we're talking about the sacredness of human judgment, the physician-patient relationship, and all the irreducibly human parts of medicine I keep insisting we mustn't lose.

And yet this chapter is, in part, an argument *against* biologicalism as an epistemology. Not against human beings. Not against clinicians. Not against judgment, trust, presence, responsibility, or love. But against the comforting idea that unaided biological cognition can still compass the full complexity of modern medicine. That distinction matters. The physician remains morally central, but the physician's brain can no longer be treated as the sole or sufficient instrument for holding the whole of biology, the literature, the patient, the workflow, and the longitudinal state of disease in view. Allow me to explain.

To understand why the resistance machinery is ultimately fated to fail—and why clinical AI isn't some passing Silicon Valley enthusiasm, some frothy demo-day artifact, some ambiently interesting tool that'll sit politely beside the existing order—we have to start with the cognitive predicament of

contemporary medicine. I've argued this at length elsewhere, but the point bears repeating because everything in this chapter depends on it. Medicine isn't struggling because physicians are unintelligent, inattentive, lazy, or insufficiently devoted. Quite the opposite. Clinicians are heroic, self-sacrificing, brilliant, and humanitarian. But medicine is struggling because its *object of study*—biology—has become too vast, too high-dimensional, too data-rich, too combinatorial, and too fast-moving for unaided biological cognition to hold. The human brain didn't fail medicine. Medicine succeeded so dramatically that it eventually outran the human brain. And our institutions haven't quite caught up to this reality.

As I argued in the preceding *Generative Epistemology* chapter, for roughly four hundred years, the scientific method was humanity's most advanced cognitive technology. It served us splendidly because data was scarce, our instruments were crude, and nature was hostile, opaque, and mostly incomprehensible. Bacon, Galileo, Newton, Boyle, Pasteur, Koch, Lister—they helped build and elaborate the most successful epistemic process we've ever devised for rendering the world legible. The scientific method was a compression algorithm for reality. It gave us a way to take the sprawling confusion of nature and discipline it into observation, hypothesis, experiment, falsification, and cumulative knowledge. And in medicine, the consequences were astonishing.

Vaccination. Germ theory. Antisepsis. Anesthesia. Antibiotics. Imaging. Insulin. Genomics. Immunotherapy. Each a genuine medical miracle. Each a compression of nature's complexity into something the human prefrontal cortex could hold, reason about, teach, and eventually standardize. And collectively, a quadrupling of human lifespan from the pre-industrial era to today.

That's no longer our central problem. Our problem is the opposite. We're positively *drowning* in biological complexity at precisely the moment our cognitive equipment, having been adequate for several centuries, becomes inadequate to the task. The human body is a dynamic, nonlinear, high-dimensional, multi-omic, immunologic, metabolic, behavioral, social, environmental system, evolved across roughly four billion years of biological and evolutionary tinkering. As I expand on in my BCI chapter, the human prefrontal cortex is an evolutionary marvel; let's give it its due. It gave us language, memory, planning, civilization, poetry, war, mercy, mathematics, medicine, and, regrettably, hospital committee meetings. But it wasn't designed to calculate the nonlinear interactions of a trillion-cell organism while simultaneously assimilating the global medical literature, integrating multi-modal patient data, tracking drug-drug interactions, adjusting for individual genetic variation, interpreting imaging, reconciling the medication list, remembering what was discussed six months ago, and then compressing all of that into a fifteen-minute clinical encounter that has to be documented, coded, justified, defended, and perhaps litigated. Oh, and figuring out how to get paid for it all by the Blues.

The clinician is being asked to synthesize more than the biological organism is capable of synthesizing. That isn't a failure of clinicians. It's a structural mismatch between the complexity of the work and the cognitive equipment we've brought to it. And once you see that clearly, a great deal of modern medicine suddenly becomes more intelligible—not less frustrating, perhaps, but more intelligible. The hyperspecialization, the guidelines, the quality metrics, the committees, the paperwork, the clinical variation, the defensive crouch, the burnout, the pajama time, the procedural asphyxiation. These aren't random pathologies. They are, quite reasonably, the predictable institutional adaptations of a profession trying to manage complexity beyond human scale.

Three Coping Mechanisms for Cognitive Defeat

Medicine, like every sufficiently complex knowledge domain, has responded to this avalanche of information through three coping mechanisms: *hyperspecialization*, *homogenization*, and *bureaucratization*. I don't mean to present these as moral failures. They're rational, locally adaptive responses to cognitive overload. But rational responses can become societally costly. They can begin as clever adaptations and harden into orthodoxies. They can help us survive complexity for a while and then prevent us from transcending it. I diagrammed these phenomena out in the *Future of Science* chapter, so I'll just quickly restate them here in the current clinical context (and for those readers who skipped directly to this chapter and didn't suffer through the others).

Hyperspecialization is the most obvious. We draw smaller and smaller circles around what one mind can plausibly master. The kidney. The receptor. The lesion. The tumor subtype. The arrhythmia. The molecule. *Experts know more and more about less and less*, not because they're narrow people, but because the field forces the narrowing. Biology is too large, so we look at the body through a straw. *And the consequence is that the body remains one thing while the clinical enterprise built around it becomes a thousand*. The diabetic patient pinballs among endocrinology, nephrology, ophthalmology, podiatry, pharmacy, nutrition, behavioral health, social work, cardiology, vascular surgery, and God knows who else, and the system—having assembled this phalanx of individuated experts, each in command of one tiny sliver of the person—then asks itself, somehow, magically, to aggregate all of that back into a coherent account of a single human being. It doesn't work. At least not well.

That's the first pathology. We carved the patient into cognitively manageable pieces and then built an enormous coordination apparatus (the *coordination tax* I discuss—and quantify economically—elsewhere) to reconstruct the person we had just disassembled. EHRs, referrals, handoffs, care managers, documentation requirements, population health registries, risk-adjustment machinery, prior authorization teams, CDI queries, and on and on. *We hyperspecialized in order to survive the complexity, and then had to build a bureaucratic superstructure to manage the fragmentation that hyperspecialization produced*. This is why the patient herself so often becomes the integration layer, which is an unenlightened thing to do to a sick human being. The richer, whiter, more urban, more educated, more persistent patient can sometimes navigate that maze. Everyone else gets lost in it.

The second coping mechanism is *homogenization*, and I want to sharpen what I mean by that. I don't primarily mean protocols, checklists, guidelines, and care pathways here; those belong more naturally to bureaucratization, and I'll come to them in a moment. By homogenization I mean the *intellectual convergence of a field* under the tyranny of the citation, the prestige journal the fundable question, the approved method, the socially sanctioned research agenda. A mature knowledge domain, overwhelmed by its own output, begins to cluster around what has already been blessed. Everyone gloms on to the same citation-rich center of gravity. The same questions get asked with better methods. The same paradigms get extended. The same safe topics receive funding. The same senior people (often the gerontology leading prestigious peer-reviewed journals or department heads at ivy-covered east coast universities) decide what counts as serious. Heterodoxy becomes career endangerment; orthodoxy becomes career management. Read the book *Boom* by Hobart and Huber for a revelatory set of insights on this.⁶⁰

This isn't a conspiracy I'm documenting. It's more, well, rational sociology. It's also how good ideas get delayed. Wise to remember that CRISPR was fringe, messy, and treated as insufficiently canonical not so long before it was awarded a Nobel prize. Google my interview with Stephane Bancel, CEO

⁶⁰ Byrne Hobart and Tobias Huber: "Boom: Bubbles and the End of Stagnation" published on November 19, 2024 ([Link](#))

of Moderna, where we observe that mRNA spent half a century rustivating in the wilderness before becoming, in the space of a pandemic, one of the defining technological triumphs of modern medicine. A profession built around consensus is good at protecting the average; it's less good at recognizing discontinuity before the discontinuity becomes safe enough to commemorate. The tyranny of citation gives us rigor, maybe, but it also gives us intellectual herding. It makes the field more legible, more fundable, more governable—and sometimes less alive.

The third coping mechanism is *bureaucratization*, and here I'm far less sympathetic (as my TowerBrook partner Ian Sacks will attest) though I understand its origin. Once a field becomes too complex for any one mind to hold synoptically, institutions do what institutions always do. They standardize. They *proceduralize*. They create protocols, checklists, guidelines, care pathways, prior authorization rules, coding regimes, quality metrics, committee structures, documentation requirements, credentialing processes, and escalating layers of review. Interesting to note that the guy who came up with handwashing checklist was voted 'Time 100' just a few years ago (no disrespect to my friend Peter Pronevost who's a brilliant thinker—I'm just commenting on our enshrinement of and reverence for 'the process'). Some of this is necessary. I'm not arguing for medicine as anarchy. But past a certain point, bureaucracy stops being a safety mechanism and becomes human organizational kudzu. It grows over the work, around the work, through the work, and eventually people start mistaking the kudzu for the institution.

The Mercatus Center counted 49,312 healthcare regulatory restrictions in the federal code alone and more than 805,000 across state codes. That's a procedural drag coefficient no other major industry suffers under in quite the same way, and we've somehow naturalized it as the cost of doing business. I'm 53, as I've reminded you a couple times already—yes, I'm trying to rationalize my recent birthday—and I've been working inside or adjacent to American health systems for almost three decades. The administrative weight has gotten heavier, more Kafkaesque, more suffocating every single year. There's no honest argument that this accretion has improved patient outcomes commensurate with its cost. There is, however, a very honest argument that it's slowly extinguished the joy of practice.

And this is the part we need to be brave enough to say. Hyperspecialization, homogenization, and bureaucratization were rational adaptations to cognitive defeat. They weren't evidence that the system was healthy. They were evidence that the system could no longer be held together by synoptic human intelligence. Smaller experts, narrower questions, thicker rules, more forms, more committees, more coordination machinery. This is what a knowledge system looks like when its subject matter exceeds the unaided processing power of the biological brain. It doesn't collapse immediately. It compensates. It subdivides. It standardizes. It builds institutional scaffolding around cognitive insufficiency.

And now, with the advent of GenAI, we've instantiated a non-biological cognitive faculty whose strengths are *precisely* complementary to the weaknesses that drove us into those adaptations in the first place. Polymathy. Total recall. Synthesis across domains. Synoptic understanding. Analogic thinking. Pattern recognition at the scale of the global literature. Continuous availability. Tireless attention. The ability to hold the image, the lab, the genome, the note, the prior-auth rule, the medication list, the patient message, the longitudinal trend, the guideline, the trial data, and the thousand relevant papers in one representational space. None of these are things the unaided human prefrontal cortex can provide. All of them are things a sufficiently capable clinical model increasingly can.

This is why clinical AI is so profound. It doesn't merely automate part of the existing system. *It offers a way out of the coping mechanisms that became the system.* It can help reintegrate what

hyperspecialization fragmented. It can search beyond the tyranny of the citation and surface heterodox, interstitial, analogic relationships that no guild would have thought of (or thought to prioritize). It can reduce the bureaucratic load by making the coordination tax computationally tractable rather than administratively endless. It can, in the best case, let medicine become synoptic again—not by asking any single human to become omniscient, but by surrounding the human with a non-biological intelligence capable of holding the whole.

Biological Defeat Is Not Physician Failure

That doesn't mean we discard the human. This isn't about ego protection, but it also isn't about human humiliation. The point isn't to lament the fact that clinicians are finite. The point is to acknowledge that the game has changed irrevocably and our institutions haven't adapted commensurately. We've wrapped the physician in an almost theological reverence for clinical judgment while denying her the tools needed to exercise that judgment at the scale of modern knowledge. Then we've buried her in documentation, forced her through prior authorization rituals, surrounded her with half-integrated systems, and asked her to take moral (and legal) responsibility for a care model that's cognitively and administratively impossible.

No wonder the profession is burned out, cynical, defensive, and increasingly bureaucratic. No wonder the art of medicine has become, for too many, the art of *surviving the machine*. No wonder physicians retreat into autonomy, specialty identity, defensive practice, and skepticism toward tools that arrive from outside the guild. They've been asked to bear the burden of a system that has already exceeded the biological brain's capacity, and then they're blamed for not transcending biology.

That's why I think the phrase "biological defeat" matters, even if it sounds a little dramatic. It names something we otherwise avoid saying. Medicine has been cognitively defeated by its own success. It generated more knowledge than it could absorb. It subdivided the body because no one could hold the whole. It standardized because variation became too hard to govern. It bureaucratized because complexity needed to be managed somehow. These weren't signs of failure in the stupid sense. They were signs of success becoming unmanageable. Biology won. Or more accurately, biology grew too complex for biological cognition alone.

Clinical AI is the first plausible deliverance from that predicament.

Not a perfect deliverance. Not an infallible deliverance. Not one without risk, hallucination, opacity, liability, bias, demagoguery, misuse, and all the other landmines I'm trying to take seriously elsewhere in this chapter. But a deliverance nonetheless. If you don't see medicine as already cognitively defeated by its own success, you'll see clinical AI as a threat to a working system. You'll reach for the usual language: safety, validation, physician autonomy, ethics, professional standards, liability, trust. And some of that language will be sincere and important. But if you do see the defeat, clinical AI looks different. It looks less like a threat and more like the first real chance to reconstitute medicine around the scale of the knowledge it's created.

That's the new reckoning. *Biology outran the biological brain*. Medicine compensated by narrowing, converging, and proceduralizing. Now a silicon intelligence has arrived with the very properties medicine lacked: polymathy, memory, synthesis, attention, and scale. The question isn't whether we should protect the old coping mechanisms out of sentimental loyalty to the institutions that built them. The question is how quickly, safely, and humanely we can redesign those institutions around the new cognitive substrate.

Because the old arrangement isn't coming back. The body has outgrown the brain. The task now is to build medicine for the intelligence age.

Medicine as Engineering, Not Science

There's a deeper possibility underneath everything I've just argued, and I want to name it because it may be the actual hinge of the chapter. In the *Generative Epistemology* chapter, I described biology as a high-dimensional, seemingly infinite, perhaps functionally incompressible space: too combinatorial for AlphaGo-style Monte Carlo tree search, too messy for self-play, too nonlinear for our tidy human categories, too ontologically promiscuous for the biological brain. That remains the view from inside biologicalism. But what if that's only *our* limitation, not nature's? What if medicine and biology aren't finally intractable scientific mysteries, but *information-organization problems* at a scale we've simply never possessed the intelligence to traverse?

That's where my overused phrase *functional verifiability* becomes relevant again. In math, code, theorem-proving, many engineering domains, and other right-or-wrong answer spaces, automation accelerates because the system can *test itself against ground truth*. The answer either compiles, proves, runs, predicts, heals, or fails. Biology has *seemed* different because its answer space is terrifyingly high-dimensional and our ground truth has been slow, expensive, wet, noisy, human, ethical, and delayed by morbidity and mortality. But if a sufficiently powerful *Jupiter Brain* can organize the state space of biology, infer mechanism, propose interventions, and verify against empirical outcomes with enough precision, then biology begins to acquire the property that made software and mathematics so amenable to machine intelligence. It becomes, at least partly, functionally verifiable.

If biology becomes functionally verifiable, medicine changes category. It becomes an engineering problem, not a science problem. I don't mean that science disappears; the wet lab, clinical trial, animal model, organoid, biological twin, and lived patient remain reality's veto. I mean that the center of gravity shifts from *artisanal discovery* to *recursive design-build-test-learn*. Science asks what's true. Engineering asks what can be made to work, repeatedly, safely, economically, and at scale. Medicine has historically been a science-flavored craft practiced inside institutions. Clinical AI makes it possible to become an engineering discipline wrapped in sacred trust.

This sounds like science fiction now. But will it remain science fiction indefinitely? I don't think so. Follow the exponentials. Or, to invoke Leopold Aschenbrenner, whom the 150 will remember from last year's paper, and a smaller number will remember from the dinner I organized for him before the hedge-fund lore, before the Wall Street Journal hagiography, before the rapid canonization of the young prophet of the OOMs (orders of magnitude): *follow the OOMs*. That was Leopold's simple and trenchant exhortation that night: don't narrate the future linearly when the relevant substrate is compounding geometrically. And whatever one makes of the surrounding mythology—the former OpenAI researcher, the Situational Awareness manifesto, the AI-infrastructure fund, the reported leap from a few hundred million dollars of early public exposure into \$20 billion AUM of disclosed holdings—the underlying lesson is the same. *The people who saw the orders of magnitude early saw the world more clearly than the people extrapolating yesterday's slope*. Follow the OOMs in medicine, and the thing starts to look Kurzweilian: healthcare, or at least the cognitive and biological-discovery layer of healthcare, begins to behave less like a labor-bound guild economy and more like an information technology.

That's the ontological leap. If biology remains an unbounded, incompressible, non-verifiable fog, then clinical AI helps around the edges: better documentation, faster literature review, fewer missed

refills, cleaner coding, more humane inbox management. Useful, but not civilization-altering. If biology is a gigantic information-organization problem that a Jupiter Brain can progressively render tractable, then clinical AI doesn't merely augment medicine. It colonizes the core. It turns diagnosis, therapeutics, prevention, senescence, pharmacology, and longitudinal care into recursive engineering surfaces. And then everything I say elsewhere in this chapter about liability, standard of care, capitation, home, physician-architects, and the *Universal Doctor* becomes not speculative ornamentation but institutional preparation for a new cognitive substrate of medicine.

Part III—The Discovery Shock

This part explains why the old diffusion cadence is no longer defensible. The 17-year graveyard, AlphaFold, the Bitter Lesson, and the medical data project all point in one direction: clinical knowledge has to move faster, and some of it will be generated by machines rather than merely summarized by them.

The 17-Year Innovation Graveyard

Reader Note: this section also reaches back to the stagnation argument in *Future of Science*. There the problem was Eroom's Law, institutional sclerosis, and the slowing machinery of biomedical discovery. Here the same delay becomes morally more concrete: years of translational latency aren't just sad graphs in innovation economics; they're years in which patients wait, deteriorate, relapse, and die before the system converts insight into usable care.

If biology has outrun the biological brain, then the next question is brutally practical: what does that cognitive mismatch cost patients? Not metaphysically. Not philosophically. Not in some airy "future of medicine" panel-discussion way. What does it cost in missed diagnoses, slow diffusion, stale guidelines, outdated habits, delayed therapy, avoidable deterioration, and the silent attrition of human life that never shows up in a headline because it happens one chart, one visit, one missed clinical update at a time?

Here's the number that should keep every one of the 150 awake at night: the average lag from clinically validated discovery to routine clinical adoption and diffusion in the United States is approximately *seventeen years*. Yup. Seventeen years.⁶¹ And sure, one can quibble with the methodology or this or that study and this or that specific timeline. But the inarguable point it is that we're sloooow at diffusion in our industry. People outside US healthcare are incredulous when they hear just *how* slow we are, and they should be. Diagnostics or treatments that manifestly work, supported by persuasive evidence, peer-reviewed literature, and enough clinical validation to deserve movement, can languish for almost two decades in the purgatorial gap between what the literature says and what the clinic does. Some never make it. Some die quietly in the guideline committee, or the EHR build queue, or the local medical-staff politics, or the payer-policy labyrinth, or the simple fact that no human being can keep up with the deluge.

This is a moral catastrophe we've learned to file under the soothing administrative category of "academic medicine, slowness of." That filing system has been very convenient for us. It allows us to treat delay as a feature of seriousness, or safety, or virtue, rather than a source of harm. It lets the

⁶¹ Zoë Morris et al.: "The answer is 17 years, what is the question: understanding time lags in translational research" published in December 2011 ([Link](#))

current system present itself as careful, sober, and appropriately evidence-based, while patients spend years—sometimes decades—living outside the reach of knowledge that already exists. The current system isn't safe; it's merely familiar. There's a difference, and we've been ignoring it for a generation.

The publication machine only makes the absurdity more visible. Biomedical science now produces roughly 1.5 million peer-reviewed papers a year, before we even get to preprints, trial updates, guidelines, real-world evidence, genomic datasets, imaging datasets, pathology datasets, claims data, and the growing deluge of patient-generated information from wearables and remote monitoring. One rather evocative estimate suggested that primary care physicians trying to stay current would need something like twenty-nine hours of reading in a twenty-four-hour day. The arithmetic is comic, but the consequences aren't funny at all. *No one can pay the reading debt.* Not the diligent PCP. Not the subspecialist. Not the academic department chair. Not the guideline committee. Certainly not the exhausted clinician doing two hours of pajama-time documentation after dinner and pretending she still has the psychic space to ingest the frontier of cardiometabolic medicine before bed.⁶²

So the system does what human systems do when the information environment exceeds the mind. It fragments. It specializes. It formalizes consensus. It builds committees. It waits for specialty societies. It waits for payers. It waits for local clinical leadership. It waits for the next EHR update. It waits for clinicians to change habit. It waits, and calls the waiting prudence.

But if the previous section's argument is right—if *hyperspecialization*, *homogenization*, and *bureaucratization* are coping mechanisms for cognitive defeat—then the 17-year lag is perhaps best understood as the operational expression of that defeat. Knowledge exists but isn't absorbed. Evidence is published but not diffused. Data accumulate but don't metabolize into care. The body of medicine learns faster than the delivery system can act. And the patient lives in the gap.

The 17-Day Mandate

So I want to propose, partly seriously and partly as directional provocation, what I'll call the *17-day mandate*: new validated clinical knowledge should move from discovery to clinical availability in days or weeks, not decades. Not every paper, every preprint, every breathless abstract from a lab with a hyperactive PR department. And certainly not recklessly. Not without validation, adjudication, and clinical judgment. But directionally, yes. The operational tempo of medical knowledge has to change. The old cadence—paper, commentary, guideline, CME, local committee, EHR build, payer recognition, slow physician habit change—is incompatible with a world in which medical knowledge compounds at machine speed.

Clinical AI is the only plausible mechanism for that compression. A model can ingest the literature, compare emerging findings against patient state, identify what has changed, score evidence quality, surface relevance to a specific cohort, flag uncertainty, route to the right clinician, and support incorporation into care far faster than unaided human committees can. That doesn't mean the model becomes sovereign and clinicians become subordinate. It means the model becomes the compression layer between global medical knowledge and local clinical practice.

That's what the 150 need to build: not another innovation program, but a *clinical diffusion engine*. A way to move from knowledge to bedside without waiting for a generation. Every major system should have a standing machinery—their own internal R37 lab (more on that in subsequent chapters)—that continuously scans validated science, compares it against internal outcomes, identifies where

⁶² Brian Alper et al.: "How much effort is needed to keep up with the literature relevant for primary care?" published in October 2004 ([Link](#))

practice lags evidence, produces model-supported pathway updates, measures adoption, and instruments patient outcomes. The old guideline pilgrimage can't remain the only mechanism by which medicine updates itself. The 150 need an always-on clinical observatory, a translation cell, a pathway foundry, and a diffusion apparatus, all governed by clinicians but metabolizing knowledge at silicon speed.

This is practical advice. When a new heart-failure therapy works, who in the system knows which eligible patients aren't on it? When a pharmacogenomic signal affects antidepressant selection, who finds the patients still trapped in trial-and-error medication roulette? When a new cancer biomarker changes therapeutic relevance, who searches the living population, not just the next patient seen by the right subspecialist? Today, we mostly rely on humans, meetings, dashboards, care managers, and *heroic local process*. Tomorrow, this should be agentic, continuous, measured, and embedded.

That's the 17-day mandate. It doesn't abolish clinical judgment; it rescues clinical judgment from the impossible task of reading everything, remembering everything, updating everything, and operationalizing everything manually. It lets physicians adjudicate the frontier instead of drowning beneath it. It lets hospitals move at the speed of evidence rather than the speed of committee metabolism.

And this is where the 17-year graveyard starts to connect to the liability argument I'll make shortly. Once a health system has the capacity to know, update, and diffuse faster, the old excuse that "medicine just takes time" becomes less persuasive. A system that can move validated knowledge to the bedside in weeks will eventually be compared against one that still moves at the speed of committee metabolism. The future standard of care won't be whatever your committee finally approved after three years of careful deliberation. It'll increasingly be what the best validated model, linked to the best clinical governance and the best liability architecture, can safely deliver now.

The alternative is the system we currently have: a thousand acts of clinical inertia justified and defended as prudence.

From Move 37 to AlphaFold: Medicine's AlphaZero Moment

Reader Note: yes, AlphaGo and AlphaFold return here. I'm not repeating the parables because I've forgotten I already used them; I'm returning to them because they're the cleanest bridge from epistemology to medicine. In the prior chapter, they showed that machine intelligence can produce nonhuman (transhuman?) insight. Here they ask the harder clinical question: what happens when the model makes a medical Move 37—a recommendation that looks wrong to the guild because it violates canon, and then turns out to be better?

The 17-year graveyard is about the failure to diffuse what we already know. The deeper and more destabilizing claim is that clinical AI won't merely diffuse *existing knowledge* faster. It will begin generating *new* clinical and biological knowledge in ways that challenge the inherited sovereignty of human expertise itself.

From Game Canon to Medical Canon

That's why the AlphaGo analogy matters. Not because Go is cleanly analogous to medicine (although there are some fascinating convergences...Go has roughly 10^{170} possible board positions—more than there are known atoms in the universe—but medicine's diagnostic possibility space is plausibly even *larger* once you include disease states, genetics, comorbidities, medications, labs, imaging, behavior, environment, time, and treatment response.). I'm certainly not suggesting doctors are game players and patients are stones on a board. The analogy matters because it marks the moment when *human expertise becomes the substrate rather than the ceiling*. It's the moment when the system doesn't merely learn the canon; it begins to see through and beyond it.

I talked about this at length in last year's essay, so I'll just quickly recap here: on March 9, 2016, in Seoul, Korea, 280 million people watched Demis Hassabis' AlphaGo play Lee Sedol, the eighteen-time world champion of Go, in the second game of their historic match.⁶³ On move 37, AlphaGo placed a stone on the fifth line. To the Korean, Chinese, and Japanese commentators watching, it looked amateurish, almost embarrassing (Demis and team wondered momentarily if the machine had gone haywire). Essentially, the accumulated wisdom of four millennia of Go pedagogy held that you simply *didn't* play a stone there in the early middle game. Lee Sedol left the room for fifteen minutes, nonplussed. What he'd just witnessed, though he didn't yet know it, was a *one-in-ten-thousand* move—a move no human player would have made precisely because it violated the accumulated heritage of the game. AlphaGo won. Later analysts called Move 37 an act of beauty. The first celebrated instance of machine creativity.

By the by, if you haven't seen the AlphaGo documentary, stop reading and go watch it. I'll wait. Lee Sedol's facial micro-expressions during Move 37 are themselves a small civilizational artifact. You can see the human mind encountering, in real time, an alien intelligence that hasn't made a mistake, but has *stepped outside the human history of the field*. That's the thing to marvel at. Not the game, but the epistemic shock.

The first AlphaGo learned, in part, by studying human games. It *encoded human mastery* and then, after lots of self-play and millions of positional experimentations, exceeded it. But AlphaZero went further. It played itself exclusively, unencumbered by *any* human teaching. It learned from the structure of the game rather than from the accumulated human canon. It developed strategies no human had conceived. This is where Rich Sutton's Bitter Lesson becomes so important. Sutton's point, distilled across decades of AI research, is that human expertise, encoded too reverently, becomes a constraint. The systems that win over time are the ones that use more raw compute and more raw data to discover patterns *unburdened by our human, handcrafted assumptions* about what the patterns should look like. Human knowledge is useful at first. Eventually it can become an impediment to transcendence.

For medicine, that's a destabilizing proposition. For 2,400 years, medicine has been about encoding, accumulating, transmitting, and formalizing human clinical wisdom. The credentialing structure, the prestige hierarchy, the senior physician, the fellowship, the board exam, the case conference, the sacred intonation of "in my experience"—all of this rests on the premise that accumulated human expertise is the irreplaceable resource. Sutton's lesson suggests that in bounded domains, and eventually in larger ones, we may want systems that first discover the way we do, and *then the way we can't*. Not now, not universally, and not without risk. But the direction of travel is getting clearer.

⁶³ Google DeepMind: "AlphaGo at 10: How AI Innovation Is Paving the Path to AGI" published in March 2026 ([Link](#))

AlphaFold as the Bridge

And the reason AlphaGo and AlphaZero matter for medicine isn't only because they showed machine intuition in a game. *They matter because they led to AlphaFold.* That's the bridge. That's the reason this whole analogy isn't just a curious intellectual side quest. DeepMind's movement from Go to protein folding is the modern parable: a system first demonstrates that it can transcend human canon in a constrained symbolic domain, and then a related family of methods begins to unlock one of the most important problems in biology. AlphaFold was the epistemic earthquake.

For decades, predicting a protein's three-dimensional structure from its amino-acid sequence was one of biology's grand challenges. The sequence tells you the ingredients; the folded structure tells you much more about function, interaction, disease relevance, and therapeutic possibility. Experimentally determining protein structures through X-ray crystallography, cryo-EM, NMR, and related techniques was slow, expensive, laborious, and human-intensive (it took an estimated five years for a PhD to map a single protein). Then AlphaFold arrived and changed the field's horizon entirely. AlphaFold 2's performance in CASP14 stunned structural biology. DeepMind later released predictions for 230 million protein structures, putting a vast new map of biology into the hands of researchers around the world. Then AlphaFold 3 pushed toward broader biomolecular interactions: proteins, DNA, RNA, ligands, and the relational machinery of life.

Demis Hassabis and John Jumper, along with David Baker for computational protein design, won the 2024 Nobel Prize in Chemistry. That Nobel matters symbolically. It's the establishment acknowledging, perhaps a little belatedly but still profoundly, that machine intelligence has crossed from tool into scientific protagonist. I'll be provocative: AlphaFold is arguably the greatest scientific achievement of the past half-century. Not because it solved all biology. It didn't. Not because wet labs are obsolete. They aren't. Biology still gets the last word. Reality still gets to veto the model. But AlphaFold showed that machine intelligence can traverse a biological possibility space too vast for unaided human cognition and produce outputs that reorganize the work of an entire field.

AlphaFold isn't clinical AI in the narrow sense. It doesn't sit across from a patient, reconcile the medication list, or decide whether the febrile but as-yet-undiagnosed patient can go home. But it changed the epistemic credibility of machine-discovered biology. It showed that a model could traverse a biological possibility space too vast for unaided human cognition and produce outputs that reorganized the work of an entire field. That's why the analogy matters. Not because protein folding is medicine, but because medicine is downstream of biology, and biology has now seen what machine intelligence can do when it stops merely reciting the human canon and begins moving through nature's search space on its own terms.

That's why medicine's AlphaZero moment is so important. *AlphaFold is the proof of concept for a new regime of biological discovery.* It's Dario's compressed 21st century in embryo (incidentally, who did Dario dedicate *Machines of Loving Grace* to? Demis). It shows that the path from human-encoded expertise to machine-discovered structure isn't a fantasy. It has *already* happened. In structural biology. At Nobel scale.

Now bring that logic back to clinical medicine. The early clinical AI systems—AMIE, MedPaLM2, the first GPT-4-era iterations—largely worked by encoding human domain expertise. Train on medical exams. Train on clinical notes. Train on literature. Train on what doctors know. This was reasonable. *It was the AlphaGo phase of clinical AI.* The system learns the canon, the question style, the human explanation, the diagnostic schema, the standard medical reasoning moves. That phase matters. It

gets us competence. It gets us trust, or at least the beginnings of it. But it's anything but the endpoint.

The actual endpoint is more unsettling: models that discover relationships in clinical and biological data that humans haven't named, couldn't see, or would have dismissed because the relationship didn't fit the canon. A treatment pattern that seems counterintuitive but works in a molecularly defined subpopulation. A diagnostic feature in pathology invisible to the human eye but predictive of outcome. A drug-response signal distributed across genotype, imaging, labs, and social context. A deterioration pattern emerging from low-grade deviations in sleep, gait, heart rate variability, medication adherence, and bathroom use. A rare disease cluster hiding in the messy interstices of EHR notes and lab trends. A "Move 37" in medicine won't look like a stone on a board. It will look like a recommendation that violates the accumulated intuition of the guild and then turns out to be right.

That's the eventual subordination of some forms of human knowledge and judgment to machine medical intelligence. I know that sentence will make people recoil. It should. It's a big claim. But better to say it directly than euphemize our way around it. This isn't about abolishing physicians, and it's not about bowing down to a machine oracle (although I do warn us about that in my Deification chapter, so take note). It is about admitting that in domains where the data are too dense, the interactions too nonlinear, the literature too vast, and the human mind too bounded, machine intelligence may become epistemically superior. The physician then becomes less the solitary generator of truth and more the steward, verifier, interpreter, moral agent, and accountable human presence in a larger clinical intelligence system.

That's professionally disorienting. But I believe it to be the direction of travel. The clinical evidence, even if you haircut the lab claims and discount the marketing bravado, is already formidable. Early 2026 benchmarks now look more like the floor than the ceiling. Early reasoning models demonstrated correct diagnoses in complex case studies at multiples of unaided human clinicians. Newer clinical benchmarks such as OpenAI's HealthBench are beginning to show performance gaps between frontier reasoning models and ordinary human clinical performance that would have seemed fantastical two years ago. Academic literature seems increasingly willing to use the word *superhuman* without scare quotes. Sober people, peer-reviewed people, people who don't generally sound like my San Francisco techno-optimist buddies, are now reaching for vocabulary that used to belong to Sam Altman's more exuberant moods.

Benchmarks Are Weather, Not Care

Now, let's be clear-eyed about benchmarks. Benchmarks certainly don't *care*. A clinical vignette isn't a dying patient, looking a doctor in her eyes and asking for rescue. A correct answer on a test isn't a workflow, a bedside conversation, a liability envelope, an EHR integration, a reimbursement model, or a grieving family. We shouldn't confuse benchmark superiority with deployable medicine. But we should also not commit the opposite error and pretend that benchmark superiority means nothing. *Benchmarks are early weather*: they tell us the pressure is changing. And the pressure is changing quickly.

The important point for the 150 is that machine superiority will arrive first in bounded domains. That's where medicine's AlphaZero moment will actually enter the enterprise. Not as the *Universal Doctor* on day one. Not as some autonomous superintelligence strolling into the ICU and taking over rounds. It'll arrive as second reads, risk stratification, clinical trial matching, diabetic retinopathy, pathology pre-screening, medication optimization, sepsis escalation, home-monitoring signals, refill renewals, cancer biomarker interpretation, depression screening with escalation, and narrow, validated clinical

envelopes where the model's performance can be measured against human practice. Those are the bounded Move 37s.

And this is why I put the themes of the 17-year graveyard and the AlphaZero moment back-to-back. One is about speeding diffusion of what humans *already* know. The other is about admitting that the frontier of knowing itself is moving *beyond* purely human generation. Clinical AI will both compress the time from discovery to bedside and begin generating discoveries that need to move to bedside. That *double acceleration* is what the current institutional architecture is least prepared to metabolize.

So what should the 150 do with this? First, stop treating clinical AI as a futuristic curiosity. Second, stop assuming that the medical guild's accumulated prestige will give it veto power over machine-discovered truth. Third, build the validation machinery now: model-versus-human benchmarks by domain, median physician, top-decile physician, local institutional performance, resource intensity, safety, equity, latency, outcome. Fourth, create clinical AI translation cells that can move from validated model output to supervised deployment in weeks, not years. Fifth, instrument every use case so that when the standard of care begins to invert, your organization has evidence rather than vibes. Sixth, prepare your physicians psychologically for the idea that clinical judgment remains sacred, but unaided clinical judgment doesn't.

That last line is the difficult one. The physician isn't being erased—nothing like it. But the physician is being relocated inside a new epistemic hierarchy. The old model placed the physician at the apex of clinical knowing and judgment. The new model places the physician inside a human-machine system that may know more than any individual clinician can understand unaided. The best physicians won't be diminished by this; they'll be amplified by it. The worst institutions will experience it as humiliation and resist. The best institutions will experience it as deliverance and govern it.

That's medicine's AlphaZero moment. Not that machines beat doctors in a contest. That's the shallow reading. The deeper reading is that medicine is beginning to encounter forms of clinical and biological intelligence that don't merely encode the human canon, but exceed it. AlphaGo showed the move. AlphaZero showed the method. AlphaFold showed the biological consequence. Clinical AI is the next migration.

The model doesn't merely remember medicine. It begins to design and discover medicine. The question for the 150 isn't whether to participate. It's how, on whose terms, and at what pace.

The Bitter Lesson in the Body

I'll close this section with a bit of a leap in imagination, starting with Rich Sutton's canonical 2019 Bitter Lesson essay.⁶⁴ To my healthcare audience, this essay is required bedside reading in Silicon Valley—it's short, hard-hitting, and profound. Worth a quick Google read. Sutton's exhortation belongs here because medicine is precisely the kind of field that will be tempted to encode its own human superstition into the machine and call it wisdom. Sutton's argument, in its most brutal form, is that the long-run winners in AI haven't been the systems most lovingly hand-engineered with human domain knowledge, but the systems that exploit general methods, more data, more compute, and more search. The hand-coded expert system impresses early. The scaled learner eventually dominates it. AlphaGo was already disorienting because it absorbed human games and surpassed the canon. AlphaZero was more unsettling because it learned by self-play, from the structure of the game itself, without reverently ingesting the guild's accumulated pedagogy.

⁶⁴ Rich Sutton: "The Bitter Lesson" published on March 13, 2019 ([Link](#))

Clinical medicine will rebel against this lesson before it metabolizes it. We'll try to encode guidelines, pathways, expert consensus, specialty-society preferences, department-chair intuition, prestige-journal orthodoxy, and local clinical folklore into models and then congratulate ourselves for having made them safe. Some of that is prudent, even necessary, in the first act. But in future snowstorms much of it may look like superstition: elaborate, credentialed, humane, sincere, and wrong. The Bitter Lesson says: add more data, more compute, more inference time, better feedback loops, better simulations, better empirical verification, and let the system search the possibility space we can't.

That means biology has to be penetrated, not merely summarized. The clinical AI stack needs ground truth: interventions, outcomes, time, sequence, dose, context, comorbidity, genotype, phenotype, adherence, social reality, imaging, pathology, claims, pharmacy, proteomics, metabolomics, microbiome, epigenetics, exposome, behavior, and *all* the heterogeneous messiness that makes actual human life so inconvenient to the randomized controlled trial. The *connectionist project* is to rip through the data, not because correlation is magic, but because causation, correlation, recombination, and mechanism can be triangulated at scales no human tumor board, guideline committee, or department retreat can hold.

This is why the god models need access—governed access, consented access, privacy-preserving access, federated where necessary, sovereign where required, but real access—to the exabytes and perhaps zettabytes of clinical, claims, imaging, multiomic, pharmacy, behavioral, and home-generated data now scattered across proprietary tombs. We should be building data trusts, model-evaluation networks, synthetic controls, federated-learning rails, and clinical-outcome registries as though this were a great national project, because it is. The Human Genome Project looks, in retrospect, like a small and noble preface to the *world medical-data project we now need*.

And this project shouldn't stop at the water's edge. High-dimensional biology isn't American, Chinese, Emirati, Indian, Nigerian, Brazilian, or European. It's human. The model needs heterogeneous world data: genetic diversity, molecular diversity, phenotypic diversity, epigenetic diversity, environmental diversity, dietary diversity, infectious-disease diversity, socioeconomic diversity, and the chaotic heterogeneity of care delivery itself. A model trained only on affluent American academic medical centers won't become the *Universal Doctor*. It'll become a very impressive Boston consultant with poor global manners.

So perhaps this is also a diplomatic project. Imagine a Sino-American-Gulf-European-African medical intelligence consortium—not naive, not borderless, not ignorant of biosecurity or sovereignty, not a kumbaya hallucination—but a serious humanitarian collaboration around medical ground truth. A Jupiter Mind ripping through the world's medical data to reduce suffering, discover mechanisms, extend healthspan, and update the standard of care. In a world otherwise stumbling toward a Thucydides trap, the shared project of healing could be a rare de-escalatory substrate. The PRC and the United States may compete in chips, models, weapons, energy, and industrial policy. They might still cooperate, or at least interoperate, around the civilizational prize of not letting people die senselessly from tractable disease.

The longevity implications are the most vertiginous. Bowhead whales live more than two centuries. Greenland sharks may live more than five. Nature has already solved versions of aging that human medicine has barely learned to describe. Ray Kurzweil and Peter Diamandis call the target *longevity escape velocity*: the moment when science gives back more than a year of life expectancy for every year lived.⁶⁵ That may be overly exuberant on the timeline, and I can already hear the sober geroscience people sharpening their scalpels. Fine. But the question isn't whether a conference-stage

⁶⁵ Peter H Diamandis: "Longevity Escape Velocity: Nearing Immortality" published on May 2, 2024 ([Link](#))

version of immortality arrives this year. The question is whether senescence is an engineering surface. Elon has made the point, admittedly in a reductionist way, that your right arm doesn't age faster than your left arm; something systemic is *coordinating the clock*. My spellcheck, horrifyingly, changed *senescence* to *armament* as I typed this, which may be either a warning from the machine or simply a Freudian autocorrect. Either way, the point remains: if aging is clock, pathway, damage, repair, immune surveillance, metabolism, and information loss, then it's not metaphysical. It's biology. And biology is the domain to which we're now bringing Jupiter-scale intelligence.

DOUBLE HUMAN LIFE SPAN sounds unserious until one remembers that mammalian and vertebrate life already contains far stranger outliers than our current actuarial imagination permits. I'm not predicting a 500-year human at the end of this paragraph. I'm saying that *once medicine becomes an information technology with functional verification loops*, the morally relevant time horizon changes. The job of the 150 isn't to debate whether this sounds like science fiction. It's to build the institutional substrate that lets the science fiction become validated, governed, equitable, and human before someone else diffuses it without us.

Part IV—The New Clinical Intelligence Stack

This part is about market structure and deployment architecture. The god models are entering medicine, the startup layer is being repriced, and the 150 have to decide whether they are architects of clinical AI or merely substrate for someone else's system.

The God Models Come for Medicine

If the last section was about medicine's AlphaZero moment—the point at which machine intelligence stops merely reciting the human canon and begins to discover around and beyond it—then the next question is no longer philosophical. It's institutional and strategic. Who's going to own this new intelligence layer? Who's going to distribute it? Who's going to govern it? Who gets the workflow, the data, the liability, the trust, the economics, the standard-of-care influence? Because the movement from *AlphaGo* to *AlphaFold* to *clinical AI* isn't happening inside the stately committee metabolism of the Healthcare 150. At least not yet. It's happening in and around the frontier labs, the hyperscalers, the venture-backed insurgents, and the handful of healthcare-native companies that were early enough, lucky enough, or obsessive enough to build data and distribution moats before the god models turned their full attention toward medicine. Probably relevant to point out here that DeepMind, originator of AlphaGo and AlphaFold, is owned by Google, and Sir Demis Hassabis now leads AI at the mothership.

So let's time travel back from Move 37 and AlphaFold to June 2026, which is to say to the rather less romantic but strategically decisive world of market structure, capital, capex, model access, and startup survival. The ground is shifting under the feet of half the venture-backed clinical AI industry as I write this sentence, and in many of the conversations I'm having, I'm not sure the founders fully realize it. Some do. The very best ones absolutely do. But many still believe they're building independent clinical AI companies in a world where the god models will remain polite substrate providers, happily renting intelligence to the application layer while staying above the fray. I think that's a dangerous and probably wrong assumption.

The god models are coming for clinical AI regardless of what the specialists, venture capitalists, hospital CEOs, or Healthcare 150 would prefer. This isn't because the frontier labs have suddenly discovered medicine's sacredness, though some of them speak about it with genuine seriousness. It's

because healthcare is too large, too knowledge-intensive, too economically consequential, and too obviously exposed to machine intelligence to be ignored. The Mag 7 now tips the scale at something like \$22 trillion, larger than every national economy on earth except the United States itself.⁶⁶ The major hyperscalers are spending at sovereign scale on AI infrastructure. OpenAI and Anthropic are valued not like software vendors, but like emerging geopolitical actors. Even the healthcare “startups” are no longer especially *startuppy* in the old garage-and-pizza-box sense; OpenEvidence at a \$12 billion valuation isn’t exactly two kids in Palo Alto building a widget between problem sets. When behemoths of this scale decide medicine is worth their attention, one shouldn’t greet the news with mild curiosity. One should get ready.

And evidently, they’ve decided. In the last two weeks of April 2026 alone, the signs were comically coordinated and concentrated (as if the PR teams of the frontier labs were basically on the same Signal thread). OpenAI launched ChatGPT for Clinicians, a free offering for verified physicians, nurse practitioners, and pharmacists, evaluated on HealthBench and built on GPT-5.⁶⁷ OpenEvidence, trained exclusively on peer-reviewed medical literature and very smartly collecting the imprimatur of NEJM and JAMA while embedding directly into Epic at Mount Sinai, as an example, crossed \$100 million in annualized revenue at a ‘decacorn’ valuation.⁶⁸ And on April 30 Google DeepMind announced its AI Co-Clinician initiative, powered by Gemini 2.5 Ultra: a triadic care architecture that puts AI under physician authority during the patient encounter, reportedly recording zero critical errors in 97 of 98 realistic primary-care queries.⁶⁹ The WHO is forecasting a shortfall of 10 million health workers by 2030. Our friend Demis is proposing, in a tone no one who has met him would mistake for casual, to close part of that gap.

Healthcare is 18.0% of American GDP. No self-respecting hyperscaler or frontier lab running a trillion-dollar valuation thesis can leave that domain to specialists indefinitely. The big boys and girls are turning their attention to medicine, and the dynamic isn’t simply the familiar story of classical creative destruction—the iconoclast in the garage dethroning the sleepy incumbent, though there is still some of that romance at the frontier. *What we’re also seeing is creative agglomeration.* The big are getting bigger. The gravitational fields intensify. The hyperscalers become sovereign-scale centers of compute, capital, distribution, model talent, and liability-bearing capacity. The insurgents that survive are increasingly the ones that find a stable orbit around Google, Microsoft, Amazon, OpenAI, Anthropic, or some other deity in the pantheon. In clinical AI, this is playing out on a tempo the boards of the 150 aren’t remotely prepared for.

⁶⁶ Motley Fool: “The Magnificent Seven Stocks” published on June 8, 2026 ([Link](#))

⁶⁷ OpenAI: “Making ChatGPT better for clinicians” published on April 22, 2026 ([Link](#))

⁶⁸ CNBC: “OpenEvidence, the ‘ChatGPT for doctors,’ doubles valuation to \$12 billion” published on January 21, 2026 ([Link](#))

⁶⁹ Google DeepMind: “Enabling a new model for healthcare with AI co-clinician” published on April 30, 2026 ([Link](#))

The Startup Landscape Matters

That's why I need to shift the gaze, briefly, from the 150 to the startup community. Not because the 150 should become venture tourists, or because every hospital board needs a half-baked view on Series B clinical AI valuations. *The reason to understand the startup landscape is that it determines the partnership landscape for incumbents.* The 150 aren't going to build the foundation models. They're not going to out-compute Google. They're not going to recreate DeepMind, OpenAI, Anthropic, or Microsoft in the basement of the innovation center, no matter how inspirational the internal hackathon may have been. *They'll have to partner.* But if they partner naively—if they scatter procurement across departments, fall in love with thin wrappers, mistake a delightful demo for a defensible company, or confuse a “better model” with a durable moat—they'll build their clinical AI strategy on sand (and not that kind that goes into silicon).

Thin Companies Get Eaten

Here's the uncomfortable message I most want venture-backed clinical AI to contemplate: *pure clinical AI software is increasingly uninvestable.* Or maybe less provocatively, pure clinical AI software is *becoming* uninvestable unless it sits on top of something the god models can't easily replicate. The moat that once justified pure software (as the prescient investor Naval Ravikant argues)—*the difficulty of feature replication*—has collapsed under the god models' capacity to commoditize application functionality. AI can replicate features. It can replicate interfaces. It can replicate workflows at the surface layer. It can absorb a clever prompt chain, a document-generation flow, a triage assistant, a clinical-sounding chatbot, a “copilot” wrapped around a foundation model, and then *give away something very close to it for free* as part of the base product. I'm sympathetic here: the god models must do this to justify their trillion dollar valuations. What AI can't so easily replicate are communities, proprietary data flywheels, deep distribution relationships, regulatory depth, hardware integration, EHR entrenchment, longitudinal workflow ownership, and, in clinical medicine, a real liability posture. I'll invoke a phrase here I'll use elsewhere in different contexts: this is the *anti-disintermediation* stack for the insurgents, but this time against the god models.

Absent that, the *thin-wrapper companies* are in trouble, as I've pointed out in my papers for the last three years. A prompt, a UI, a clinical-sounding name, and a foundation model underneath don't constitute a moat. *They constitute a temporary feature waiting to be swallowed.* And some of the god model providers aren't even pretending otherwise. OpenAI's free clinician-facing product on GPT-5 isn't just an interesting research side-project. It's not a genteel contribution to the literature. It's a committed foray into the 18.0% of U.S. GDP that's healthcare. It says, with the brutality of a free product launched by an imminently trillion-dollar valuation company, that the application layer shouldn't assume the foundation layer will remain celibate.

Again, the durable clinical AI companies are the ones that have built something the algorithm can't simply copy. Proprietary clinical data pipelines. EHR integration at scale. Workflow ownership. Regulatory depth. Specialty-specific distribution. Clinical relationships. Longitudinal outcomes.

A trust footprint. A liability architecture. Tempus, currently hovering at around a \$10 billion market capitalization, is interesting because its moat isn't just a prettier interface. It's a massive, multimodal, *clinically contextualized* data flywheel: research records, cardiology patients, neuropsychiatric datasets, radiology images, genomic sequencing, molecular profiling, digitized pathology, clinical notes, real-world outcomes. OpenEvidence is interesting because burrowing into Epic at Mount Sinai is more than a channel; it's an embedded workflow. It lives closer to the clinician's actual epistemic

moment. Paige and VIRCHOW, domain-specific foundation models, are interesting because computational pathology isn't generic summarization. It's something much more complicated: a domain-specific model trained on enormous quantities of high-fidelity pathology data, operating in a modality where curated data and clinical-grade validation matter enormously. These are real moats. A "better model" isn't.

The VIRCHOW example is especially important because it clarifies the first phase of the market. Domain-specific models had, and in some areas still have, a head start. A pathology foundation model trained on millions of H&E-stained whole-slide images, built from carefully curated clinical material, is doing something a general chatbot doesn't initially know how to do. VIRCHOW and PRISM-like models point toward a world in which small, curated, high-quality, domain-specific data can, *for the moment*, outperform much larger general-purpose models in clinical-grade tasks: rare cancer detection, biomarker prediction, cell identification, pan-cancer detection, molecular inference from morphology, treatment-response prediction. The pathology slide is an almost absurdly dense object. The human eye sees tissue. The model may see prognosis, mutation, recurrence risk, therapy response, and patterns not yet named by the human pathologist. That's why the domain-specific models matter. They're closer to the biological signal.⁷⁰

But—and this is the strategic "but" many startups don't want to confront—the god models aren't standing still at the edge of medicine with their hands folded respectfully. Once they gain access to proprietary medical data, once they can reason over multimodal EHR, imaging, pathology, genomics, claims, notes, outcomes, and workflow data, the distance between the general model and the vertical specialist may compress very quickly. *The god model with access to the right clinical data will rival, then surpass, the domain-specific model that once looked safe behind its narrow moat.* I view this as a question of *when*, not *if*. The question isn't simply whether a domain-specific model is better today. The question is whether its advantage survives when the frontier lab gets access to comparable data and wraps it in a more powerful reasoning engine, better tooling, broader distribution, and perhaps, eventually, liability-bearing capacity.

Harvey's Law for Clinical AI

Let me provide an adjacent example. Harvey, the celebrated legal-AI darling that vaulted to an \$11 billion valuation, started by building its own proprietary legal foundation model.⁷¹ Then, no surprise, the frontier models advanced remorselessly, replicating more and more of the expertise that a vertical foundation model was supposed to own. Harvey then did something intelligent: it stopped trying to compete with the gods. It rode them. Harvey's model selector now routes legal tasks to whichever underlying model—OpenAI's, Anthropic's, Google's—performs best on that task. It built 25,000 custom workflows on top, integrated with Microsoft 365 Copilot, processes hundreds of thousands of agentic queries a day, and became one of the most valuable AI companies in the world without owning a foundation model. I'm going to call that Harvey's Law (let's see if this catches on): the winning vertical AI company isn't the one that competes with the god models; it's the one that builds proprietary workflow, data, distribution, and customer intimacy on top of them.

Clinical AI has to internalize Harvey's Law; but also understand why medicine makes it harder. Harvey's error domain is mostly documents. A wrong legal brief costs money, embarrassment, sanctions, maybe a lost case. A wrong clinical recommendation, to state the obvious, can end a life. That liability asymmetry is the reason the ride-the-god-models strategy in clinical AI requires something

⁷⁰ Vorontsov et al.: "A foundation model for clinical-grade computational pathology and rare cancers detection" published on July 22, 2024 (Link)

⁷¹ Harvey AI: "Harvey: Raises at \$11 Billion Valuation to Scale Agents Across Law Firms and Enterprises" published on March 25, 2026 (Link)

Harvey never needed: a liability posture the creator is genuinely willing to stand behind with capital, reinsurance, monitoring, post-market surveillance, and indemnification inside validated envelopes. The second vulnerability is even more existential. In legal AI, Harvey rides the models. In clinical AI, OpenAI just launched ChatGPT for Clinicians directly. The *substrate has decided to become the application*. The foundation model provider—the very ground on which Harvey-style businesses stand—can, at any time, climb out of the ground and occupy the market above it.

So the clinical AI equivalent of Harvey needs a deeper moat than workflow alone. Workflow is necessary, but not sufficient. The moat must include proprietary data, EHR integration, clinical relationships, specialty-specific validation, regulatory competence, distribution, and a liability-bearing deployment posture. It needs to *own the last mile of clinical trust* and *the first mile of clinical data*. It needs to become, in some sense, indispensable to both the god model below it and the health system beside it. Harder than it looked three years ago.

This is why I think the future will be one of *stratified coexistence*, at least for a while. The god models will dominate general clinical reasoning, broad clinician-facing assistance, patient-facing explanation, literature synthesis, and cross-domain integration. The specialists—Tempus, Paige, oncology-specific models, cardiology-specific models, radiology foundation models, maybe a handful of pathology and neuropsychiatry players—will hold their edge in molecularly specific or modality-specific tasks where proprietary data and clinical validation create real defensibility. Aidoc, where I serve as an advisor, will continue its pioneering ways. The workflow-and-distribution companies—OpenEvidence being the most obvious current example—will capture value by embedding into the clinician's actual daily epistemic pathway. And then the hyperscalers will periodically reach down, replicate functionality, acquire what matters, partner where they must, and commoditize what they can. That's the race. The specialists have a head start in curated data and vertical expertise. The god models have the compute, reasoning breadth, capital, distribution, need to grow into their parabolic valuations, and appetite to do so. The 150 have the bedside, the workflow reality, the clinical legitimacy, the patients, and the local data exhaust. The winners will be the companies and health systems that figure out how to join those assets before the center of gravity hardens somewhere else.

For the startup community, the implication is bracing. Stop telling yourself that your moat is "clinical expertise" if that expertise is mostly encoded in a prompt chain sitting on top of GPT, Claude, or Gemini. Stop telling yourself that your model is safe because it performs better on a narrow benchmark if the god model can eventually access the same data and distribute a "good enough" version for free. Stop telling yourself that physicians liking your UX is enough if you don't own data, workflow, integration, and liability. The clinical AI companies that survive will be unusually thick companies: thick with data, thick with distribution, thick with workflow, thick with regulation, thick with trust, thick with actuarial seriousness. Thin companies get eaten by predators.

Enterprise Doctrine for the 150

For the 150, the implication is equally important but different. Don't become passive consumers in this market. Don't let each department buy its favorite shiny clinical AI tool until you've recreated, at the AI layer, the same balkanized mess you already hate in your EHR and revenue-cycle stack. Don't build your clinical AI strategy around companies that can't survive the next frontier-model release. Don't sign vendor agreements that *push all liability back to your physicians* while the vendor hides behind "decision support" disclaimers and an ArXiv benchmark paper. And don't mistake model performance for deployment readiness. A clinical AI product isn't ready for enterprise deployment because it's clever. It's ready when it has data provenance, workflow integration, model-monitoring architecture, clinical validation, escalation pathways, equity evaluation, user training, post-market surveillance, and a liability posture.

That's why the partnership architecture matters so much. The 150 should partner with frontier labs where broad reasoning, enterprise-scale model access, and safety culture matter. They should partner with vertical specialists where proprietary data, modality-specific performance, regulatory depth, and specialty workflow create real advantage. They should partner with workflow companies where clinician adoption and EHR integration matter. But they should impose a *common enterprise doctrine* across all of it: no balkanized tool sprawl, no shadow clinical AI, no liability-free clinical recommendations, no isolated pilots without measurement, no procurement without data governance, no deployment without an explicit clinical envelope.

Clinical AI isn't arriving as a tidy procurement category. It's arriving as a new, foreign stack. At the bottom: compute, foundation models, and multimodal reasoning. In the middle: proprietary clinical data, specialty models, workflow layers, EHR integration, model selection, and regulatory depth. At the top: clinical deployment, liability, physician trust, patient relationship, and standard-of-care incorporation. The hyperscalers dominate the bottom. The best startups may own parts of the middle. The 150 still own much of the top. The strategic question is whether the top and middle get joined to the bottom by design, or whether the bottom simply rises and absorbs everything above it.

That's the medical invasion now underway. Not armies, not tanks, not some melodramatic hostile takeover of the clinic. A gravitational invasion. A capital invasion. A compute invasion. A reasoning invasion. A *liability invasion waiting to happen*. The god models are coming for medicine because medicine is where intelligence has the highest moral and economic return. The startups will either deepen, orbit, or disappear. Sam, Elon, Demis, Jensen, Dario and even Zuck aren't just talking about healthcare for the platitudes, they're getting serious. The 150 will either become deployment partners or substrate.

And that, finally, is why this matters for the healthcare CEO. You're not picking software vendors. You're choosing where your institution will sit in the new clinical intelligence stack.

Choose badly, and you become a passive integration surface for other people's models, other people's economics, other people's liability frameworks, other people's standard-of-care definitions. Choose well, and you become one of the institutions that helps clinical AI *become* medicine rather than merely another product launched *onto* medicine.

The god models (and the armada of startups) have entered the clinic. The question is whether the 150 will meet them there as architects, or wait to be renovated.

Cheap Tokens and the Vanishing Small-Model Lead

One more uncomfortable market-structure point belongs in the god-model section. The proprietary small foundation models—VIRCHOW, PRISM-like pathology models, radiology models, oncology-specific reasoning engines, and the other beautiful, carefully curated specialist intelligences—have a real lead today. I don't want to trivialize that lead. A model trained on millions of whole-slide pathology images isn't a toy, and the data curation, annotation, clinical-grade validation, and modality-specific inductive bias matter enormously. But the lead may be more ephemeral than its owners hope. The relentless, inexorable, remorseless improvement of the frontier models is the strategic weather. Once the gods obtain comparable clinical data, the small specialist model has to defend itself not merely against a peer, but against a planetary reasoning engine with cheaper inference, broader multimodality, deeper tool use, better distribution, and a balance sheet capable of swallowing the risk.

Harvey remains the precautionary tale, and also the compliment. It didn't win by insisting forever that its own foundation model would defeat OpenAI, Anthropic, Google, and everyone else. It rode the gods. It built workflow, customer intimacy, legal-engineering depth, model selection, and distribution. The analogous clinical company has to do the same, but with more blood on the floor if it fails. A healthcare startup can't merely be an elegant clinical wrapper. It needs proprietary data moving at velocity, validated workflow, regulatory competence, clinical trust, liability-bearing posture, and the last mile of adoption. Workflow plus foundation plus safety plus liability plus distribution is the five-fold, compounding stack that the god models do not yet have and may eventually be forced to acquire rather than replicate.

OpenEvidence is fascinating in exactly this way. I admire Daniel Nadler's ingenuity a great deal; he saw the physician as consumer before most of healthcare's enterprise sales machinery could even parse the sentence, and he built a product clinicians actually use. But unless OpenEvidence finds a durable tailwind that's defensible against the gods—proprietary clinical-data loops, workflow entrenchment, content rights, trust, distribution, liability, or some combination of all of them—it'll live under the gravitational shadow of ChatGPT for Clinicians, Gemini, Claude, and whatever comes next. The god models can generate the tokens more cheaply and distribute them more widely. That doesn't mean OpenEvidence loses. It may mean it becomes more valuable as a target, partner, or integrated layer inside a multi-trillion-dollar juggernaut. If Microsoft, Google, or another hyperscaler eventually distributes Nadler's ingenuity over a global clinical infrastructure, that would be a lucrative exit and perhaps, paradoxically, a more impactful outcome for the world.

The Aidoc example may be the more strategically sound path to durability. Aidoc has built an unusually deep moat around the unsexy but indispensable pillar of safety. They started with a domain-specific model and then made a bigger bet: CARE, a pan-modality, pan-disease clinical foundation model trained across CT, MR, and X-ray, with the EHR, labs, and longitudinal patient context wrapped around it. Tens of millions of cases, validated through live clinical deployment rather than benchmark theater. A technical marvel, but the model is only part of the value. In 2021, a year before ChatGPT debuted, Aidoc launched aiOS, the operational substrate that enables the largest clinical AI deployments in American medicine.⁷² An enterprise-grade platform built around model safety, performance monitoring, change management, validation, and the critical integrations that turn an algorithm into a clinically deployable agent. The VIRCHOW lead is real, but an architectural lead is structurally different. It isn't a model defending a dataset; it is a model defending a *deployed clinical operating system*.

⁷² Aidoc: "Aidoc Secures FDA Clearance for Healthcare's First Comprehensive Foundation Model AI" ([Link](#))

, which is a much harder thing for a frontier lab to commoditize from the outside. The durable shape of vertical clinical AI looks less like a single beautiful model and more like the platform that grows around one: deployment, safety, validation, accountability, distribution. Call it the inverse of Netflix. Netflix built the streaming platform first and used that advantage to mint a moat of original content. Clinical AI runs the other way: the model came first, but the winners will be the ones who build the platform to diffuse it.

Hippocratic AI, another player I'm an advisor to, is a similar kind of instructive case, and I mean this admiringly. Munjal Shah and team built toward a novel platform, not just a clever demo. *Polaris*, their healthcare-focused model architecture, emerged because the AI nursing and patient-voice platform had to move fast, had to be safety-shaped, and had to live in an operational care-delivery workflow rather than in a benchmark paper.⁷³ Whether every claim survives the brutal audit of clinical diffusion isn't the point here. The strategic point is that Hippocratic built something thick: workflow, voice, safety architecture, role-specific agents, customer deployment, and a posture toward actual labor substitution and augmentation. That's what capital, urgency, and ingenuity should do. Build something the gods might want to partner with or buy, not something the gods can recreate over lunch.

And Google deserves its own paragraph because *cost per clinical token will matter* far more than most healthcare boards realize. A clinical AI system that lives inside every encounter, every refill, every home sensor stream, every care-management note, every pathology image, every message, and every medication adjustment is going to produce and consume a staggering number of tokens. The winner won't merely have the best answer. *The winner will have the cheapest safe answer at scale.* Google's tensor processing units, its ownership of more of the hardware and compiler stack, its datacenter discipline, its Gemini/DeepMind research complex, and its capacity to optimize inference economics give it a structural advantage in generating clinical AI tokens cheaply. Not magically. Not inevitably. But structurally. In a world of ubiquitous clinical agents, cost per token becomes cost per unit of medicine.

Somebody Has to Volunteer to Get Sued

Blame Allocation Is the Gate

Ok, we've finally arrived at the central argument (only took a few thousand words!). Everything before this section was setup. Everything after it is consequence. To sum up: the old scarcity has ended. Biology has outrun the biological brain. The 17-year innovation graveyard is no longer morally defensible. Medicine is approaching its AlphaZero moment. The god models are coming for the clinic. The startup community is discovering that pure clinical software is a much thinner moat than it looked three years ago. All of that leads here, to the gate in the wall. The bottleneck for clinical AI isn't capability. It isn't model quality. It's not even regulation in the abstract, though the regulatory ambient is very much part of the picture. *The bottleneck is liability. Or, more precisely, blame allocation:* the unresolved, post-industrial-revolution-deep, civilization-shaping question of who's responsible when the machine errs.

Our legal and regulatory systems simply weren't designed for exponential AI. They were designed for drugs that take decades to develop, devices that emerge slowly and incrementally, human physicians who move at biological rather than silicon speed, and granite, century-old hospitals and health systems operating inside relatively legible chains of causation. A drug has a label. A device has a

⁷³ Hippocratic AI: "Hippocratic AI Launches Polaris 5.0: The First Evidence-based AI for Healthcare Proven to Outperform Major Frontier Models on Critical Medical Tasks and Safety" published on April 30, 2026 (Link)

manufacturer. A surgeon has a hand. A hospital has a policy. A malpractice carrier has a rate table. The system is messy, but the mess has a familiar geometry. Clinical AI breaks that geometry. A *continuously (let alone recursively)* updating frontier model doesn't sit obediently inside the old categories. It may be assistive in one context, quasi-autonomous in another, invisibly embedded in a workflow in a third, and updated by a distant developer (or by the AI itself) before the local medical staff fully understands what changed. The question that freezes deployment is as simple to state as it's maddening to answer: who, exactly, is responsible when the AI makes an error?

We can only go at the speed of blame allocation. I'd like that on a slide somewhere. In 16-point font, ideally. Maybe 24-point if the board is 53 or older.

And this isn't just the Stanford 2024 landmark study anymore, though that framing was useful. The more recent 2025–2026 policy conversation has basically converged on the same anxiety. HHS' December 2025 clinical AI RFI explicitly asked what role the department should play in addressing novel legal and implementation issues for non-device AI, including liability and indemnification, and also asked how regulation, payment policy, and evaluation should change to accelerate clinical AI adoption.⁷⁴ The 2025 JAMA Summit report likewise emphasized that AI's effects are often hard to quantify, highly dependent on the human-computer interface, user training, and deployment setting, and that the system needs better mechanisms for evaluation, monitoring, and incentives.⁷⁵ The AHA's February 2026 response to HHS then said the quiet institutional part aloud: regulatory barriers, privacy-law patchworks, and unclear frameworks can inhibit hospital and health-system deployment, and clinical AI adoption needs synchronized policy, clinician involvement, and clearer federal posture.⁷⁶ In other words, our guidelines are anarchic, confused and contradictory.

So yes, the Stanford formulation still stands, but now it has company. The legal literature keeps circling the same problem: healthcare AI liability is complex because responsibility may sit with physicians, institutions, developers, software vendors, device makers, payers, malpractice carriers, and reinsurers, while causation becomes harder to narrate when the model is opaque, probabilistic, and embedded in a local workflow. The law remains underdeveloped because the field is still new, because explicit court treatment is scarce, and because no one wants to become the test case that teaches the rest of us what the doctrine means.

Clinical AI stalls not because the models lack capability, but because no one in the system wants to be left carrying the liability bag when the machine errs. The doctor doesn't want to defend why she trusted the recommendation. The hospital doesn't want enterprise exposure from a probabilistic system operating across thousands of encounters. The developer fears product liability. The device manufacturer fears regulatory exposure. The payer fears reimbursement and coverage disputes. The malpractice carrier fears actuarial uncertainty. The reinsurer fears correlated systemic risk. Each actor sits inside a different legal and economic framework, but they converge on the same rational conclusion: the downside is immediate, personalized, and potentially catastrophic, while the upside remains diffuse, collective, and uncertain. *So everyone waits.*

That's what a lawyerly society does when it encounters an exponential technology. It freezes, points at its own statutes, and waits for someone else to take the risk. Engineers build. Lawyers litigate, proceduralize, allocate, defer, and say no eternally. I'm not saying that with total contempt. Laws, rights, due process all matter. Harmed patients deserve recourse. And God save us from Silicon

⁷⁴ U.S. Department of Health and Human Services: "HHS Announces Request for Information to Harness Artificial Intelligence to Deflate Health Care Costs and Make America Healthy Again" published on December 19, 2025 ([Link](#))

⁷⁵ JAMA: "AI, Health, and Health Care Today and Tomorrow: The JAMA Summit Report on Artificial Intelligence" published on October 13, 2025 ([Link](#))

⁷⁶ American Hospital Association: "AHA Response to HHS RFI on AI in Health Care" published on February 23, 2026 ([Link](#))

Valley technologists who think tort law is merely a bug in the operating system of progress. But the legal architecture that protects patients can also become the blocker that prevents better care from reaching them. In clinical AI, that's exactly the danger.

This is why the *assistive-versus-autonomous* spectrum matters so much. Much clinical AI will remain assistive for a while: summarizing records, surfacing risk, suggesting differentials, flagging possible deterioration, drafting care plans, reconciling medications. In those cases, shared responsibility makes intuitive sense because the physician is still supervising, contextualizing, and deciding. But as systems move toward functionally autonomous diagnostic or treatment determinations inside validated envelopes, the old fuzzy allocation becomes untenable. If the model is effectively making the call, if the developer controls the design and update logic, if the hospital deploys the tool according to its intended use, and if the clinician is no longer in a realistic position to independently reconstruct the model's reasoning, then liability has to migrate toward the entity best positioned to understand and manage the risk.

That should be the governing principle: *assign liability to the party best positioned to control the failure mode*. In cars, that increasingly means the manufacturer when the autonomous system is in control. In clinical AI, it should increasingly mean the developer, device maker, software creator, or liability-bearing platform when the system is operating inside its validated envelope. *The more informed party should shoulder more responsibility*. That's not anti-innovation; it's the precondition if we want innovation to scale. If everyone disclaims responsibility, nothing truly clinical will deploy. If someone assumes responsibility, the market can finally move.

So yes, we need safe harbors for clinicians who use approved AI tools in good faith. We need clearer federal preemption or harmonized standards so fifty state tort regimes don't suffocate deployment (a real risk, by the way). We need FDA post-market monitoring suited to adaptive models rather than static devices. We need product-liability doctrine that recognizes software can cause physical harm. We need reinsurers and malpractice captives willing to price the risk instead of treating foundation-model liability as some metaphysical horror. And we need the 150 to stop treating liability as something that happens to them and start treating it as something they can help architect.

For the 150, liability isn't something to fear passively. It's something to architect. If health systems remain mere receivers of vendor disclaimers, medical-board caution, and malpractice-carrier anxiety, they'll be acted upon. If they define validated envelopes, instrumentation rules, safe-harbor conditions, indemnification requirements, escalation pathways, and captive-insurance strategies, they become co-authors of the deployment regime.

Until then, clinical AI remains trapped in pilots, disclaimers, advisory language, and performative caution. Everyone will praise the promise. Everyone will admire the benchmark. Everyone will sit on the panel. Everyone will say "guardrails" with the necessary solemnity. And then the thing won't deploy at scale, because no one has volunteered to get sued.

I'll say it again: the current system isn't safe. It's merely familiar.

The Sandbox Regime: Clinical AI Trials at Machine Tempo

The Validation Gate

The liability gate has a second companion gate: validation tempo. We need an IRB-like clinical-trials architecture for clinical AI, but radically compressed, continuously learning, and increasingly in silico. The old trial stack was designed for drugs and devices that moved slowly, changed rarely, and entered the body as relatively stable interventions. Clinical AI is different. It updates, reasons, personalizes, escalates, watches, routes, summarizes, and sometimes recommends. Evaluating it once, freezing it, and pretending the work is done is as incoherent as certifying a resident after one morning of rounds and then never observing her again.

So build sandboxes. I want that word to become banal inside every serious health system. A clinical AI sandbox isn't innovation theater. It's a governed simulation-and-deployment environment where models are stress-tested against historical cases, synthetic cases, adversarial cases, local workflows, equity strata, edge cases, and prospective shadow-mode deployment before they touch the patient autonomously. It should feel like an IRB, a flight simulator, a post-market surveillance system, a morbidity-and-mortality conference, and an AlphaZero self-play environment had a very nerdy child.

The analogy to AlphaGo and AlphaZero isn't ornamental. Monte Carlo tree search and reinforcement learning made sense in Go because the system could simulate enormous numbers of possible games against a clear reward function. Medicine is harder because the reward function is plural, delayed, ethical, expensive, and sometimes contested. But as we build biological twins, patient-state simulators, organoid-linked feedback loops, digital phenotypes, synthetic controls, and longitudinal outcome datasets, more of clinical AI evaluation can move into simulation before it moves into the frail human body. We won't replace reality. We'll amortize risk before reality has to adjudicate it.

A clinical AI sandbox should run trillions of simulated or retrospective encounters where feasible: What would the model have recommended? What did the clinician do? What happened? Where did the model under-call risk? Where did it over-call? Which subpopulation suffered? Which medication combination produced unexpected benefit? Which social determinant converted a medically correct plan into a practical impossibility? Which escalation threshold increased false positives but prevented the deaths we actually care about? That's the A/B testing logic of internet platforms, but sanctified, slowed just enough, governed, and disciplined by the moral asymmetry of clinical harm.

This is also where biological twins become more than a fashionable phrase. A biological twin isn't merely a dashboard representation of a patient. It's a dynamic, probabilistic, updating model of that patient's physiology, behavior, medication exposure, environment, risk, preferences, and trajectory. The better the twin, the more we can simulate counterfactuals: what happens if we intensify therapy, de-intensify therapy, switch medications, escalate behavioral outreach, alter diet, initiate home monitoring, send the nurse, or do nothing? The simulation must prove itself against reality, repeatedly and ruthlessly. But once it starts to prove itself, the trial cycle compresses from years to weeks, weeks to days, and eventually days to continuous learning.

This is the clinical AI trial framework the 150 should begin building now: retrospective validation, synthetic stress testing, shadow deployment, narrow prospective deployment, instrumented clinician override, equity analysis, adverse-event review, model-version tracking, and continuous post-market learning. Not one pilot. Not one committee blessing. A standing trial machine. A *living sandbox*. A clinical-learning foundry. If we can *simulate* before we expose, and then learn continuously after we

expose, we can move faster without pretending speed is safety.

Part V—Resistance, Risk, and Inversion

This part maps the opposition and the true gating problem. The guild will call it safety; some of that will be right, and some of it will be old-scarcity protectionism. The central claim is that clinical AI can only move at the speed of blame allocation, and the standard of care will eventually invert.

The Infallibility Trap

I'm going to be a little philosophical in this section, because before we get to the guild response to clinical AI, we need to name the philosophical mistake that makes the guild response sound more persuasive than it should. *We in America hold technology generally, and clinical AI specifically, to an impossible and hypocritical standard of perfection.* We expect the machine to be *infallible* in a way no human being has ever been infallible, which is, quite literally, an inhuman standard. We are, after all, all too human: tired, distracted, overworked, associative, emotionally textured, occasionally brilliant, occasionally mediocre, and permanently reliant on rather spotty biological hardware. Sorry, a little uncharitable, but directionally true. The physician forgets, anchors, overweights the recent case, misses the weird edge signal, gets interrupted, gets sued, gets burned out, gets divorced gets older, and sometimes gets it wrong (sorry, that comes across darker than I intend, but you get the point). We know this. We've normalized it. We've professionalized it. We've insured it. We've built a whole malpractice jurisprudence around the concept of reasonable, tolerable, tragic, but non-criminal human error.

And yet when an AI system errs, we treat the error as a category violation. Not merely a mistake, but a scandal. The machine crossed a line. *The machine violated the covenant.* The machine trespassed on human authority. The machine must be eliminated. That instinct is emotionally understandable, but it's analytically incoherent, morally dangerous, and deeply hypocritical. The human-only baseline isn't some Edenic sanctuary of safety into which clinical AI rudely intrudes. The baseline is already imperfect. Deeply imperfect. The relevant question isn't whether AI makes mistakes. Of course it will. The relevant question is whether, inside a validated use case, it makes *fewer* mistakes than the human alternative, at the same or lower resource intensity, with better monitoring, better auditability, and a clearer ability to improve.

That has to be the standard: *human equivalence or human superiority inside a defined clinical envelope*, with transparent validation discipline and post-diffusion monitoring. Not infallibility. Not perfection. Not zero bad cases. Not no lawsuits. Not no ugly headlines. If a clinical model meets or exceeds the median human—or increasingly the top-decile human—in a bounded domain, then the moral question begins to change. The burden shifts from "How dare you use it?" to "How dare you *not*?" Perfection is the enemy of deployment, and deployment is the only way to learn. Sam is right about this one: *iterative deployment* has merit even, or perhaps especially, in something as sacred and consequential as health. The alternative isn't risk-free. The alternative is the status quo, which is already dangerous, already inequitable, already expensive, already error-prone, and already morally compromised by its own familiarity.

This is where American anti-risk culture starts to look rigged in favor of incumbency. If every insurgent clinical AI tool has to prove perfection before getting onto the field, while the existing system is allowed to continue under the softer doctrine of reasonable human fallibility, then we haven't designed a safety regime. We've designed an incumbency-protection regime. We've wired the game so the establishment wins by default, or perhaps more precisely, we've allowed the officials to keep

rewriting the rules whenever the insurgent gets too close to the line of scrimmage. That's not a serious posture for a civilization that wants progress. It's a gerontocratic, lawyerly, risk-averse posture pretending to be ethics.

Scale Risk Cuts Both Ways

There are, of course, serious and legitimate objections to my argumentative TED-talk in the previous paragraphs. Let's take them seriously, because this chapter can't become a cartoon of accelerationism. The first objection is scale. *AI errors can propagate at scale* in a way human errors generally don't. A physician misdiagnoses one patient. A model deployed across a million encounters with a systematic flaw could harm a very large number of patients very quickly. That's categorically different from individual human error, and anyone who hand-waves this away shouldn't be allowed near a clinical deployment plan. But scale cuts both ways. *The same AI that fails at scale also succeeds at scale*. A model that improves diagnostic accuracy by even a few percentage points across a million encounters prevents tens of thousands of errors that would have occurred under the human-only baseline. Given the current baseline of diagnostic error, that's an extraordinarily high bar for the status quo to clear.

And the answer to scale risk isn't nondeployment. It's *monitored deployment* inside a validated envelope, with model-performance surveillance, escalation pathways, uncertainty flags, adverse-event review, equity monitoring, post-market discipline, and a liability structure that gives the responsible party an economic reason to fix what breaks. A hyperscaler, a robustly capitalized clinical AI startup, an enlightened payer, or a health system with a serious malpractice captive and a real liability posture isn't inherently a threat to safety. Quite the opposite. *Concentrated, disambiguated, well-insured liability is the precondition for clinical AI*. The worst world isn't one in which someone responsible owns the risk. The worst world is the current fog: fragmented tools, disclaimers everywhere, responsibility nowhere, physicians nominally "in the loop" but not actually able to reconstruct the model's reasoning, and vendors insisting everything is mere decision support until the deposition begins.

The second objection is the catastrophic case. The hyperscaler, startup, hospital, or payer that assumes clinical AI liability achieves dominant market penetration, then suffers a single catastrophic, clearly attributable AI failure—a death, perhaps several deaths—that triggers not merely litigation, which reinsurance can absorb, but legislative and regulatory backlash that sets the whole field back five or ten years. The Chernobyl effect. The Three Mile Island effect. The Vioxx effect. The Cruise effect. This is undeniably real, especially in a hyper-litigious society that metabolizes machine failure much more indignantly than human failure.

But here again the argument cuts in the opposite direction from the one the professional guilds will prefer. We've stress-tested this pattern across aviation, automotive, pharmaceuticals, nuclear energy, medical devices, and other high-stakes industries. Catastrophic failures produce legislative shock, reputational damage, litigation, and *years* of recovery. They also often produce better safety frameworks rather than permanent prohibition. Vioxx didn't end drug approvals; it forced a more sober understanding of post-market risk. Aviation disasters didn't end flight; they built one of the most impressive safety cultures in the modern world. Nuclear accidents didn't prove that splitting the atom was intrinsically impossible; they exposed the design, governance, and institutional failures that made certain deployments unsafe. The lesson isn't "never deploy." The lesson is "deploy inside a system capable of learning faster than catastrophe can metastasize."

And the catastrophic-case risk actually argues *for* liability assumption, not *against* it. The worst possible scenario is fragmented, surreptitious deployment where twenty different point solutions

operate without clear liability assignment, one fails catastrophically, and no entity is positioned to be held accountable or forced to fix the system. That's a litigation lottery. A liability-bearing actor—hyperscaler, startup, payer, hospital coalition, whoever—has the incentive to monitor, improve, insure, and contain the failure because it owns the downside. That's why the BYD analogy (coming up) matters. Better one accountable entity with capital and confidence than a fog bank of disclaimers.

And this is what everyone seems to forget: *the status quo is already its own catastrophic case*. The hundreds of thousands of deaths and permanent disabilities associated with diagnostic error aren't hypothetical. They aren't a scenario analysis. They're the baseline we're protecting. The catastrophic failure of clinical AI, if it comes, will be visible, attributable, narratable, and correctable. The catastrophic failure of clinical inertia is invisible, distributed, normalized, and tolerated. That doesn't make it less catastrophic; it only makes it less politically obvious.

So no, the standard can't be infallibility. The standard has to be human equivalence or superiority, inside a validated envelope, with relentless monitoring and a clear allocation of responsibility. If we can't accept that, then we're not protecting patients from technology. We're protecting the incumbent system from comparison.

The Guild Will Call It Safety

Let me state this as plainly as I can: DO NO HARM can't mean DO NOTHING. The Hippocratic injunction isn't a permission slip for institutional obstructionism. Non-maleficence includes the harms of omission, delay, access failure, missed diagnosis, stale standard of care, unavailable psychiatry, and the resigned abandonment of patients to a familiar but broken system. The guild will prefer to define harm as the visible machine error and ignore the invisible human-system error. That definition is morally rigged. It protects what can be narrated in a lawsuit and neglects what accumulates in cemeteries, disability rolls, bankruptcy courts, and kitchen tables.

The historical analogy I keep circling is the protective guild whose original function was real, even noble, until the technology moved past it. The scribal order protected textual accuracy before movable type; after Gutenberg, the protection became partly scarcity maintenance. London's black-cab Knowledge protected navigational competence before GPS; after ubiquitous mapping, it became less obviously a safety necessity and more obviously a prestige and access moat. Medicine isn't manuscript copying or taxi dispatch, and patients aren't parcels. But the pattern is recognizable: *a guild builds legitimate trust under conditions of scarce expertise*, then struggles to distinguish trust from monopoly when the expertise becomes technically distributable. That's the danger now.

Once you see the infallibility trap, the guild response becomes easier to decode. The first serious opposition to clinical AI won't announce itself as guild protection. It won't say, "We're protecting prestige, compensation, epistemic authority, professional scarcity, and the old ordination structure of medicine." It'll sound much more seductive than that. Much nobler. Much more enlightened. It'll call itself patient safety, prudence, ethics, equity, oversight, professional standards, responsible innovation, and protection of the vulnerable. And some of that language will be sincere. Some of it will be right. But a lot of it will function, in practice, as delay. And delay isn't morally neutral when the baseline we're protecting is already saturated with error, waste, fragmentation, disability, and preventable death.

And this is where the guild becomes dangerous, because the guild will sound right. It will say the tool isn't validated enough. The patient relationship is too sacred. The model is opaque. The consent

language is insufficient. The regulator hasn't spoken. The medical board hasn't been consulted. The professional society hasn't issued guidance. The liability framework is immature. Again, some of this will be true. I'm not arguing for reckless deployment, and I'm certainly not arguing for "move fast and break patients." But the aggregate effect of these objections, if left unchallenged, will be to preserve the incumbent order by dressing professional self-protection in the vestments of patient safety. We should respect legitimate caution. We should also call out capture when caution becomes capture.

Behavioral Health as Tell

Behavioral health is a sad clarifying example, and one I care deeply about as an investor, strategist, and simply as a citizen of this democracy: because the need is so desperate and the regulatory instinct so revealing. Governor JB Pritzker signed preposterous legislation restricting the use of AI in mental-health therapy in 2025, barring licensed therapists from using AI for treatment decisions or client communication and restricting companies from offering or advertising AI therapy services without licensed-professional involvement.⁷⁷ The stated concern is safety. Fine. Safety matters. But this is also a country where huge areas have inadequate psychiatric access, where loneliness, depression, anxiety, addiction, and adolescent distress are everywhere, and where millions of people are already turning to AI systems for emotional support, companionship, symptom interpretation, and late-night conversational holding. The use case doesn't disappear because Illinois disapproves. It migrates into the shadows, outside clinical integration, outside instrumentation, outside supervision, outside any enlightened architecture of escalation and safety. Pritzker may mean well (although I think in this case he just cravenly caved to the unions). The impulse may be protective. But the effect is to stand in opposition to one of the most powerful potential access technologies in behavioral health and shout stop, while the patients simply route around the law. (Please see my behavioral chapter for more on why ChatGPT may actually be a fantastic therapist).

So that's the pathology. A blanket posture of prohibition can feel morally superior to supervised deployment, but it may be less safe in practice. If patients are going to use AI companions and therapeutic agents anyway—and they are—the real question is whether these systems get integrated into a clinical, ethical, monitored, escalation-capable framework, or whether we pretend to ban them while vulnerable people use them alone at 2 a.m. without our knowledge, consent, or integration into a broader, unified care plan.

This pathology isn't confined to the United States. I wrote about this last year: South Korea offered a guild-politics preview in 2024, one that I still think prefigures what might happen here in the US. The government proposed expanding medical-school admissions to address physician shortages, including rural and essential-care shortages; physicians and trainees responded with mass walkouts and protests, causing hospital disruption and public backlash. Public support for the striking doctors cratered. That matters because it shows what happens when professional scarcity becomes too visible. The public will tolerate credentialed authority for a long time. It'll tolerate scarcity for a while. It'll tolerate paternalism when it believes the profession is acting as guardian of the patient. But *once the guild looks like it's protecting its own monopoly* against access, cost, and measurable performance, the moral prestige of the profession starts to erode.

This is coming to the United States. As clinical AI advances, medical guilds will increasingly defend their monopoly with patient-safety arguments that the public, employers, payers, and eventually the data will find less persuasive. I say this with sadness, not contempt. The AMA, specialty societies,

⁷⁷ Illinois Department of Financial and Professional Regulation: "Gov Pritzker Signs Legislation Prohibiting AI Therapy in Illinois" published on August 4, 2025 ([Link](#))

medical boards, academic medical centers, malpractice carriers, and clinical professional associations have stewarded real wisdom over generations. They're not villains, quite the opposite. But their institutions are built to protect an old equilibrium, and it seems they struggle to distinguish wisdom from self-preservation when the equilibrium changes. Medicine has a great deal to lose from superior machine intelligence: prestige, compensation, social esteem, interpretive authority, epistemic primacy, and the old sovereign status of unaided clinical judgment. So of course it'll resist. But resistance becomes morally compromised when the technology being resisted can outperform unaided clinicians in bounded domains, expand access where humans are scarce, and reduce harm relative to the actual baseline. At some point, professional self-protection becomes a form of patient harm.

Utah gives us the crisp, contemporary, domestic illustration. In October 2025, the Utah Office of Artificial Intelligence Policy approved a regulatory mitigation agreement with Doctrionic for an AI-enabled prescription-renewal pilot.⁷⁸ I was frankly thrilled about this. The pilot was deliberately narrow: prescription renewals, routine refills, patients who already had valid prescriptions from a human physician. Exactly the kind of bounded, low-acuity, validated envelope one should use to begin. *The Utah Medical Licensing Board* called for suspension.⁷⁹ The language was, predictably, the language of risk and patient safety. But the underlying reflex was also discernible: the guild wasn't consulted, and the guild didn't approve. Utah dismissed the call. The pilot continues.

That's how the Overton window moves. Not through grand declarations about AI medicine. Not through another responsible-innovation symposium with a barely-edible continental breakfast. Through bounded deployments that force the old objections to become specific. Prescription renewals. Diabetic retinopathy. Mammography second reads. Medication adherence. Low-acuity refills. Post-discharge check-ins. Depression screening with escalation protocols. Sepsis prediction with clear human response pathways. Narrow use cases aren't timid. They're how the profession learns, how the liability system gathers experience, how the public sees benefit, and how blanket objections begin to look self-interested and performative.

The future of medicine can't be vetoed indefinitely by the ancien régime of medical societies, committees, and credentialed incrementalism. Safety is essential. But safety can't be allowed to become the rhetorical cover for conservatism and stasis. The morally serious posture is neither reckless acceleration nor guild-protective delay. It's supervised deployment inside validated envelopes, with instrumentation, escalation, indemnification, post-market monitoring, and brutally honest comparison to the human baseline. Hold clinical AI to human equivalence or human superiority, not to machine infallibility. Protect patients from bad AI, yes. But protect them also from the old system's delays, misses, access failures, and invisible harms.

That's the distinction the 150 have to learn to make. Every time someone invokes safety, ask: *compared to what?* Compared to the idealized hospital in the policy memo, or compared to the real one—with missed diagnoses, unavailable psychiatrists, 17-year diffusion lags, fragmented care, exhausted clinicians, and patients serving as their own integration layer? Once you ask the question that way, the moral terrain changes. Clinical AI becomes less a threat to safety than a test of whether our institutions are still capable of recognizing opportunity when it arrives.

⁷⁸ Utah Office of Artificial Intelligence: "Doctrionic AI Regulatory Mitigation Agreement" published on October 24, 2025 ([Link](#))

⁷⁹ STAT News: "Utah medical board calls for immediate suspension of state's AI doctor experiment" published on April 24, 2026 ([Link](#))

The Lawyerly Society Meets the Engineering State

I've alluded to the Sino-American competitive dynamic repeatedly—relentlessly? tediously?—throughout this essay, but here it lands with particular force. The previous sections were about blame allocation, the infallibility trap, and the guild response. This section is about what happens when one civilization treats those as central gating questions, while another treats clinical AI as an engineering, state-capacity, and diffusion problem. That's the whole divergence. In the United States, the question is: who's liable? In China, and to a similar extent in the GCC, the question is: how fast can we install it?

I mention this in my China chapter, and I'll raise it again here: Dan Wang's juxtaposition between China as an "engineering state" and America as a "lawyerly society" keeps reappearing because it's such a useful and uncomfortable frame. China builds, scales, installs, tolerates messiness, and deals—often coercively, sometimes invisibly—with the downstream consequences. America reviews, litigates, proceduralizes, equilibrates among veto points, convenes stakeholders, and produces a memo explaining why the pilot needs another pilot. I'm exaggerating, obviously, but only slightly. And in clinical AI, that cultural difference may determine where the first large-scale health-improvement benefits actually accrue.

China's Clinical AI Diffusion

Reader Note: This China / engineering-state material is developed fully in Chapter 9 and foreshadowed in Chapter 1's installation argument. I include it here because clinical-AI liability and diffusion aren't only domestic governance questions; they're competitive-installation questions.

China is already sprinting. Tsinghua's Agent Hospital isn't a perfect analog to a U.S. academic medical center deploying autonomous clinical AI in the wild, and we shouldn't pretend it is. Much of it is still virtual, simulated, staged, and designed for training, research, and controlled pilot operations. But that's precisely the point. Tsinghua isn't waiting for every doctrinal question to be adjudicated before building the learning environment. Its Institute for AI Industry Research describes *Agent Hospital* as a virtual hospital in which patients, nurses, and doctors are represented by autonomous agents simulating the entire care process—pre-hospital, in-hospital, post-hospital, triage, registration, consultation, examination, diagnosis, prescription, rehabilitation, and follow-up. Its AI doctors diagnosed nearly 10,000 virtual patients and achieved 93.06% accuracy on a MedQA respiratory-disease subset; Tsinghua later formally inaugurated the AI Agent Hospital and said it would proceed through phased pilot operations at Beijing Tsinghua Changgung Hospital and its Internet Hospital, beginning with general practice, ophthalmology, radiology diagnostics, and respiratory medicine.^{80 81} *That's the engineering-state posture:* build the closed-loop system, simulate the workflows, train the agents, run pilots, move toward physical deployment, and use state-aligned institutions to compress the experimentation cycle.

⁸⁰ Institute for AI Industry Research, Tsinghua University: "AIR Creates a Virtual Hospital, Enabling AI Doctors to Self-Evolve" published on May 24, 2024 ([Link](#))

⁸¹ IT Home: "Tsinghua University Establishes AI Hospital to Cultivate a New Generation of 'AI Collaborative Doctors'" published on April 28, 2025 ([Link](#))

Alibaba is moving through the consumer and medical-assistant layer. *Quark Health*, powered by Alibaba's Qwen family, reportedly passed China's deputy chief physician qualification exam across twelve common medical disciplines, reached chief-physician qualifying scores in several fields, and was integrated into Quark, Alibaba's flagship consumer-facing AI assistant app.⁸² Quark itself is being repositioned as Alibaba's AI consumer platform, with Qwen-powered search, chat, reasoning, voice, and task execution. Again, this isn't the same thing as a fully autonomous physician treating real patients without oversight. But it's exactly how diffusion begins: consumer interface, medical reasoning layer, hospital and institutional partnerships, exam benchmarks, public familiarity, and a national technology platform that can push functionality into the daily lives of hundreds of millions of people far faster than an American hospital committee can approve a vendor questionnaire.

So no, I'm not espousing a CCP diffusion timetable or its political ethos. Let's be clear before someone clips this paragraph for a board packet and performs the usual ritualized outrage. I don't want American medicine imported into an authoritarian model. I don't want adverse consequences sublimated, denied, or absorbed by citizens without recourse. I don't want safety theater replaced by state-capacity theater. *But I do want us to look soberly at the counterfactual.* If China can move clinical AI from simulation to pilot to broad consumer and hospital deployment faster than we can, it will generate real-world data faster, iterate faster, train models against more patient interactions faster, normalize AI-mediated care faster, and *perhaps capture health-improvement benefits faster.* The advantage won't merely be technical. It will be *cumulative*. Diffusion creates data. Data improves models. Better models justify more diffusion. More diffusion creates more political and clinical legitimacy. That flywheel compounds.

That's what should make American healthcare profoundly uncomfortable. The United States may retain the most powerful AI research infrastructure on earth. The preeminent foundation models are disproportionately American. The scientific talent is disproportionately American or America-trained. The capital markets, especially for early-stage venture, at least for now, remain vastly superior on this side of the Pacific (and Atlantic, for that matter). But invention isn't installation. Invention isn't diffusion. Invention isn't adoption. The nightmare isn't that America fails to invent the medical brain. The nightmare is that America invents the medical brain, then loses the clinical diffusion war, then licenses superior medical intelligence back from regimes and sovereign-capital ecosystems that moved faster because they were less lawyerly, less guild-captured, less procedurally tangled, and more willing to let the technology run.

The American legal tangle matters here not because it's intellectually interesting—though heaven help us, to some people it is—but because it creates undue delay. The precise doctrine matters less than the cumulative effect: preemption uncertainty, software-as-product ambiguity, attribution fog, fifty state tort regimes, malpractice exposure, FDA uncertainty, hospital risk aversion, medical-board caution, payer incentives, and the generalized American instinct to treat every new capability as a liability event waiting to happen. Fear of lawsuits plus regulatory ambiguity isn't just a nuisance. It's a cripplingly expensive *innovation tax*. And innovation taxes of this magnitude are how frontiers migrate.

The AI system that might be piloted carefully in Boston gets installed more aggressively in Beijing. The workflow that would spend eighteen months in legal review at an American academic medical center gets tested under sovereign mandate in Abu Dhabi or Riyadh. The population-scale data flywheel that could have belonged to American institutions starts spinning somewhere else. We

⁸² South China Morning Post: "Alibaba's healthcare AI model scores as high as senior-level doctors in medical exams" published on May 28, 2025 (Link)

flatter ourselves that delay is prudence, but from the perspective of technological hegemony, delay is often forfeiture.

Installation is the differentiator. This is one of the leitmotifs of this entire essay, and it lands here with particular force. Across major technological revolutions—the railroad, electrification, the internal combustion engine, mass computation—the country that invented the technology wasn't always the country that captured the full long-run economic and civilizational benefit. Britain gave the world the steam engine; America industrialized with a vengeance. America invented the integrated circuit and the internet; East Asia built the manufacturing substrate and is now contesting the deployment substrate. The lesson isn't that invention is irrelevant. *The lesson is that invention without diffusion is an unfinished act.* Whoever installs the advance most horizontally, most comprehensively, and most economically wins the compounding benefits.

The disagreeable truth, the one I've circled in chapter after chapter and sometimes been reluctant to name with the force it deserves, is that autocracies are structurally better than messy, pluralistic, lawyerly democracies at the installation phase of a major technological revolution. Centralized authority. No vetocracy. No fifty-state tort regimes. No specialty societies with letterhead, lobbyists, and a Pennsylvania Ave. address in DC. No five-year procurement cycles. No academic committees mistaking deliberation for virtue. The Gulf monarchies and the CCP can simply decree that clinical AI is a national priority and push it across the system on a timeline an American academic medical center couldn't contemplate without 100 committees.

Again, this isn't admiration; it's analysis. Some autocracies may also hide errors, suppress dissent, flatten individual rights, and sometimes mistake speed for wisdom. Their tolerance for adverse consequences isn't a virtue. The Chinese model can sublimate failure in ways that should make any liberal democrat hesitate. But if we refuse to study the diffusion advantage because we dislike the politics, we're committing the same kind of moralized analytic error that has made American institutions so slow in the first place. The point isn't to import their politics. The point is to learn from their installation capacity before our own patients pay for our procedural narcissism.

That's what's unfolding in front of us. I wrote this last year: **America innovates. Europe regulates. China appropriates.** And, I should add, deploys (although that destroys my little 'ates' symmetry, much to my authorly chagrin). The European Union, predictably, has responded to the AI revolution with the *AI Act*—the precautionary, self-defeating, timid principle dressed as digital sovereignty. China isn't having the same tremulous conversation. China is building agent hospitals, consumer medical assistants, national model infrastructure, and hospital pilots. The Gulf states—UAE, Saudi Arabia, Qatar—are deploying AI-augmented clinical infrastructure under sovereign-led mandates, with sovereign-wealth capital behind them and far fewer procedural veto points. Cleveland Clinic's partnership with G42 isn't incidental (again, have a listen to my podcast with Cleveland Clinic CEO Tom Mihaljevic for color on this). It's what early reverse importation looks like: a world-class American clinical institution recognizing that some of the future may be built in jurisdictions where capital, data, regulatory posture, and state capacity allow faster movement.

The health consequences of that divergence could be enormous. Clinical AI isn't only an economic prize; it's *a population-health prize*. Earlier diagnosis, fewer missed cancers, faster deterioration detection, better medication management, more accurate triage, more continuous chronic-disease monitoring, broader psychiatric access, faster guideline diffusion, and *national deflation in health-care costs* are anything but trivial benefits. If China and the Gulf diffuse these capabilities sooner, they may capture years of real-world learning and health improvement while we're still debating liability allocations in CLE webinars. Their models will see more cases. Their systems will learn from

more interactions. Their clinicians and patients will become culturally habituated to AI-mediated care sooner. Their regulators will gain experience supervising the technology in practice. Their institutions will begin to understand what works and what fails, much faster than our institutions will.

And the United States? We'll still have extraordinary science. We'll still have brilliant frontier labs. We'll still have the best acute care in the world for those who can reach and afford it. But we risk ceding the deployment advantage—the clinical diffusion, the real-world data flywheel, the population-scale learning, the institutional normalization, the GDP-level healthcare deflation—to regimes less litigious, less captured by guild interests, and more willing to tolerate the discomfort of iteration. The nightmare isn't that America lacks ingenuity. *The nightmare is that we become a country that invents the future but can't install it.* And there's another unpleasant side effect that whichever ideological regime spreads its 'free doctors' and 'free teachers' infrastructure around the world, especially to the Global South, will diffuse its own political regime. I'd rather that be ours.

Pluralism Can't Become Paralysis

This isn't an argument for autocracy. It is, with the appropriate moral seriousness, the opposite. The civilizational argument is that we want these benefits here. We want the standard of care of an American academic medical center diffused across rural America, urban America, the underserved America that has spent two generations being failed by the artisanal sick-care system we currently run. We want the deflationary, lifespan-extending, disparity-mitigating benefits of clinical AI delivered to American patients first and most comprehensively—not to patients whose deployment speed is purchased at the cost of political freedom. The pluralistic democracy isn't the obstacle to that vision. It's the reason for it.

But pluralism can't become paralysis. Due process can't become performative delay. Patient safety can't become guild protection with more elevated language. Litigation can't become the operating system of medicine. If we want the benefits of clinical AI to accrue inside an American moral framework, then we need an American installation strategy: liability safe harbors, harmonized standards, post-market monitoring, hospital captives willing to price risk, FDA pathways for adaptive tools, and the 150 acting not as passive complainants but as clinical deployment architects. The question isn't whether we can move as fast as China. We probably can't, and maybe shouldn't. The question is whether we can move fast *enough* not to lose the compounding benefits of diffusion.

That should concentrate some minds in Washington. It probably won't. But it should.

A Great National—And Multinational—Project

This is where the China argument has to become more than anxiety. The United States is behaving, in medicine, like a late-stage empire: procedurally ornate, legally exquisite, institutionally tired, over credentialed, under-diffusive, and oddly proud of the friction. China is behaving like a *rising engineering power*: impatient, state-aligned, willing to build the railroad while the lawyers are still debating the easement. The Thucydides-trap dynamic is everywhere in AI, chips, energy, Taiwan, industrial policy, and now health. But medicine offers one of the few domains where competition and collaboration can coexist without sounding entirely naive.

The project should be American first only in the sense that our institutions must finally mobilize. We need a national clinical AI diffusion program: model-evaluation networks, safe harbors, liability sandboxes, data trusts, multi-system clinical registries, patient-consent architectures, privacy-preserving computation, home-data standards, and AI-ready public datasets that aren't an embarrassment. But

the biology itself demands multi-nationality. A model that learns only from one late-stage empire's obese, overmedicated, administratively tortured, racially stratified, fee-for-service patient data will learn something real and also something parochial. *It needs the world.*

This isn't a call for an unconstrained global medical-data bazaar. Biosecurity is real. Data sovereignty is real. Privacy is real. Authoritarian misuse is real. But a humanitarian clinical AI project can be structured: federated learning, sovereign nodes, differential privacy, audit rights, consent layers, shared benchmarks, global disease registries, and domain-specific collaborations around cancer, rare disease, infectious disease, maternal mortality, neurodegeneration, and aging.

The Human Genome Project was international because the genome wasn't the property of one country. The Jupiter Brain for medicine should be international because suffering isn't sovereign.

And to use my preferred word: this is about diffusion, not mere installation. Installation is what the engineering state does to a technology: put it in the ground, connect the wires, deploy the app, mandate the pilot. Diffusion is slower, deeper, and more human. Diffusion means trust, liability, workflow, payment, physician adoption, patient understanding, governance, cultural permission, and institutional legitimacy. American democracy may not be able to install like China. But it can diffuse better if it decides to take diffusion seriously as a discipline rather than an afterthought.

Strategy Moves Before Doctrine: Why My Bet Is Google

Which is why we should look across the Pacific for instruction, or at least an indication—not because China's political model is admirable, and certainly not because American medicine should import authoritarian risk tolerance, but because strategy sometimes moves before doctrine, and deployment often creates the legal architecture that commentators later pretend was inevitable.

Liability as Strategy

BYD is the example I keep coming back to. BYD—Build Your Dreams, for those who don't live inside the EV discourse bubble—is no longer some scrappy Chinese battery company one can condescendingly ignore. It's the largest EV manufacturer in the world, a \$100 billion industrial colossus built around batteries, software, vertical integration, and an almost unnerving willingness to compress the future. In July 2025, BYD announced that its *God's Eye* assisted-driving system had reached Level 4 smart parking under specified conditions and that the *company would assume responsibility for losses incurred while the system was in control*.⁸³ Customers, as I understand it, wouldn't need to run first through the usual insurance ritual; they could go directly to BYD. This is a corporate-stimulated strategic doctrine. It's a company saying, in effect: we have enough consummate confidence in the system that we're de-risking this for you, the consumer. We understand the system best. We can price the downside. We'll stand behind it.

Now, yes, I recognize the nonequivalence. China isn't the United States. BYD operates in a different legal, political, institutional, and regulatory environment. The rule-of-law and consumer-litigation apparatuses are different. The state-sponsored capitalism is certainly different. Again, this is no apologia for authoritarianism, and I don't want our clinical AI system absorbing the more troubling features of the PRC diffusion ethos. But the analogy is still powerful. *The entity best positioned to understand the technology, manage the failure modes, update the system, absorb the risk, and price the downside moved first and underwrote liability ex ante.* The clinical AI analogue is suggestive, and to me at least persuasive.

⁸³ CarNewsChina: "BYD announces L4 autonomous parking feature, promises full liability for losses in China" published on July 9, 2025 ([Link](#))

A company that believes in its system shouldn't hide forever behind "decision support." It shouldn't throw the model into clinical workflow, issue a few disclaimers, publish a benchmark paper, and then shove all the risk back onto the doctor. *It should stand behind the machine.* That's the posture that converts a research artifact into a deployable clinical instrument.

And this isn't all hypothetical in the United States. Digital Diagnostics—formerly IDx—did something a few years back that deserves much more attention than it has received. Its autonomous diabetic-retinopathy system, IDx-DR, was the first FDA-cleared autonomous AI diagnostician in the United States, and the company assumed liability for the AI's performance when the tool was used appropriately.⁸⁴ That's the path. *The more informed party*—the AI creator—shoulders the risk. The deployment posture is the willingness to stand behind the machine with capital.

And the category has moved beyond one lonely example. Autonomous diabetic-retinopathy screening is now a small but meaningful proving ground for the larger thesis, with Digital Diagnostics, EyeArt, and AEYE Health all part of the autonomous retinal-screening landscape. But Digital Diagnostics remains the canonical liability prototype because it paired autonomous diagnosis with the principle that the *creator owns the output* when the tool is used appropriately. That combination matters much more than the narrowness of the use case. The fact that it begins in diabetic retinopathy isn't a weakness. It's how deployment starts: narrow, bounded, measurable, validated, reimbursable, and legally intelligible. The future doesn't usually arrive first as the whole gloriously finished cathedral. It arrives as a small chapel.

This is the sine qua non for VC-backed vertical applications in clinical AI. If you don't have enough confidence in your tool to assume liability and buy reinsurance, however beautiful or miraculous or category-defining your model is, and however compelling your arXiv benchmark paper, *you have a research project.* Not a deployable clinical instrument. Clinical AI companies love to talk about trust. Fine. Trust has a balance sheet. Trust has an insurance policy. *Trust signs the indemnification.*

Corporate strategy must move before tort reform. Don't wait for Congress, fifty state legislatures, or centuries of common-law adjustment to produce a nicely-compressed doctrinal architecture. It won't arrive in time. *Liability assumption itself is the first-mover advantage.* If the technology is genuinely safe and clinically effective, the company that assumes liability first signals conviction, telegraphs readiness, and forces the rest of the market to respond on new terrain. The law will come hobbling behind, as law often does, with a clipboard and a scowl, explaining what the pioneers already built.

Which brings me—with the slightly impolitic tenacity of a guy who has been making the same argument at investor dinners for a year and a half—to the same thesis. Whoever has the balance sheet, the confidence, the fortress reinsurance, and the strategic nerve to assume clinical AI liability will win not only the clinical AI market, but a structural position in the largest knowledge industry on earth. My bet is Google. *Maybe.*

Not because Google is morally purer than everyone else, or because Sundar's recent pronouncements are uniquely more enlightened than Sam's or Satya's, or because they're steadily, inexorably becoming the most valuable company in the world—yes, Jensen's days are numbered, please send all NVIDIA-related hate mail to the usual address—but because Google has the right combination of attributes at the right moment. Demis, who I increasingly think may make his first Nobel look, in retrospect, less like a coronation than a preface, has built a research culture inside DeepMind that takes biology and medicine seriously in a way OpenAI, for all its commercial brilliance, hasn't yet

⁸⁴ Michael Abramoff et al: "A reimbursement framework for artificial intelligence in healthcare" published on June 9, 2022 ([Link](#))

matched. And yes, OpenAI is mobilizing; GPT-Rosalind isn't exactly a subtle hint about where the frontier labs are headed. But Google's position is still singular.

AlphaFold was an epistemic event. Isomorphic is the commercial continuation of that scientific program (Thrive Capital, where I'm a venture partner, led this round). Google DeepMind's Co-Clinician effort looks less like an opportunistic healthcare press release and more like an emphatic continuation of Demis' mission. Google has spent years building the scientific, technical, and institutional substrate for exactly this kind of move: DeepMind, AlphaFold, Isomorphic, Gemini, Google Health, clinical research partnerships, cloud infrastructure, and a seriousness about biology that no longer looks extracurricular. This isn't the old Google Health wandering into the hospital with good intentions and insufficient institutional understanding. This is Google returning with an AI-native science engine.

They're also swimming in cash. Alphabet ended 2025 with \$126.8 billion in cash, cash equivalents, and marketable securities.⁸⁵ That's a fortress balance sheet of the kind that allows a company to absorb a multibillion-dollar tail risk without blanching. Put the pieces together: DeepMind, AlphaFold, Isomorphic, Gemini, Google Health, Co-Clinician, enterprise partnerships, cloud infrastructure, data infrastructure, and a balance sheet that can actually underwrite liability. What you see, looking at the totality of these moves, is a company that has noticed, slowly and then all at once, that healthcare is the largest industry in the American economy, that clinical AI is the largest knowledge-work prize inside that industry, and that liability is the bottleneck. If Google has the nerve to underwrite the risk, it has the tools, talent, cash, scientific credibility, and strategic patience to own the clinical operating system.

If Google moves decisively into clinical AI with a liability-assumption posture, it's not merely winning a market. It's becoming the operating system for the largest industry on earth. Combined with its AI infrastructure, DeepMind science engine, advertising-and-data moat, cloud business, and healthcare re-entry, a Google that owns the clinical AI liability posture is plausibly the most valuable company in the world. I don't say that lightly. I say it because the math demands it.

Again, this is a big *maybe* because assuming clinical AI liability will require a serious strategic resolution, and a willingness to accept not just the indemnification financial costs, but the potential reputational blowback from our societal reticence to embrace technological adoption risk. That alone may cause enough queasiness in Mountain View to scupper the whole project. At least for now.

So if not Google, then Microsoft or Amazon. The important point isn't the logo. It's the balance sheet and the strategic fortitude to take this one. Hyperscalers can obtain reinsurance at rates startups can't touch, absorb litigation, acquire the best vertical clinical AI companies, partner with the ones too strategically precious to buy, and use liability assumption itself as the strategic unlock. This is *where the undercapitalized clinical AI startup should get very nervous*. Some of these companies have marvelous products *masquerading as companies*. Some have exquisite models, beautiful workflows, meaningful early validation, but no balance sheet capable of carrying the risk their own products create. The hyperscaler can weaponize its balance sheet, snap up the technology-forward but liability-poor startups, wrap them in reinsurance, integrate them into clinical infrastructure, and make the thing deployable.

That's the acquisition thesis hiding inside the liability thesis. The next clinical AI M&A wave won't be only about model performance. *It will be about risk-bearing capacity*. The question will be: who has the model, who has the data, who has the workflow, who has the clinical trust, and who can afford to

⁸⁵ Alphabet: "Form 10-K" published on February 4, 2026 ([Link](#))

get sued? Startups may have the first three. The 150 may have the fourth. The hyperscalers have the fifth.

And that's why the winning architecture is probably not hyperscaler alone, or startup alone, or hospital alone. It's a liability-bearing coalition. The hyperscaler underwrites and scales. The startup contributes vertical excellence, proprietary data, modality-specific performance, or workflow depth. The 150 provide the deployment substrate: patients, physicians, clinical governance, institutional trust, malpractice captives, and real-world outcomes. That's the triangle. Intelligence, liability, bedside.

She who assumes the liability wins. But she'll probably buy, partner, and indemnify her way there.

When the Tool Becomes the Standard

Let's go a click deeper here, because liability doesn't merely determine whether clinical AI gets deployed. It eventually determines what counts as good medicine. That's the second implication of this whole blame-allocation question, and it's probably even more destabilizing for the profession than the first. Right now, malpractice culture treats AI guidance as *advisory*: useful for confirmation, helpful for documentation, perhaps valuable for administrative simplification, but not something one is expected to follow when it cuts against conventional practice. The human clinician remains the epistemic sovereign. The model can whisper, summarize, suggest, flag, and perhaps nudge, but the physician is still supposed to stand above it, verify it, domesticate it, and translate it back into the familiar standard of care. That principle is stable only as long as AI remains subordinate—only as long as the tool is modestly better, inconsistently better, or merely helpful around the margins. It becomes unstable rather quickly if the tool is simply, incomprehensibly, measurably superior.

This takes us, perhaps disturbingly, back to Geoffrey Hinton's original discomfort: *no less intelligent creature has ever reliably controlled a more intelligent one.*⁸⁶ And it takes us back to Jack Clark's line that we're "growing" these systems more than building them, which is another way of saying that their internal cognitive machinery isn't perfectly legible to us, even when their outputs are useful, structured, and often astonishingly correct. This is where clinical AI becomes epistemically weird. How, exactly, is a less capable human clinician supposed to independently verify the recommendation of a system that's not merely opaque, but superior? What does "human in the loop" mean when the loop contains a machine that has read everything, remembers everything, sees across modalities, compares millions of analogous cases, and produces an answer the human can audit only partially? We keep pretending that the physician can stand above the model like a professor grading a resident's note. But in bounded clinical domains, *that hierarchy may invert*. The physician may increasingly be less the sovereign verifier of machine judgment and more the accountable human steward of a larger clinical intelligence system.

That doesn't mean we surrender to the oracle. I'm not arguing for epistemic idolatry, and I don't want clinical medicine reorganized around some dark-enlightenment faith in the machine (again, have a look at my Deification chapter for some warnings here). We need auditability, provenance, model-version tracking, citations, uncertainty flags, confidence intervals, escalation protocols, and clear documentation of whether the clinician accepted, modified, overrode, or under-rode the recommendation. We need structured reasoning artifacts the legal system can metabolize, even if we don't get full mechanistic transparency into the model's neurology. But we should be honest: auditability isn't the same as perfect interpretability, and perfect interpretability can't become the new disguised form of infallibility. *We don't fully understand the human brain either*, and yet we allow it to practice medicine after training, testing, licensing, supervision, malpractice coverage, and accumulated social

⁸⁶ The Nobel Prize: "Geoffrey Hinton Interview" published on December 6, 2024 ([Link](#))

trust. The question isn't whether the model is metaphysically transparent. The question is whether it's validated, monitored, bounded, useful, safer than the alternative, and accountable when it fails.

Once that happens, the standard of care begins to invert. Today, deviation from conventional practice protects against malpractice. Tomorrow, deviation from a validated AI recommendation may itself become the thing that requires explanation. Why was the better instrument not employed? Why was unaided judgment used when something demonstrably superior was available? Why did the patient not receive access to the model that had lower miss rates, better sensitivity, better specificity, better calibration, better longitudinal outcome prediction, or better medication optimization inside that validated envelope? As Vinod says, at some point soon it'll no longer be malpractice to use clinical AI. It may be malpractice not to use it. That's the future-malpractice inversion, and the profession is nowhere near ready for it.

This isn't an abstract legal thought experiment. Standard-of-care change isn't merely a doctrine produced by courts after the fact. It's organizational. It gets made in workflows, order sets, EHR prompts, quality measures, medical-staff policies, payer rules, malpractice underwriting, professional society guidelines, peer-review committees, residency training, and the quiet accretion of institutional habit. If a health system deploys a validated sepsis model with documented improvements in escalation and mortality, and another hospital doesn't, how long before the second hospital's inaction becomes legally and morally legible? If one oncology program uses a model that reliably identifies trial matches or molecularly appropriate therapies that another misses, how long before the missed opportunity becomes a liability event? If one primary-care platform uses AI to identify abandoned medications, rising depression risk, uncontrolled hypertension, and heart-failure decompensation before the ED visit, and another doesn't, how long before the old "standard" stops looking standard at all?

Hospitals will play a major role in shaping that deployment context because they already control a large share of the physician base—*roughly 54 percent of physicians* in some form of employment, depending on how one counts ownership, contracting, and hospital-affiliated practice.⁸⁷ *That physician employment base is a prefabricated clinical AI diffusion channel.* But the tool has to be deployed in ways that make the physician more *powerful* rather than more *surveilled*. If clinical AI arrives as another compliance instrument, another productivity cudgel, another ambient panopticon measuring the doctor's every pause and deviation, the profession will resist it with every antibody it has. If it arrives as relief—less pajama time, less inbox sludge, better pre-visit synthesis, smarter medication reconciliation, better escalation, fewer missed findings, more time with the patient—then the physician becomes the missionary rather than the hostage. The same technology can feel like liberation or surveillance depending on how the institution installs it.

Payers may be even more determinative. UnitedHealth Group employs or controls roughly 90,000 U.S. physicians, about one in ten, which remains one of the more astonishing and under-appreciated facts in American healthcare.⁸⁸ Combine that physician asset with premium-funded ambulatory networks, claims-and-clinical data at scale, home assets, pharmacy, behavioral health, and a willingness to industrialize, and you should have a structurally advantaged actor in the diffusion of clinical AI. Provided, of course, the public and regulatory backlash from the overall industry's denial machinery doesn't cause a nuclear-core meltdown—which, to be fair, is still an unresolved question (although United's smart May announcement of a unilateral removal of 30% of pre-authorizations

⁸⁷ Physicians Advocacy Institute: "PAI-Avalere Health Report on Physician Employment Trends and Practice Acquisitions: 2018-2026" published in April 2026 ([Link](#))

⁸⁸ Becker's ASC Review: "UnitedHealth's 90,000 physicians, 423 ASCs and 2,700 subsidiaries" published on July 17, 2025 ([Link](#))

mitigates this risk).⁸⁹ The vertically integrated payer has every ingredient clinical AI wants: the premium dollar, the data, the physician channel, the actuarial reason to prevent utilization, and the managerial temperament to standardize care. Hospitals shouldn't assume that because payers are politically unpopular, they're strategically defeated. Unpopular empires can still be formidable.

Now to *what gets unlocked*, because the point of liability isn't merely to move risk from one pocket to another. The point is that once somebody stands behind the machine, the care model itself can start to change. For a long time, as I outlined above, our reaction to biological complexity was to draw smaller and smaller circles around the body: organs, pathways, receptors, cell types, disease categories, specialties, subspecialties, until one person's tiny sliver of expertise became the unit of authority. Rational under conditions of cognitive limitation and data scarcity. But those same responses produced one of the deepest pathologies in American medicine: *balkanization*. Fragmentation masquerading as sophistication. *Biology is integrated*; the healthcare system built around it emphatically isn't. The whole patient disappears into the mosaic of separately reimbursed specialists. Records atomize. Incentives diverge. Responsibility diffuses. Behavioral health sits off to the side like an afterthought. Pharmacy is treated as a cost center or PBM battlefield rather than a central therapeutic lever. Social needs are documented, occasionally, and then too often politely ignored. The patient becomes the integration layer, to the detriment of us all.

Balkanization isn't just inefficient; it's unjust. The wealthy, urban, educated, white, well-connected patient can sometimes navigate the labyrinth with enough persistence, concierge assistance, and social capital. The rural patient, the poor patient, the elderly patient, the nonwhite patient, the patient with depression and diabetes and heart failure and transportation insecurity and a medication list long enough to qualify as literature, is asked to serve as the courier of context between institutions that don't really see one another. This is a disconnected series of encounters pretending, through branding and billing codes, to be a system.

AI's clinical promise is reintegration. Polymathy returned to medicine by non-biological means. A sufficiently capable clinical model can synthesize medication adherence, social needs, mood, sleep, diet, symptoms, labs, imaging, claims, messages, genomics, and clinician judgment into a coherent account of an evolving patient state. The diabetic with heart failure, depression, financial insecurity, medication complexity, and transportation barriers isn't a set of billing encounters. He's an evolving clinical state across medical, behavioral, pharmacologic, and social domains. That's *synchronization*: a care model that keeps context alive *between* visits, detects abandonment risk, prompts outreach, reconciles medications, coordinates behavioral health, and ensures the next best action isn't lost in the gap between departments.

And the gaps are everywhere. Patients abandon drugs at rates approaching 40 percent, with economic costs that plausibly run into the hundreds of billions annually. Referrals leak. Behavioral health remains peripheral. Specialty care proceeds without whole-person context. Medication adherence is treated as a patient-behavior problem rather than a system-design problem. Social determinants become an ICD code, a dashboard, a grant application, and then, too often, not much else. Clinical AI that synchronizes this finds a structurally enormous, largely unoccupied market. It can keep the longitudinal thread alive when the human system loses it. It can notice that the patient stopped refilling the diuretic, missed the cardiology appointment, wrote three anxious portal messages, gained six pounds, slept poorly, and lives forty minutes from the clinic. None of those signals alone may trigger action. Together they may be the story.

That's what liability unlocks. Not merely autonomous diagnosis. Not merely better clinical decision

⁸⁹ UnitedHealth Group: "UnitedHealthcare Cuts Prior Authorization Requirements by 30%" published on May 5, 2026 ([Link](#))

support. Not merely a clever second read or a nicer differential. It unlocks the possibility of a care-delivery architecture *organized around continuous context* rather than episodic intervention. It allows medicine to move from disconnected expertise toward synchronized intelligence. It lets the physician, nurse, pharmacist, behavioralist, social worker, and care manager operate around a shared, updating representation of the patient rather than a scattered collection of notes, claims, labs, and half-remembered conversations. It begins to make the system synoptic again.

This is where the 17-day mandate returns. A standard of care built around unaided human memory can tolerate 17-year diffusion lags because everyone is moving slowly together. A standard of care built around AI-enabled clinical systems can't. Once the system can know, surface, and synchronize faster, failure to do so becomes legible. The old standard asked what a reasonable physician would do in the room. The new standard will increasingly ask what a reasonable AI-enabled system should've known before the room became necessary.

That's the deeper inversion. The old standard of care was built around what a reasonable physician would do under the circumstances. The new one will increasingly ask what a reasonable AI-enabled clinical system should have known, surfaced, prevented, escalated, or synchronized. That's a much larger question, and it's why the 150 need to begin preparing now. Build the validation machinery. Define the clinical envelopes. Instrument the use cases. Protect clinicians inside approved workflows. Demand liability-bearing partners. Use the employed physician base as the diffusion channel. Don't wait for the standard to invert and then discover that someone else—payer, hyperscaler, regulator, plaintiff's attorney, or vertically integrated platform—wrote the operating manual.

Medicine's next standard will be built not only around human judgment, but around whether human judgment was properly augmented. The profession isn't ready for that. The institutions aren't ready for that. But readiness, as always, isn't the same thing as arrival. The inversion is coming anyway.

The Hive Mind and the Instantaneous Standard of Care

The standard-of-care inversion becomes still stranger once clinical AI *agentifies*. Elsewhere in this essay I use the phrase *Hive Mind*, and it belongs here. Imagine the care team decomposed into individuated agents: one doing original molecular research against the patient's tumor biology; another scanning the subspecialty literature; another monitoring behavioral health; another reconciling pharmacology, pharmacogenomics, and drug-drug interactions; another watching home telemetry; another navigating benefits and prior authorization; another coordinating transportation, social needs, and adherence; another synthesizing the whole thing into a next-best-action plan for the physician-architect. We'll anthropomorphize these agents because human beings anthropomorphize everything, including Roombas, markets, weather, and occasionally committees. That's fine. But their determinative property isn't human. *It's transhuman coordination.*

Humans individuate knowledge and then spend onerous hours every day trying to communicate it back into a shared state: meetings, conferences, phone calls, texts, emails, handoffs, curbside consults, inbox threads, Slack messages, Teams channels, and the grim little rituals of organizational life. Recent neuroscience estimates conscious human information throughput at roughly 10 bits per second. The exact number will be debated, *but the felt truth* is obvious: we update each other laboriously and clumsily. Agents don't have to. Agents can be an *aspen tree*: many trunks, one root system. Everything one individuated agent learns can update the hive. The system recursively, relentlessly learns and improves. As I typed that sentence too quickly, my spellcheck changed *improving* to *unionizing*, which is disconcerting in ways I'm not yet emotionally prepared to unpack.

Moltbook, anyone? My healthcare readership should pause and google this. But back to our narrative.

Now apply that to the standard of care. The future standard isn't a septuagenarian committee of specialty luminaries in a wood-paneled boardroom meeting every two years to bless a guideline PDF. I mean no disrespect to septuagenarians, luminaries, wood paneling, or PDFs. But that cadence is ludicrous in a world of recursive clinical agents. The standard of care becomes a living object: updated by new evidence, local outcomes, global outcomes, model performance, adverse events, medication signals, trial results, patient-reported outcomes, and the hive's continuous experience. The update interval moves from years to months to weeks to days to, in some domains, something close to instantaneous. Again, have a look at Anthropic's confessional June blog on recursively improving AI systems. Worth a read. A stiff drink recommended.

In short, this is both magnificent and institutionally terrifying. What happens to credentialism when the credentialed body no longer owns the frontier of clinical knowledge? What happens to ordination when expertise is demonetized and redistributed through agents? What happens to specialty societies when their guidelines become downstream summaries of a learning system rather than upstream sovereign commands? The institutions that once protected patients through deliberative scarcity may begin to harm patients through deliberative delay. The question isn't whether the institutions are evil. They aren't. The question is whether they can survive the loss of epistemic primacy without converting their remaining legitimacy into obstruction.

This is also why UnitedHealthcare, and entities like it, shouldn't be treated as a side character. *Should a vertically integrated payer underwrite clinical AI risk?* The idea sounds perverse until you look at the assets: premium dollar, claims data, physician employment, ambulatory networks, pharmacy, home, behavioral health, actuarial machinery, and an economic reason to prevent avoidable deterioration. The obvious objection is trust. A payer-built clinical AI can become a denial machine in a lab coat. That risk is real. But structurally, an entity like United is one of the few private actors capable of pricing, distributing, and underwriting clinical AI at national scale. The question is whether society can force that underwriting into a pro-patient liability architecture rather than a utilization-management panopticon.

The Hive Mind standard of care isn't optional if the technology works. It'll arrive first in fragments: oncology pathways, pharmacology, imaging, deterioration detection, specialty triage, home monitoring, refills, behavioral escalation. But the fragments will connect. And once they connect, the old standard-of-care machinery will look less like prudence and more like latency. Latency in a software system is an annoyance. Latency in medicine is harm.

Part VI—The Reconstitution

This final part turns from diagnosis to design. If clinical AI becomes real medicine, then the physician's role, medical education, payment, home-based care, presymptomology, liability, and CEO-level diffusion all have to be rebuilt around the new cognitive substrate.

The Physician-Architect and the New Flexner Settlement

If the standard of care begins to invert, and if liability-bearing clinical AI starts moving from advisory curiosity into the actual machinery of medicine, then the next question is unavoidable: what happens to the physician? Not in the cartoonish sense—will AI replace doctors?, that dreary panel-discussion question that somehow refuses to die—but in the deeper professional, epistemic, and moral sense. What's the physician *for* when clinical intelligence is no longer scarce, no longer exclusively biological, and no longer monopolized by the guild? What does medical judgment mean when the model can see more, remember more, synthesize more, and, in bounded domains, outperform the unaided human? And how do we preserve the *sacredness of care* while admitting that the old physician-centered architecture of medicine was built around a scarcity that's now beginning to dissolve?

For too long we've relegated our most expensive and highly trained human assets to the role of glorified data-entry clerks. Physicians type. They click. They reconcile. They hunt for records. They draft prior-auth letters. They answer portal messages. They fight the EHR, the payer, the schedule, the inbox, the documentation requirement, the coding logic, the administrative barnacles that've attached themselves to the practice of medicine with the tenacity of some parasitic marine organism. And then we wonder why they're burned out. We take the most cognitively trained, socially trusted, expensively educated people in the labor market and bury them in clerical sludge, then commission another wellness webinar when morale collapses.

First Act: Reclaim the Physician

Clinical AI should *reclaim the physician*, at least in its first and most humane phase. A restoration of joy to the practice of medicine. If AI handles the mundane, the physician doesn't become irrelevant. She becomes more central, but in a different way: less data-entry clerk, more high-stakes strategist; less documentation engine, more edge-case interpreter; less routine biological pattern-matcher, more steward of ambiguity, empathy, ethics, risk, trust, and patient-specific trade-offs. The early promise is real and shouldn't be dismissed as mere vendor optimism. Chart synthesis, inbox triage, prior authorization support, documentation relief, medication reconciliation, pre-visit planning, literature synthesis, risk stratification, and clinical-gap closure can give physicians back hours of time and, more importantly, psychic room. A physician who isn't drowning in pajama-time documentation may actually have the capacity to think, listen, explain, persuade, comfort, and decide. Imagine that. Medicine as medicine.

But we shouldn't tell ourselves bedtime stories either. The cheerful version of this transition says AI will do the paperwork while doctors get restored to the high priesthood of judgment. That's true for a while, and it's worth fighting for, but it isn't the whole story. As the models inexorably become smarter, they won't confine themselves to the mundane. They'll *move upward* into the most complex, interdisciplinary, frontier domains of medicine—the places we currently enshrine as the apex of professional judgment. Diagnosis in complex multisystem disease. Oncology therapy selection across genomics, pathology, trial eligibility, and patient preference.

Neuropsychiatric treatment stratification. Drug interaction and pharmacogenomic reasoning. Rare disease detection. Multimodal deterioration prediction. The model won't obediently remain in the clerical basement. It'll come for the sanctum.

Second Act: Relocating Judgment

That's the part the profession needs to start preparing for now. Clinical judgment won't disappear, but it will be *relocated*. It will be less sovereign, more architectural; less solitary, more system-mediated; less the private origination of truth, more the accountable human stewardship of a larger intelligence system. The physician's highest role may increasingly become the edge case: the moral case, the ambiguous case, the family meeting, the frail elder with five comorbidities and three competing goals, the cancer patient choosing between marginal survival benefit and quality of life, the psychiatric patient whose risk is legible only through relationship, the patient whose social reality makes the "right" medical answer impossible to implement. This is where humans still matter most. Not because we can out-compute the model. Because we can bear responsibility, interpret meaning, and hold trust.

And critically, clinicians shouldn't merely adopt AI tools designed elsewhere. *They should help architect them.* Clinical logic, safety envelope, workflow, failure modes, escalation paths, evaluation criteria, patient-facing implications, model override norms, uncertainty thresholds, consent language, documentation expectations, liability boundaries—these can't be designed solely in Silicon Valley shared office space by brilliant people who've never watched a resident try to discharge a medically fragile patient at 6 p.m. on a Friday while the pharmacy, payer, family, home-health agency, and transportation vendor all conspire, unintentionally but effectively, to make reality impossible. *Clinical AI has to be co-developed in clinics*, not shipped into the EHR like one more workflow indignity. This goes back to the exhortation I keep making to the 150: get into the game. Co-design the paradigm shift. Don't abdicate yet another technological revolution and then spend the next twenty years complaining that the tool doesn't understand medicine.

The design standard should be clear. *If a clinical AI tool doesn't reduce cognitive load within the first sixty seconds of use, it's a failure of design.* Physicians don't need more dashboards, alerts, explainability theater, workflow ornaments, or morally improving pop-ups. They don't need another tool that requires them to become part-time implementation consultants. The tool has to make the work easier, safer, faster, more coherent, more clinically alive almost immediately. And the physician has to experience it as *relief and force multiplication*, not surveillance. That distinction matters enormously. The same AI that feels liberating when it removes sludge will feel authoritarian when it arrives as another productivity cudgel. If the physician experiences clinical AI as a panopticon, the antibodies will be ferocious. If she experiences it as liberation from the clerical regime, she becomes the missionary.

This is what I mean by the *physician-architect*. Not physician veto. Not passive adoption. Not the old guild posture of "nothing touches the patient unless we bless it after five years of committee review." And not the opposite pathology either, where technologists ship the future into medicine and ask the clinicians to adjust. Physician-architects design, validate, govern, and deploy clinical AI. They understand the bedside and the model, the workflow and the safety envelope, the patient relationship and the data structure. Their role is no longer merely to practice medicine inside an inherited system. It's to reconstruct medicine.

Some of the old priesthood will be *defrocked*—sorry, yet another eruption of my long-suppressed Catholic indoctrination. But we should be honest about this. Expertise, as Vinod has argued in various forms, is *moving asymptotically toward free*, and the prestige professions are going to feel that first and most spiritually. Law, consulting, finance, medicine, academia—professions whose economic and social prestige rested on scarce judgment, scarce knowledge, scarce synthesis—will all face

some degree of *demonetization* as intelligence becomes abundant and distributed. Medical expertise won't be exempt because we declare it sacred. The more ubiquitous and democratized an asset becomes, the more it's commoditized. That's not an insult to doctors. It's an economic eventuality we would do well to acknowledge in advance.

The ATM analogy is useful, though only up to a point. ATMs didn't eliminate bank tellers in the 1980s; they *changed what tellers did* and helped banks open *more* branches, so teller employment actually rose for a while. Then automation came for more of the work, branch models changed, and the labor story became more complicated—teller jobs started to disappear. The clinical version will likely play out over a decade or more, not overnight. Initially, AI will be a boon. It will streamline administrative tasks, restore joy to parts of the practice, expand panels, lengthen patient interactions, improve reach into underserved geographies, and make physicians feel less like exhausted typists in an infinite billing machine. The typical physician spends only a minority of her time on direct clinical care; too much of the rest is documentation, EMR management, prior authorization, inbox work, and the famous two to three hours of nightly pajama time. *De-bureaucratization* could meaningfully mitigate physician scarcity and geographic maldistribution. Vinod's "100 touches in 100 days" idea for primary care (see my Advisory Board print interview with Vinod from a few years ago) captures the upside: an AI-mediated engagement model in which the physician and care team can maintain continuous, lightweight, clinically intelligent contact with patients rather than episodic, fifteen-minute tollbooth medicine. If that model diffuses, primary care could finally move from the current meager 5 to 7 percent of U.S. medical spending toward something more like the 20 to 25 percent it probably deserves.

But again, the second act is harder. If *expertise moves toward the marginal cost of compute*, some of the compensation premium attached to scarce clinical cognition will come under pressure. Not all of it. Not the embodied work, not the procedural skill of knife cutting skin, not the moral presence, not the accountability, not the trust relationship. But the idea that the physician is uniquely valuable because she alone can retrieve, synthesize, and apply the relevant clinical knowledge will become harder to defend as the models improve. The profession should therefore avoid the most self-defeating posture available to it: pretending this is all just documentation relief. It isn't. It's a reallocation of clinical intelligence across human and non-human agents. The sooner physicians take authorship over that future, the better the future will be for both physicians and patients.

The Sacred Trust

Now let me slow down and say the thing that has to be said, because otherwise this starts to sound like a heartless staffing manual written by someone who doesn't understand the profession. Not the case. I have *reverence* for clinicians. Veneration, even. Nurses, physicians, therapists, pharmacists, aides, and social workers occupy a sacred relationship with patients at moments of fear, pain, dependency, shame, hope, grief, and vulnerability. Healthcare isn't merely a labor market. It isn't merely a workflow machine. It is *consecrated work*. That phrase may sound too theological for a strategy essay, but I mean it quite literally. The work carries an interiority, a responsibility toward embodied suffering and human frailty, that most sectors never have to bear.

That sacred trust has to survive the intelligence age. AI can augment the work, but if we use it to replace the human relationship with a simulacrum—if we reduce the encounter to an uncanny-valley chatbot with a cheerful tone and no soul—we'll have committed a category mistake of civilizational proportions. *The point of de-bureaucratization is to restore humanity, not strip it away.*

Get the nurse out of the chart and back to the bedside. Get the physician out of the inbox and back

into the conversation. Get the pharmacist out of fax warfare and into medication optimization. Get the care manager out of manual list-building and into actual intercession. If the administrative-labor savings simply become margin while the patient gets a machine instead of a human at the moment of need, the politics and morality of this transition will curdle quickly—and they should.

This is the new covenant that has to replace the old labor covenant: fewer humans wasted on administrative drudgery, more human attention available for the parts of care that still require presence, trust, judgment, and love. Clinical AI shouldn't become an austerity machine with a bedside avatar. It should become the instrument by which we rescue the human encounter from the administrative regime that's been slowly suffocating it. That's how the profession can support this transition without feeling that it's collaborating in its own desecration.

A New Flexner Settlement

Which brings me to something I increasingly think we actually need: *a new Flexner Report*. I know this is a freighted assertion, and I know the original Flexner Report carries all sorts of complicated legacy and stigmatization. Hear me out. The original report, commissioned by Carnegie and published in 1910, set fire to the old order of American medical education, excoriating weak standards, poor facilities, laughable admissions criteria, and a professional culture insufficiently grounded in science and clinical training.⁹⁰ Half the schools closed. Medicine became more standardized, more scientific, more regulated, more professional—and yes, more exclusionary too. Much of modern medical education still sits inside the world Flexner built.

We now need an AI Flexner Report. Not because Flexner was wrong, but because the premises of his world no longer fully hold. His model was built for an era in which memorization, hierarchy, gatekeeping, and physically localized expertise made sense. Ours is moving toward a world in which omniscience, total recall, simulation, interpolation, extrapolation, and hypothesis generation sit in your pocket. The old question was how to make medicine scientific. The *new* question is how to make medicine humane, rigorous, and trustworthy when the raw materials of scientific and clinical cognition are increasingly synthetic, ubiquitous, and superhuman.

A new Flexner would ask how we should reconceptualize the education, deployment, and evaluation of physicians and care teams if AI becomes foundational to future medical practice. Stop selecting medical students primarily on MCAT scores, memorization capacity, and the ability to perform excellence inside an educational system built around information scarcity. Select instead for clinical judgment under uncertainty, ethical reasoning, human connection, adaptive learning, epistemic humility, taste, agency, and the *capacity to work fluently alongside AI systems* that will rapidly surpass them in raw recall. The physician of the future doesn't need to be a worse version of the model. She needs to be the person who knows when the model is right, when it's wrong, when it's technically correct but socially impossible, when it's medically elegant but morally obtuse, and when the patient needs a human being more than an answer.

Residency has to be rebuilt around AI-augmented decision-making, not the illusion that the resident's unassisted judgment is the gold standard. The old training culture valued endurance, recall, pattern recognition, and progressive autonomy under supervision. Some of that remains essential. But we also need residents trained in model supervision, probabilistic reasoning, workflow design, failure-mode analysis, patient communication in AI-mediated care, and the ethics of human-machine decision-making. They need to learn how to interrogate model output without pretending they can

⁹⁰ Abraham Flexner: "Medical education in the United States and Canada. From the Carnegie Foundation for the Advancement of Teaching, Bulletin Number Four" published in 1910 ([Link](#))

independently reproduce everything the model knows. They need to practice the residual human role before it atrophies into ceremonial sign-off.

Continuing medical education needs to become a living, learning system rather than a compliance ritual. AI-generated case data should surface individual physician error patterns, local practice variation, missed opportunities, outdated habits, and deviations from emerging evidence. Every cancer patient should get the benefit of the last thousand relevant publications in her disease, not just the three her oncologist happened to read. Every cardiac patient's anticoagulation, dose, and risk profile should be modulated by genomic, transcriptomic, real-world, and longitudinal evidence rather than stale clinical habit. Every patient encounter should contribute to a learning system rather than vanish as one more isolated anecdote in a chart.

And we need to reexamine the economics of medical training. The \$250,000 debt burden that binds the next generation of physicians to a compensation structure that *will not survive the decade* isn't some side issue. If the profession's knowledge premium compresses while training remains ruinously expensive, we'll have created a grotesque intergenerational trap. We can't ask young physicians to mortgage their futures for an economic model built around cognitive scarcity if cognitive scarcity is precisely what this technology erodes. That mismatch will become politically and morally untenable.

The new Flexner settlement should also define benchmarks for when AI is good enough to enter clinical practice. *Perfection isn't the standard; human superiority is.* The National Academies, or some similar body with enough legitimacy to matter, should help articulate domain-specific baselines: median physician, top-decile physician, institutional best practice, error rate, resource intensity, equity effect, latency, longitudinal outcome. Once AI meets or exceeds those baselines inside a validated envelope, the path to deployment should open. No impossible standard. No machine infallibility test masquerading as ethics. Human superiority, not perfection.

And yes, if the patient grants it—and I would—conversational AI with memory, social-medical context, behavioral patterning, schedule data, medication history, sleep patterns, and even the odd little digital breadcrumbs of daily life begins to create a much richer, stranger, more continuously adaptive model of care. Not to dehumanize medicine, but to give it continuity and memory it's never possessed at scale. The irony is almost too perfect: the machine may help restore the longitudinal human relationship precisely because the human system could no longer remember the human being.

That's the Flexner question now. If medicine's epistemology changes, medical education has to change. If clinical intelligence becomes abundant, the physician's role has to change. If judgment is no longer monopolistically human, then training has to produce humans who can steward machine judgment without surrendering moral agency. The 1910 report still heavily influences medical curricula today. That's no longer sufficient. We need a new settlement for the intelligence age.

And this brings us naturally to the next problem: even if we redesign the physician's role, preserve the sacred trust, and rebuild medical education, clinical AI still won't diffuse correctly unless the financing substrate changes. A system that pays for activity will resist intelligence that prevents activity. Which means the next question isn't pedagogical, but economic: what payment model actually rewards the medicine clinical AI makes possible?

Capitation Becomes Ascendant

Ok, let's land the plane. Way too many notes in this one, to channel *Amadeus*, but the functional question now becomes unavoidable. If everything I've argued so far is directionally right—if clinical AI begins to invert the standard of care, if liability-bearing models move from advisory curiosity into actual deployment, if the physician becomes more architect than solitary oracle, if the sacred trust has to be preserved while medicine becomes more computational—then what does this mean for the system as it actually exists? Not the metaphysics of intelligence. Not the AlphaFold sermon. Not the grand civilizational anxiety. What happens to reimbursement, care delivery, primary care, the home, and the distribution of expertise itself?

To be fair, none of this works at scale under the current financing logic. That's the first hard landing. The business-model problem with clinical AI is that it doesn't fit naturally into fee-for-service. At all. *Fee-for-service pays for activity, and clinical AI's most beautiful use cases often destroy activity.* At least the activities that predominate today. It prevents admissions. It avoids procedures. It keeps the heart-failure patient from decompensating. It gets the hypertensive patient back on medication before the stroke. It closes the behavioral-health loop before the ED visit. It coordinates care in ways that make the next billable intervention unnecessary. Wonderful for the patient. Wonderful for the country. Awkward for a system that still treats avoidable utilization as revenue. Clinical AI therefore works best under capitation, delegated risk, case rates, Medicare Advantage, Medicaid risk, provider-sponsored health plans—any arrangement where preventing utilization is economically legible. Capitation becomes supreme, or at least *more* supreme (if this phrasing is grammatically allowed) than it has been. More—much more—on this reimbursement hydraulic later in the essay.

Why VBC Finally Gets Its Operating System

I know, I know. Value-based care has been one of the more disillusioning policy sagas in modern healthcare: fifteen years of evangelism, billions in CMMI experimentation, endless conference panels, heroic PowerPoints, and tangible savings still underwhelming enough to induce existential fatigue. *The fair question is what's different this time.* The specific answer is *tools*. Prior VBC efforts didn't fail because the underlying theory was wrong; they failed, in large measure, because the operating technology was inadequate. You can't manage a risk-bearing population of 50,000 patients with care managers spelunking through spreadsheets, static registries, delayed claims feeds, and heroic human follow-up. *The cost of intervention*—even if it worked in these cases, which it mostly doesn't—*overwhelms the value of prevention.* Clinical AI changes that ratio. An AI agent that monitors a panel of 2,000 patients continuously—flagging decompensation risk, closing medication gaps, coordinating follow-up, catching the patient who stopped refilling her blood pressure medication three weeks before her stroke—does the work of dozens of care managers at a fraction of the cost. Same theory. Different tool. Or maybe more precisely: same theory, finally with an operating system.

Another way to say this is that *clinical AI may resurrect value-based care.* VBC didn't fail because prevention is a bad idea; it failed because we tried to operate a longitudinal, risk-bearing, continuously adaptive theory of medicine with batch claims, static registries, manual care management, and human beings dialing through lists like it was a 1998 call center. Clinical AI finally gives VBC a nervous system: panel surveillance, patient-state prediction, automated outreach, medication-gap closure, behavioral nudges, home-data ingestion, risk stratification, and escalation at the moment the curve starts to bend rather than after the hospitalization has already consumed the savings. Same ideology, different operating substrate. *This time, the tools may actually match the theology.*

This is why advanced primary care may yet have its resurrection. The sector has had a rough and recently humiliating run. V28 decimated parts of it. Cano's story turned sad and ignominious. Oak Street got absorbed, and then got the acquiring CEO fired. CareMax limped. The whole category has oscillated between evangelism and disillusion, and primary care without the premium dollar has too often looked like a noble but financially marginalized asset. And yet the deeper thesis was never wrong. *It was under-instrumented*. Advanced primary care, at its best, already approximates the architecture the new world wants: risk-bearing, team-based, ambulatory, longitudinal, PCP-centered, behavioral-health-aware, pharmacy-integrated, and oriented toward prevention rather than episodic rescue. Doctors, nurses, behavioralists, pharmacists, medical assistants, care navigators, home-based resources, digital support—all orchestrated over time rather than episodically. Clinical AI gives that model its missing nervous system.

That's also why provider-sponsored health plans matter more than the recent rough run suggests. Health systems—don't sell your health plan! Read my payer chapter for the why (not). I *do* understand why systems are tempted: onerous RBC requirements, network adequacy headaches, SG&A burden, actuarial irritation, payer economics, regulatory complexity, the whole miserable alphabet soup of reasons provider-sponsored plans are annoying and non-core. But if clinical AI's value accrues most naturally to the owner of the premium dollar, then divesting that asset at precisely the moment longitudinal synchronization becomes operationally feasible may be a category error. *The winning combination is premium dollar, ambulatory assets, employed or aligned physicians, advanced primary care, home-based care, pharmacy integration, behavioral health, and clinical AI*. Hospitals that still think of physician employment primarily as referral capture are missing the point. The future is force multiplication of the physician base you already own, married to risk-bearing economics that let avoided utilization become value rather than lost revenue.

And yes, vertically integrated payers understand this, which is why providers should be very careful about celebrating the payer super-cycle's apparent stumble as if the war is over (again, see my payer chapter). UnitedHealth and its peers have the premium dollar, physician assets, ambulatory networks, pharmacy, claims and clinical data at scale, home assets, capital, and the willingness to industrialize. Combine all of that with clinical AI, and the old hospital-less IDFN (*integrated delivery and financing network*) strategy gets an operating system it never had before. The hospital counter isn't nostalgia for the old inpatient center of gravity. It is to build, preserve, or partner *into* the financing and care-delivery substrate in which clinical AI actually makes sense. Pure hospitals, especially those still overwhelmingly fee-for-service and not consummate operators like HCA, remain vulnerable. Systems with premium dollar, primary care, home, pharmacy, behavioral health, and AI-enabled longitudinal management become much more interesting.

Home as the Epicenter

De-Institutionalization as Strategy

The deeper word for this section is *de-institutionalization*. Home isn't merely another site of service; it's a *repudiation of the assumption that medicine must be institutionally concentrated to be safe*. The hospital remains indispensable for acuity, procedure, trauma, complex diagnostics, and the many forms of suffering that require density. But much of modern healthcare has been institutionalized because intelligence, monitoring, coordination, and escalation were scarce outside the institution. Diffuse intelligence, and the institutional boundary starts to move.

The home also becomes the *engine of personalization* because the data become continuous,

high-velocity, and personal. A lab value every six months is a postcard from a distant country. A wearable, CGM, cuff, scale, medication dispenser, voice interface, gait sensor, sleep sensor, and patient-agent conversation create an evolving stream. Whether the daily number is 118,000 biomarkers, 118,000 observations, or some other device-dependent count, the conceptual point is the same: velocity and volume transform the patient from a sparse episodic chart into an N-of-1 learning system. The last mile of medicine becomes the first mile of data.

That continuously updating data is also a moat. Check out the company all.health for a master class in this area. Static proprietary data are useful; moving proprietary data are powerful. The patient's state isn't a noun. It's a verb. It changes with sleep, food, stress, medication adherence, infection, grief, weather, movement, social contact, and the thousand unglamorous inputs that determine whether a care plan lives or dies. The company or health system that owns, governs, and ethically uses this evolving data stream owns personalization. Not generic personalization as marketing copy, but actual N-of-1 medicine: dosing, timing, outreach, behavioral intervention, escalation, and prevention tuned to the person in motion.

Finally. I've been arguing for this transition for years, and I believe it may indeed be here. As clinical AI matures, the site of care changes in one direction: toward the home. This isn't merely "hospital at home" as a reimbursement waiver, or remote patient monitoring as another vendor category with too many alerts and not enough workflow. This is something bigger. GenAI, multimodal LLMs, biological foundation models, wearables, computer vision, contextual sensors, and increasingly cheap home-monitoring infrastructure make it possible to see the patient outside the institution with a fidelity we've never had before. Not perfectly. Not magically. But enough to change the care model. Decompensation becomes more predictable. Intercession becomes earlier. *The home becomes legible:* movement patterns, variance from baseline, bathroom usage, fridge alerts, ambulation, sleep, medication adherence, heart-rate variability, gait deterioration, voice changes, weight trends, blood pressure, glucose, oxygen saturation, and all the weird little signals that, in isolation, look like noise but in combination may reveal a patient starting to fall off her clinical trajectory.

There's a strange historical irony here, which I rather like. *Clinical AI may return medicine, in some sense, to the home.* Preindustrial medicine was overwhelmingly domestic: the sick were treated in the home, cared for by family, visited by clinicians when clinicians were available, embedded in a social context that medicine later abstracted away. The hospital then became the great industrial institution of care: powerful, concentrated, technologically dense, professionally staffed, lifesaving, and also expensive, *depersonalizing*, and overused as a site for things that didn't always require institutionalization. Now, with intelligence diffused through devices, sensors, agents, remote monitoring, and AI-enabled escalation, we may return some portion of medicine to the home—but this time with the diagnostic and monitoring capacity of a much larger system wrapped around it. A preindustrial site of care, reanimated by postindustrial intelligence. Heal in place. Age in place. Die in place, when that's the humane and desired thing. There are worse futures.

This is the great GenAI deinstitutionalization. Costs should come down as more care safely decants from high-fixed-cost institutional settings into virtual, hybrid, community-based, and home-based venues. The labor model changes too. As AI lowers the fixed human cost to serve, the economics of alternative care venues improve. Home care, hospice at home, hospital at home, chronic-disease monitoring, post-discharge surveillance, behavioral-health companionship, medication support—all of these become more coherent when wrapped in continuous intelligence rather than episodic human check-ins. Roughly \$265 billion of services for Medicare beneficiaries are expected to move into the

home in the coming years, driven by deregulation, unambiguous patient preference, the desire to age and heal in place, and an institutional-care setting the pandemic stigmatized in ways we haven't fully worked through.⁹¹ Clinical AI reinforces every one of those tailwinds.

The payer strategy is already visible. Ambulatory and home assets become more valuable when they can be orchestrated by intelligence. UnitedHealth's vertical integration of LHC and Amedisys signals where this can go. Paradoxically, that may make unaffiliated home-care companies more partnerable for health systems that don't want the payer to own every home-based touchpoint (TowerBrook has invested in *Compassus*, and so naturally I'm a big evangelist for this home-care company). Other PE-backed and private home-care companies may become flexible nodes in a more distributed care architecture. Public home-care aggregators may face the same suspicion all vertically integrated healthcare assets face when they sit inside payer empires. But the direction is unmistakable: *home is no longer peripheral*. It's the strategic site of care. The hospital shouldn't experience this only as volume leakage. That would be the old reflex. The better posture is to treat home as a clinical substrate to be governed, instrumented, and synchronized. If the hospital doesn't own or partner into the home, someone else will own the data, the relationship, the escalation, the medication adherence, the behavioral-health touchpoint, the longitudinal context, and eventually the premium-dollar economics.

The home-care thesis and the clinical AI thesis are the same thesis at different altitudes.

Precognition and Presymptomology

The Anticipatory Regime

Before we get to the *Universal Doctor*, we need a word for the anticipatory regime that precedes it. I keep coming back to *presymptomology*. Not diagnosis after complaint. Not screening after guideline eligibility. Not population health after the claims feed limps in ninety days late. Presymptomology: the science and practice of identifying the preclinical trajectory before the patient experiences the symptom, names the symptom, reports the symptom, or even knows there's a symptom to report. Again, check out all.health, who is pioneering this notion.

Clinical AI gives medicine a kind of constrained precognition. Not mystical clairvoyance, not prophecy, not the machine as oracle, but probabilistic anticipation at a resolution humans have never had. The model sees sleep deterioration, gait variance, weight drift, voice changes, refill interruption, glucose volatility, portal-language anxiety, subtle lab shifts, ambient behavior change, and a pattern of social withdrawal. The patient hasn't yet fallen. The heart failure hasn't yet decompensated. The depression hasn't yet become suicidal. The cancer hasn't yet announced itself. But the trajectory has changed. *Medicine should act before the symptom becomes an event*.

This will require a new bedside vocabulary because telling a patient about a risk that hasn't yet become a disease is ethically delicate. Doctors are trained to *respond to suffering*, interpret signs, and explain diagnoses. They're less prepared for AI-mediated clairvoyance: 'Your future isn't determined, but your trajectory has changed, and we should intervene.' That sentence needs a pedagogy. It needs consent norms, anxiety safeguards, false-positive discipline, escalation thresholds, and a culture that doesn't convert every probabilistic whisper into a billable panic.

The physician-architect of the presymptomology era must learn to manage anticipatory uncertainty. Some risks should trigger action. Some should trigger watchfulness. Some should be hidden from

⁹¹ McKinsey & Company: "From facility to home: How healthcare could shift by 2025" published on February 1, 2022 (Link)

patients until clinically meaningful because not every probabilistic shard deserves to become a human worry. That's not paternalism; it's information ethics. The Jupiter Brain may see too much. The human physician will need to decide what the patient should bear, what the system should monitor silently, and when the future has become actionable enough to name.

The Universal Doctor

Reader Note: one final echo of the previous chapter: if Future of Science argued that machine intelligence can hold biology more synoptically than we can, the Universal Doctor is the patient-facing form of that claim. This isn't a second epistemology chapter. It's the clinical and institutional design problem that follows from the epistemology.

Let me close this stretch by asking you to suspend disbelief and imagine with me a little. What would an optimal healthcare system look like with near-infinite, almost-free expert resources? What if the moral imperative here weren't simply cost reduction, though heaven knows we need that, but *democratization*? Right now perhaps one billion people on earth have anything like dependable access to a primary doctor, while six or seven billion carry a smartphone. That asymmetry should inspire us, and perhaps haunt us a little. *The form factor of medical expertise is moving toward the phone, not the marble tower.*

The legitimate objection—that a smartphone in rural Bangladesh isn't the same as a smartphone in Manhattan—is fair. The path to the *Universal Doctor* in low-resource settings runs through infrastructure investment: affordable connectivity, edge-computing architectures that don't require continuous cloud access, local language models fine-tuned on regional disease burden, culturally competent interfaces, data-sovereignty frameworks, clinical escalation pathways, community health workers, and some actual grown-up implementation work. None of that is trivial. None of it is science fiction either. The convergence point—where a community health worker in Malawi has access to a clinical AI rivaling the diagnostic capability of an academic medical center in Boston—is further away than the utopians suggest and closer than the cynics will admit.

The reason this is no longer fantasy goes back to the nature of the silicon intelligence I've been describing elsewhere in this essay. These systems are increasingly polymathic, omnidisciplinary, quasi-omniscient in their recall, and analogic in their reasoning. They can move across domains in ways that medicine, with its guilds and subguilds and sub-subguilds, has made almost impossible for humans. They can integrate the cardiology question with the nephrology question, the psychiatric question, the pharmacogenomic question, the social-needs question, the nutrition question, the medication-adherence question, and the "why has this patient's life become impossible?" question. And if Demis is right about the trajectory from *interpolation* to *extrapolation* to *hypothesis*, then the implications for medicine are mind-bending. We don't merely get a model that interpolates within existing medicine. We get one that extrapolates from the known frontier and eventually participates in conceptualizing new science, new medicine, new therapeutic possibilities, new ways of seeing disease itself.

Once AI meets or exceeds the quality and error rate of the top-decile physician in a defined clinical domain, the dream becomes radical: a universal doctor. A polymathic, multispecialty, cybernetic physician intelligence that reunifies what hyperspecialization fragmented. No more endless bouncing from PCP to endocrinologist to neurologist to otolaryngologist to oncologist to whoever else

happens to hold one shard of the puzzle. *From fragmentation to synchronization.* From artisanal silos to a coordinated frontier of medicine. The best care, now effectively reserved for the wealthy, the urban, the persistent, and the well-connected, becomes diffusible to everyone, in every geography, in every language, at every socioeconomic stratum, through the form factor people already have.

The cost of medical expertise, currently hierarchically rationed and painfully hard to access, moves asymptotically toward free. Or more precisely, toward the marginal cost of compute. Not zero. But close enough to feel like a civilizational change. The multispecialty primary doctor becomes, in some sense, ubiquitous. Prevention, behavioral health, early intervention, nutrition, chronic-disease management, medication optimization, social-needs synchronization—all championed by a superhuman longitudinal agent not rationed by how many fellowship-trained humans happen to be practicing in the right ZIP code. By turning healthcare into an information technology, we democratize not just healthcare access but health itself.

That phrase matters: *healthcare as an information technology.* Channeling Ray Kurzweil, a dollar of healthcare in 2000 buys about 81 cents of comparable healthcare in 2025. A dollar of effective compute buys something like 137,000 times the inflation-adjusted equivalent (thank you, Claude, for helping me with my math here). Technology benefits from the law of accelerating returns and rides the exponential. Healthcare has stubbornly refused to, because healthcare is human-labor-intensive, credentialed, regulated, artisanal, institutional, and protected by guild and statute. Blood isn't software. Surgery isn't a download. A frail elder still has to be lifted, bathed, held, listened to, and cared for. But cognition, triage, monitoring, prevention, coordination, education, medication support, behavioral coaching, and much of longitudinal management can start to behave more like information goods. And once parts of healthcare become information goods, they inherit some of the deflationary properties of information technology.

The global implications are enormous. Diffusing *medical superintelligence* at the marginal cost of compute isn't just healthcare policy; it's development policy. Poor health drags down societies through absenteeism, disability, early mortality, lost labor-force participation, informal caregiving, educational impairment, malnutrition, cognitive impairment, and all the other losses that never make for especially moving PowerPoints but bleed civilizations dry. Better health produces more labor-force participation, more schooling, more cognition, more productivity, more resilience, more food security, and more political stability. *Health is growth policy.* If we can deploy medical intelligence globally—against AIDS, malaria, parasitic worms, maternal morbidity, malnutrition, untreated depression, epilepsy, hypertension, diabetes, neurodevelopmental delays—we're not only relieving suffering. We're changing the economic trajectory of entire societies. One can imagine, not entirely fancifully, an AI finance minister and an AI health minister converging on the same conclusion: the best growth strategy is better biology.

And yes, Jevons' paradox applies. *If expertise gets radically cheaper, people will use more of it.* We'll have more healthcare, not less. More questions answered. More monitoring. More prevention. More medication adjustments. More behavioral nudges. More early interventions.

More clinical touchpoints. That's not a contradiction of the labor-substitution thesis. It's part of the equilibrium. Some labor dislocation at the administrative and clerical layer will coexist with more interactions, more prevention, more monitoring, more patient guidance, and more human care in the places where human care still matters most. Expertise moving toward free will stimulate utilization. Good. We have artificially scarce expertise and artificially swollen administration. Clinical AI lets us invert that.

That's the prize. Not just cost reduction, though yes, cost reduction. Not just efficiency. Not just better margins. A better system. A system where the best care isn't hoarded. A system where the science of medicine begins to catch up to its practice. A system where the map of expertise is finally democratized. If that sounds utopian, perhaps it is. But medicine has always oscillated between priesthood and democratization, between artisanal scarcity and scientific spread. We may be entering the most dramatic swing yet.

But the Universal Doctor only works if it remains connected to human care. The smartphone can become the form factor of expertise, but it can't become the whole moral architecture of healing. The point is to democratize intelligence while preserving the human relationship where the human relationship matters most.

And to get there, somebody has to make the turn.

She Who Assumes the Liability, Wins

The Indemnification Thesis

Let me state the central claim, finally, as starkly as I can.

She who assumes the liability, wins.

In the race to deploy clinical AI, the company—or coalition—that moves first to assume product and malpractice-style liability within a validated envelope will gain the decisive comparative advantage of this generation. That move signals confidence. It signals safety. It signals readiness. It tells clinicians and health systems that the creator is willing to stand behind the product with capital, reinsurance, monitoring, and the implicit pledge that the machine has been tested, observed, supervised, and is ready to be answered for. It transforms clinical AI from an advisory curiosity into a deployable clinical instrument. And it forces every other actor in the ecosystem to respond on new terrain rather than continue litigating the old.

The hyperscalers have the best odds. Google, Microsoft, Amazon—their trillion-scale balance sheets aren't financial trivia. They're strategic weapons of a kind no actor in the medical-industrial complex has ever possessed. They can obtain reinsurance at rates startups can't touch. They can absorb litigation. They can acquire, partner, indemnify, and scale. They can do what smaller firms, however brilliant, often can't: stand behind the machine with capital sufficient to make the standing meaningful.

But this doesn't mean the hyperscalers are the only possible liability-bearing actors. A robustly capitalized clinical AI startup with a narrow enough clinical envelope, enough validation, and enough reinsurance can begin to assume liability. Digital Diagnostics already showed the embryo of this model. Enlightened payers can assume or share liability where they control the premium dollar, the clinical workflow, the ambulatory network, and the economics of prevention. Enlightened providers, especially large systems with malpractice captives, can participate in liability-bearing coalitions by defining the deployment envelope, monitoring outcomes, and pricing risk. The future is probably not one actor heroically assuming all risk in splendid isolation. It's more likely a liability stack: hyperscaler balance sheet, startup vertical excellence, payer economics, provider deployment substrate, reinsurance, and captives woven into a structure that finally lets clinical AI move.

That point matters because "she who assumes the liability" isn't only a company-level thesis. It's an institutional posture. It's the difference between admiring a model and deploying a clinical instrument. It's the difference between a startup with a beautiful benchmark and a company with a product the medical system can actually trust. It's the difference between a health system running

pilots and a health system becoming the domestic deployment layer for clinical AI. It's the difference between saying "human in the loop" as a liability incantation and actually assigning responsibility in a way that lets patients benefit from superior performance.

The window for incrementalism has closed. Not because every AI system is ready. Not because risk has disappeared. Not because the models are perfect. *Perfection is the wrong standard.* It has always been the wrong standard. We never imposed it on the human physician, the surgeon, the academic medical center, or the FDA-approved drug. The standard, properly understood, is human equivalence or human superiority inside a validated envelope, with rigorous post-market monitoring, explicit liability assignment, and a willingness to learn from failure rather than litigate it into silence. A lawyerly society can preserve rights and still build. We're in danger, right now, of becoming the land of the eternal no.

The status quo is already drenched in error. The technology is already arriving. The patients are already waiting. The smartphone is already in their pocket.

Here is the thought I want to leave you with. The question we're facing isn't, in the end, strategic. It's not regulatory. It's not even economic, though the economic implications are almost incomprehensible. *It's a moral question.* We have, for the first time in the species' improbable history, a tool capable of bringing the standard of care of a top-decile academic medical center into the pocket of the rural Bangladeshi farmer or the single mother in a Detroit ZIP code where the nearest psychiatrist is a six-month wait and a forty-minute drive. We have the tool. We've built it. It's sitting on a shelf, surrounded by lawyers, waiting for a serious adult—or better yet, a serious coalition of adults—with a serious balance sheet to take it down and stand behind it.

History won't remember your new patient tower. It won't remember your rebranding campaign, your strategic plan, or your latest population-health press release. It'll ask a harder question: did you, in the moment when the tools became available, have the courage to replace a broken, linear, hierarchically rationed sick-care system with an exponential, distributed, longitudinal, synchronized, universally accessible one? Did you have the courage to underwrite the future? Or did you defer to the priesthood, the committee, the quarterly earnings call, the legal department, and the ancien régime of medical orthodoxy?

The only question left is who has the courage to sign the indemnification.

She who assumes the liability, wins.

And the rest of us—the patients, the families, the clinicians, the country, the watching world—will be, in the most literal and the most civilizational sense, in her debt.

CEO-Level Diffusion and the Risk-Embrace Society

The CEO Mandate

A final operational admonition before I summarize: clinical AI isn't an innovation-department initiative. It's not a CIO side quest. It's not something to be delegated to a politely credentialed AI governance committee that meets monthly, says *responsible* six times, and suffocates the future with smiles and professionalized hesitation. *The CEO has to get her hands dirty.* Chair the AI governance committee personally. Put clinical AI on the board agenda. Demand validated envelopes. Demand liability-bearing partners. Demand timelines. Demand measurement. Demand diffusion, not pilots. The bureaucrats and subordinates aren't bad people, but bureaucracies are machines for converting executive ambiguity into institutional delay.

It's irresponsible not to use these tools now where they're already superior or plausibly equivalent inside bounded domains. Not because every tool is ready. Not because the models are magic. Not because risk has disappeared. *Because asymmetric upside is real*, the baseline is harmful, and supervised deployment is the only way the system learns. We need more societal risk embrace, not adolescent recklessness, but adult courage: the willingness to compare AI against the real human baseline rather than against an imaginary medicine that has never existed.

The CEO's job is to create the conditions under which the organization can move: liability architecture, sandbox validation, physician co-design, patient trust, payer alignment, home-data strategy, risk-bearing economics, and an explicit doctrine for when diffusion begins. Waiting for perfect doctrine is how lawyerly societies lose engineering revolutions. The clinical AI leader doesn't ask whether the institution is comfortable. She asks whether the institution is morally serious enough to become uncomfortable in the direction of better care.

The Chapter, Compressed

For those who want the argument distilled—and mercifully spared another several thousand words of my metaphysical interior monologue—here is the chapter in compressed form.

Clinical AI isn't another administrative technology. It's not revenue-cycle automation with a stethoscope. It's not ambient documentation with better manners. It's the *second door* in the broader AI-healthcare architecture I've been describing: administrative simplification, clinical augmentation, computational and synthetic biology, and consumer empowerment. This chapter is about clinical augmentation, but it inevitably bleeds into the other three because clinical AI needs the enterprise simplified, biology accelerated, and the patient reconnected into a longitudinal, personalized, synchronized system.

The old scarcity has ended. Modern medicine was built around scarce human clinical intelligence: scarce expertise, scarce synthesis, scarce access to the full body of medical knowledge, scarce capacity to hold biology synoptically. Credentialing, specialization, prestige, malpractice, reimbursement, and physician sovereignty all emerged inside that scarcity. Now a non-biological intelligence is beginning to dissolve it. The scaffold doesn't automatically collapse, but it has to justify itself again.

Biology outran the biological brain. *Hyperspecialization, homogenization, and bureaucratization* were rational adaptations to cognitive defeat. They weren't proof that the system was healthy. They were proof that medicine could no longer be held together by synoptic human intelligence. Clinical AI offers the first plausible path back toward synoptic medicine.

The status quo isn't safe; it's merely familiar. Roughly 371,000 deaths and 424,000 permanent disabilities annually from diagnostic error alone—something like 30,000 preventable deaths every month— isn't an ethically innocent baseline. The 17-year graveyard from discovery to bedside isn't "translation." It's tolerated harm. The 17-day mandate is the directional target: validated clinical knowledge should move toward patients in days or weeks, not decades.

We're in medicine's AlphaZero moment. Move 37 matters because it showed a machine doing something the human canon regarded as wrong, and being right. AlphaFold matters because it moved that lesson from game space into biology. Clinical AI is the next migration: models that don't merely remember medicine, but begin to design and discover medicine.

The god-model invasion is here. Co-Clinician, ChatGPT for Clinicians, OpenEvidence, GPT-Rosalind, Gemini, Claude, and the rest of the pantheon aren't extracurricular curiosities. The Mag 7's \$22 trillion in market cap and the hyperscalers' staggering capex tell you who has decided healthcare is the prize. The 150 aren't going to out-compute them. The 150 have something else: the bedside, the physician base, the workflows, the trust envelope, the malpractice captives, the patient relationship, and the legitimacy without which clinical AI remains a product rather than medicine.

Pure clinical-AI software is increasingly uninvestable: *unless it has something the god models can't easily replicate.* Thin wrappers will be discarded. Data flywheels, EHR distribution, clinical workflow ownership, regulatory depth, specialty validation, and liability-bearing capacity are the durable moats. Harvey's Law applies: the winning vertical AI company rides the god models rather than competing with them. But clinical AI's version of Harvey's Law requires something legal AI never needed: the willingness to stand behind the answer with capital, insurance, and indemnification.

The bottleneck is liability, not capability. Clinical AI can't scale because no one wants to accept blame when the machine errs. Not the doctor. Not the hospital. Not the developer. Not the payer. Not the malpractice carrier. Not the reinsurer. Solving blame allocation is the master unlock. We can only go at the speed of blame allocation.

The infallibility trap is the philosophical mistake underneath the opposition. We hold clinical AI to an inhuman standard we've never imposed on human medicine. We tolerate human fallibility because it's familiar, professionalized, insured, and dispersed into the background noise of the system. But when the machine errs, the error becomes a scandal. That's not a safety regime. It's incumbent protection masquerading as ethics.

The guild will call it safety. Some of that caution will be sincere, and some of it will be necessary. But much of it will function as delay. Professional self-protection will dress itself in the language of ethics, prudence, equity, oversight, and responsible innovation. The 150 need to learn the difference between legitimate safety and old-scarcity protectionism.

Installation is the differentiator. *America innovates. Europe regulates. China appropriates* and, yes, *deploys.* Autocracies and Gulf monarchies can install major technological revolutions faster than messy, pluralistic, lawyerly democracies. This isn't an argument for autocracy. It's the opposite. The civilizational argument is that we want the deflationary, lifespan-extending, disparity-mitigating benefits of clinical AI diffused here, inside an American moral and institutional framework, not licensed back from regimes whose speed is purchased at the cost of political freedom.

Corporate strategy must precede tort reform. BYD didn't wait. Digital Diagnostics didn't wait. The first mover who assumes liability inside a validated envelope creates the practical framework; the law follows. My bet remains Google—not because the logo matters, but because balance sheet, scientific ambition, DeepMind, AlphaFold, Isomorphic, Gemini, Google Health, and clinical seriousness give

it a plausible path to underwrite the risk. If not Google, then Microsoft or Amazon. The important point is the liability-bearing balance sheet.

The standard of care will invert. Today, deviation from convention protects against malpractice. Tomorrow, deviation from a manifestly superior AI recommendation may become the thing that requires explanation. The question shifts from “Why did you use the AI?” to “Why didn’t you?” That inversion won’t arrive as a philosophical proclamation. It’ll arrive through workflows, order sets, payer rules, quality measures, malpractice underwriting, medical-staff policies, and validated outcomes.

Medicine moves from practice toward science. Channeling Vinod here! Subjective clinical gestalt doesn’t disappear, but it loses sovereignty where models can outperform unaided humans. The physician isn’t morally dethroned; she’s epistemically relocated. Less data-entry clerk. Less solitary diagnostician. More steward, architect, interpreter, trust-bearer, verifier, and moral agent inside a larger intelligence system.

The physician-architect is the new role. Clinicians can’t merely adopt tools designed elsewhere. They have to co-design the clinical logic, safety envelope, workflow, failure modes, escalation paths, evaluation criteria, patient-facing implications, and liability boundaries. Clinical AI must be built with the bedside, not shipped into the EHR as one more indignity from people who’ve never watched a Friday evening discharge collapse under the combined weight of pharmacy, payer, family, transport, and reality.

The sacred trust must survive the intelligence age. Healthcare is consecrated work. If clinical AI becomes an austerity machine with a bedside avatar, this whole transition will curdle politically and morally, and it should. The point isn’t to replace the human relationship with a cheerful simulacrum. The point is to get the nurse out of the chart and back to the bedside, the physician out of the inbox and back into the conversation, the pharmacist out of fax warfare and into medication optimization, the care manager out of manual list-building and into actual intercession.

We need a new Flexner settlement. The 1910 Flexner Report rationalized medicine for an age of scientific professionalization, memorization, hierarchy, gatekeeping, and physically localized expertise. The AI Flexner Report should ask what medical education becomes when omniscience, total recall, simulation, interpolation, extrapolation, and hypothesis generation sit in the pocket. Stop selecting primarily for memorization. Select for judgment under uncertainty, ethical reasoning, human connection, adaptive learning, epistemic humility, and the ability to work fluently with systems that will surpass humans in raw recall.

Capitation beats fee-for-service for clinical AI. Fee-for-service pays for activity; clinical AI’s most beautiful use cases often destroy activity. Capitation, delegated risk, Medicare Advantage, Medicaid risk, case rates, advanced primary care, and provider-sponsored plans make prevention economically legible. This time the tools finally match the theory. Whoever owns the delegation wins.

Advanced primary care may get its resurrection. The old APC thesis wasn’t wrong. It was under-instrumented. Clinical AI gives team-based, longitudinal, risk-bearing primary care its missing operating system. Provider-sponsored health plans shouldn’t be sold reflexively. Premium dollar plus ambulatory assets plus employed or aligned physicians plus home-based care plus pharmacy plus behavioral health plus clinical AI may become one of the most valuable combinations in healthcare.

Home becomes the epicenter. Clinical AI, multimodal models, wearables, sensors, computer vision, and home-monitoring infrastructure make it possible to predict decompensation and intervene earlier, remotely. Heal in place. Age in place. In a strange historical irony, the most advanced intelligence technology in history may return medicine to its preindustrial site of care: the home, now

wrapped in postindustrial intelligence.

The end state is the *Universal Doctor*. A polymathic, multispecialty, cybernetic physician intelligence that reunifies what hyperspecialization fragmented. The best care, currently rationed by ZIP code, income, race, geography, institutional proximity, and social capital, becomes diffusible at the marginal cost of compute. The smartphone becomes the form factor of expertise. The rural Brazilian farmer, the single mother in Baltimore, and the community health worker in Malawi all get access to something that begins to approximate the reasoning capacity of an academic medical center.

Turning healthcare into an information technology democratizes health, not merely healthcare access. Healthcare has refused the deflationary logic of technology because it's human-labor-intensive, credentialed, regulated, institutionalized, and guild-protected. Blood isn't software. Surgery isn't a download. But cognition, triage, monitoring, prevention, coordination, medication support, behavioral coaching, and much of longitudinal management can start to behave more like information goods. That's where the exponential finally enters medicine.

Jevons' paradox will apply. If medical expertise gets radically cheaper, people will use more of it. We'll have more healthcare, not less: more questions answered, more monitoring, more prevention, more nudges, more medication adjustments, more behavioral coaching, more early interventions, more clinical touchpoints. That's not a contradiction of labor substitution. It's the humane equilibrium: less human labor wasted on administrative sludge, more human attention available where presence, trust, judgment, and love still matter most.

The Healthcare 150 must become the clinical AI deployment layer. They won't own the frontier model layer. They can own the deployment layer. Their work is to create board-level command structures, physician-led co-design labs, validated clinical envelopes, liability-bearing partnerships, malpractice-captive strategies, instrumentation of model use and override, employed-physician diffusion channels, risk-bearing payment models, home-based care substrates, and sacred-trust governance.

Medicine may become engineering, not merely science. If biology is functionally verifiable—if interventions can be recursively proposed, simulated, tested, and measured against ground truth—then medicine becomes a design-build-test-learn discipline at Jupiter Brain scale.

The Bitter Lesson applies to the body. Human medical wisdom is indispensable early and dangerous when fossilized. More data, more compute, more inference, better simulation, and better verification will eventually discover patterns the canon can't.

The world needs a great medical data project. Clinical, claims, imaging, pharmacy, multiomic, behavioral, home, and wearable data should be organized through privacy-preserving, consented, sovereign-aware infrastructures as a national and multinational humanitarian project.

The small-model lead is real but may be temporary. Domain models like VIRCHOW matter; so do platforms like OpenEvidence and Hippocratic AI. But cheap tokens, frontier reasoning, data access, workflow, liability, and distribution will decide durability.

Clinical AI needs sandboxes. Retrospective validation, synthetic stress tests, shadow mode, biological twins, continuous post-market learning, and simulation-heavy trial frameworks should compress validation without pretending patients are software users.

The standard of care becomes a hive mind. Individuated clinical agents will share a common learning substrate, updating standards continuously rather than waiting for wood-paneled committees and biennial PDFs.

Home is de-institutionalization. Continuous, high-velocity, personalized data turns the patient into an N-of-1 learning system and moves intelligence from the hospital's walls into the life of the person.

Presymptomology is the next clinical frontier. The Jupiter Brain will see trajectories before symptoms; doctors must learn the ethics and communication of probabilistic clinical precognition.

Clinical AI may resurrect value-based care. The thesis wasn't wrong; the operating system was missing. Agentic longitudinal medicine makes prevention economically and operationally plausible.

CEOs must lead diffusion. Clinical AI can't be delegated into paralysis. Board-level command, liability architecture, sandbox validation, physician co-design, and risk-bearing deployment are executive responsibilities.

She who assumes the liability, wins. But she'll probably not win alone. The winning architecture is likely a coalition: hyperscaler balance sheet, startup vertical excellence, payer economics, provider deployment substrate, reinsurance, and malpractice captives woven into a structure that finally lets clinical AI move. Intelligence, liability, bedside.

What she does with the winnings is the civilization question. Clinical AI can become a margin-recapture machine, an austerity machine, a denial machine, or a surveillance machine. Or it can become the instrument by which medicine is reconstituted: less fragmented, less bureaucratic, less geographically rationed, less dependent on human memory, less tolerant of invisible harm, and more worthy of the sacred trust at its center.

That's the choice. Clinical AI is coming either way. The question is whether the 150 shape it, govern it, indemnify it, diffuse it, and humanize it—or whether they wait politely while someone else does.

BEFORE WE TURN THE PAGE

Clinical AI shows what happens when machine intelligence enters the act of medicine itself. But before healthcare leaders can decide how to deploy it, they have to understand the broader labor convulsion underneath the entire economy. The next chapter steps back from the clinic to the macro bargain: what happens when the machine learns the ladder workers were supposed to climb?

Chapter 4 • AI Labor Convulsions

When the Machine Learns the Ladder

“The instrument of labour, when it takes the form of a machine, immediately becomes a competitor of the workman himself... machinery not only acts as a superior competitor to the worker... but it also makes him superfluous.”

— KARL MARX, DAS KAPITAL, 1867

A WORD ON NAVIGATING THIS CHAPTER

This chapter is the broader labor storm system before it makes landfall in healthcare. It asks whether GenAI breaks the old technological bargain by attacking cognitive scarcity itself, and why the prestige professions, apprenticeship ladders, and the intelligentsia should be much less comfortable than they are.

Another Note to My (By-Now-Exhausted) Reader

Friends, this labor chapter—the first of two parts—is, with intention, the broader and more meditative one. If you’re here only for the hard-hitting healthcare operating implications—revenue cycle, payer bots, clinical labor, the 150’s playbook, attrition strategy, management-layer compression, what to do with open recs, and all the other things that make hospital executives sit up a little straighter during otherwise perfunctory board meetings—you can go ahead and skip to the next chapter without violating any law of God, literature, or management consulting. I won’t be offended. Much.

Part II of this is the healthcare chapter. But I feel like I should preface that chapter with something more upstream first.

This chapter is about labor, yes, but not only labor. It’s about capital, cognition, class, expertise, professional identity, apprenticeship, technological salvation, concentrated harm, the old bargain by which modern societies survived automation, and the possibility that GenAI breaks that bargain by attacking the one scarcity that kept *reabsorption* plausible: *human cognition*. It’s the chapter about the storm system before landfall. The next chapter is where the storm hits healthcare.

So if this chapter feels more sociological, philosophical, historical, and occasionally macroeconomic than operational, that’s by design. I want the reader to understand why the healthcare labor argument can’t be reduced to “AI will automate some tasks” or “health systems should use copilots.” Those are small statements. True, maybe, but small. The larger claim is that we’re witnessing the *industrialization of intelligence* inside a service economy built around expensive human cognition. And if that’s right, then healthcare isn’t merely adopting a technology. It’s about to confront a change in the production function of care.

That’s the argument I want to build here: slowly enough that it doesn’t degenerate into panic, but forcefully enough that nobody can mistake the stakes.

Marx, Catastrophically Wrong—Except About the Machine

I'll start this labor meditation with Marx, not because I've suddenly converted, in my early middle age, to historical materialism, the labor theory of value, or any of the other dreary and poisonous doctrines that did so much harm once they migrated from unreadable tomes to state power. The man was catastrophically wrong about a great deal, and catastrophically wrong in ways that mattered: morally, politically, economically, and, in the twentieth century, murderously. And we'll see his most recent incarnation—in New York City, California, and with increasing recidivism in other geographies across the US in the ballot box—similarly refuted across time (let's hope sooner rather than later, before too much damage is done).

But on the point of this chapter's epigraph—on machinery becoming not merely the *tool* of the worker, but the *competitor* of the worker—Marx saw something with prescient clarity. The machine doesn't simply amplify labor. At a certain point, it rivals labor, disciplines labor, reorganizes labor, and in some domains makes labor superfluous. That sentence, sitting at the top of this chapter, lands here with more force than I'd like.

Let me state the argument as directly as I can, because the AI-and-labor debate gets chaotic almost immediately. We're snowed in by euphemism, corporate and government propaganda, bad-faith reassurance, TED-talk optimism, and the now-familiar ritual in which very smart people explain why a technology explicitly designed to perform human cognitive work will somehow *not* displace human cognitive workers. I've said this repeatedly throughout the essay: technology has been humanity's great engine of liberation, and I'm defiantly not joining the guild of professional doomers who treat every new machine as a sacrilege. *I love technology*. I think it has saved us, repeatedly over our civilizational history, from scarcity, disease, ignorance, drudgery, distance, darkness, and the impersonal subjugations of nature.

But GenAI isn't merely another machine in the long procession of machines. I've characterized it in other chapters as a *multiplication of intelligence*. For this chapter on labor, I'm going to switch up the formulation a bit and call it the *industrialization of intelligence*. The distinction matters because this technology doesn't merely touch the modern service economy in all of its imposing magnitude (an estimated \$60 trillion out of the \$127 trillion global GDP).^{92,93} It goes after the substrate on which that economy has been built: expensive human cognition, coordination, narration, verification, judgment, and administrative intermediation. In other words, it comes not only for the *bottom* of the occupational hierarchy, where mechanization historically began, but for the *top*.

The historical pattern is familiar enough that we can compress it. The steam engine came for muscle. The spreadsheet came for arithmetic. The search engine came for retrieval. The loom obsoleted the weaver; the automobile dismantled the horse-powered economy; the ATM changed the bank branch; the internet disintermediated some clerical and retail work and then created whole new occupational ecosystems. *That was the old bargain*. Destruction first, regeneration later. Channeling Ray Kurzweil: *concentrated pain, diffuse redemption*. Harsh in the middle, but ultimately expansionary. And, taken over the long arc, good for progress and civilization.

GenAI is altogether different. It comes for synthesis, judgment-adjacent work, pattern recognition, professional narration, and the priestly scarcity of expertise. That's why the Marx line is so relevant and disconcerting. *The instrument of labor has once again become a competitor of the worker*, only

⁹² World Bank Group: "Services, value added (% of GDP)" accessed on June 9, 2026 (Link)

⁹³ International Monetary Fund: "IMF Data Explorer" accessed on June 9, 2026 (Link)

this time the worker isn't the weaver, the blacksmith, or the carriage driver. This time the worker is the analyst, the coder, the claims specialist, the consultant, the administrator, the physician, the nurse performing documentation, the lawyer, the revenue-cycle employee, the care coordinator, the credentialed knowledge worker whose economic value depended on cognition remaining scarce.

So the structure of my argument is simple, even if the implications are anything but. Technology's historical bargain with labor was cold, remorseless, and often cruel, but, over time, *regenerative*. It destroyed jobs, guilds, communities, and entire ways of life, and then, usually after some period of dislocation and suffering, reabsorbed workers into *new bottlenecks* (more—much more—on this concept later). The worker moved from the hand to the machine, from the farm to the factory, from the factory to the office, from clerical routine to knowledge work. That was the ladder. The machine attacked one layer; labor climbed to the next.

The uncomfortable thesis of this chapter is that GenAI may break that bargain, not because it automates another set of tasks—every major technology does that—but because it attacks the *cognitive scarcity* that historically allowed the worker to move somewhere else. It's coming not only for the task, but for the reason the worker was scarce in the first place. It's coming for the intelligentsia, the apprenticeship professions, the prestige labor edifice, the knowledge-work categories that were supposed to be protected from mechanization because they lived in the *neocortex* rather than the *hand*. If the steam engine amplified muscle, and the microprocessor amplified calculation, GenAI amplifies—and then begins to substitute for—the very faculty that made lawyers, consultants, analysts, administrators, coders, doctors, and institutional bureaucrats economically scarce. That's something altogether different than another labor-market adjustment. It's a category break.

And yes, I'm very aware of the vigorous, inflamed, often ad hominem debate over whether AI will be primarily *augmentative* or *substitutive* of human labor. That debate is real. It's nuanced, complex, emotionally fraught, and still beset by ambiguous data, vendor propaganda, ideological priors, and the usual fog that surrounds a technology before the labor market has finished metabolizing it. So let me acknowledge the uncertainty upfront: this won't be one thing everywhere. There will be augmentation, substitution, task recomposition, new demand, new bottlenecks, and all the messy middle-state phenomena economists quite properly insist we take seriously.

But I'll foreshadow where I come down, *especially in US healthcare*: I think the effect will fall asymmetrically toward *substitution*. Not instantly, not uniformly, not without resistance, not without propaganda, not without demagoguery and not without new work appearing elsewhere in the system. But directionally, and with enough force that the 150 need to stop treating this as futurism and start treating it as a board-level operating assumption.

And if that is even partially right, then we need a sober and methodologically rigorous approach—to say nothing of sensitivity and compassion—as we map what can be automated, what should be augmented, what must be protected, what has to be redesigned, and what we owe the human beings whose old work the machine may soon make unnecessary.

The Dilemma: Margin Recapture or Labor Reconstitution?

This is where the argument becomes moral, not merely economic.

The central question isn't whether AI will be powerful enough to substitute for human labor. In enough domains, it already is, and the exponent is still moving—vertically, and not by gentle, linear increments. The deeper question is whether institutions that benefit from replacing expensive human labor with machine intelligence can do so without betraying the people who built those institutions.

Not: can we automate? Of course we can automate. The question is whether we can *reconstitute labor* rather than merely eliminate it.

That phrase matters. *Labor reconstitution* isn't a euphemism for layoffs, or at least it had better not become one. It's the recognition that the underlying unit of economic value is changing. Some work *should* disappear because it's stupid, duplicative, spiritually deadening, or an artifact of broken systems. Some roles should shrink because they exist only to compensate for bureaucratic encrustations. Some human capacity should be redirected toward work that remains genuinely and irreducibly human: trust, touch, presence, persuasion, mercy, judgment under ambiguity, leadership, moral accountability, and the hard social work of helping other human beings actually live differently. And some workers, let's be honest, will need a generous transition because the new system simply won't need their old role in anything like the same quantity.

The lazy version of AI labor strategy—the one now being advertised by a growing number of billionaire tech CEOs announcing AI-driven layoffs in a tone of faintly unbearable self-congratulation, in many cases using AI as a blanket to cover more *analog* strategic missteps—is *labor reduction for margin*. The serious version is *labor reconstitution for abundance*. That distinction will matter enormously in healthcare, but it matters beyond healthcare too. She who converts labor savings into affordable abundance and humane transition wins; she who converts them only into margin invites revolt, and perhaps a level of societal destabilization we're not yet being honest enough to contemplate.

I know that sounds a little armageddonish. But a sector that employs roughly 24 million people across healthcare and social assistance—about one out of every seven US workers—anchors regional economies, absorbs trillions in labor expense, and props up large parts of the middle class can't quietly substitute capital for labor without political and social consequences. That's magical thinking. Or worse, spreadsheet thinking.

And since last year's *Juggernaut* essay, when I first speculated on these impending labor dislocations, the data have become harder to wave away. Still noisy, still conjectural in important ways, still viewed through the chronically fogged-up and imprecise instrument panel of modern macroeconomics, but materializing nonetheless. AI is now distorting GDP, trade flows, capex, equity concentration, labor share, market sentiment, and the emotional experience of the economy. That distortion is very much the point. A technology that substitutes capital for labor should, if this thesis is right, show up first as a capital boom, a compression of labor's share, a divergence between GDP and employment, and a confused, anxious, despondent public living inside a nominally strong economy while wondering why the prosperity somehow doesn't feel like theirs. And that's exactly what we're starting to see.

So I want to strengthen the chapter's core claim with an empirical hinge: this is no longer only a theory of future substitution. The *great divergence* between capital and labor, productivity and compensation, GDP and employment, has begun to show up in the data. Not conclusively. Not cleanly. Not beyond dispute. The evidence is still noisy, still vulnerable to revision, still entangled with capex circularity, pandemic overhang, Fed chaos, tariffs, war-induced inflation, sectoral weirdness, and the usual macroeconomic fog machine. But the directional signal is harder to dismiss than it was a year ago.

Predictably, partisans and ideologues on all sides are already seizing on the ambiguity to propagandize their preferred conclusions: AI is adding hundreds of basis points to GDP growth! AI is just a bubble! AI is salvation! AI is theft! AI is the end of work! AI is a scapegoat CEOs are using to excuse

old-fashioned layoffs! Fine. Let's have the argument.

But abstract out of the noise and the directional conclusion is hard to escape: it's time for serious leaders to stop treating AI labor substitution as some academic hypothetical and begin planning for a post-diffusion labor architecture. Not because every job disappears. Not because the data are dispositive with the satisfying neatness of a regression table delivered from heaven. But because the risk is now large enough, proximate enough, and asymmetrically consequential enough that leadership requires treating it as a high-probability outcome on the distribution table.

If capital can increasingly instantiate cognition, and cognition can increasingly perform labor, then every labor-intensive institution has to rethink what labor is for.

That's the architectural spine of this chapter. I begin with the old technological bargain, because a labor argument without history becomes either Luddite panic or techno-utopian evasion. I then ask why AI may *break that bargain*, because cognition—not just labor—is now being industrialized. From there I move into the macro decouplings: *population from progress*, *GDP from employment*, *productivity from compensation*, and *capital from labor*. Then I turn to the prestige professions and the apprenticeship problem, because the first great psychological shock of AI labor substitution will come not from the automation of low-status labor, but from the automation of the people who thought credentials, intelligence, and proximity to difficult knowledge would keep them safe.

Only after that can you and I responsibly talk about healthcare. That's the next chapter.

For now, the work begins by understanding the old bargain before we decide how, and whether, it can survive the new machine.

Labor, Prestige, and the Capital-Labor Inversion

Reader Note: This reprises the multiplication-of-intelligence argument from Chapters 1 and 2, but changes the object from science to labor. The same cognitive substrate that alters discovery now threatens the prestige professions and the apprenticeship ladder.

The Automation of the Intelligentsia and the Prestige Professions

There's something categorically different about this one. I went into great depth on this question in the *generative epistemology* chapter, so I'll be more succinct and to-the-point here: this particular GPT is coming for the intelligentsia. It threatens a *commodification* and *demonetization* of the prestige professions—those perched at the top of the social, cultural, and financial hierarchies.

Past automations mechanized human *musculature*. We engineered amplifications of our human and animal physical power to effectuate things in the world: Arkwright's mill, Watt's steam engine, Bessemer steel. These technologies acted in the world of *atoms*. They gave us dominion over space and time, annihilating distance with canals and railroads and internal combustion engines and airplanes, multiplying our physical powers of production and allowing humanity to reshape the material world on a civilizational scale.

Then came the advent of technologies that *automated low-level cognition*. The microprocessor enabled the calculator, the spreadsheet, and eventually the entire architecture of the modern information economy. These technologies were *augmentative of our cognitive power*; and to the extent

they were substitutive of tasks humans previously performed, the displacement occurred largely toward the less sophisticated end of the skill spectrum. Clerks, typists, bookkeepers, and other routine information workers found portions of their labor absorbed by machines, while the prestige professions—the lawyers, doctors, consultants, scientists, engineers—remained largely insulated from the technological turbulence below them.

But this is different.

If this technology truly represents a *multiplication of intelligence*—the very property that has assured our supremacy as a species and determined the stratification of those within—then the implications are qualitatively and categorically *different* from every previous automation. And so it's harder to simply extrapolate from past automations and 'play it forward' for this one. This one threatens to democratize high cognition itself. It threatens to demonetize expertise by making it ubiquitous. It's coming for cerebral, non-routine, analytic activities that have long been the province of high-status, high-compensation knowledge workers. Accountants, lawyers, consultants—and yes, doctors—have little immunological protection against this development. Ironically, unlike their blue-collar counterparts who have historically collectivized and organized into unions with some measure of deterrence and reprisal, this dislocation—and perhaps disemployment—is impending for many within the professional classes. A commodification of domain expertise. A leveling, democratizing force.

This is coming for the intelligentsia. The cognoscenti. The knowledge workers.

Sure, the absolute pinnacle of each profession will continue to dominate for the foreseeable future—an *amplification and force multiplier for the preeminent*. The best surgeons, litigators, researchers, and strategists will become dramatically more powerful when paired with intelligent systems that extend their reasoning capacity and radically increase their productivity. But for the broader distribution of professional labor, the dynamic is likely to be very different. These systems will preponderantly elevate the competency of the *least skilled* while compressing the distance between the bottom and the upper quartiles of performance. The middle of the distribution—where much of the economic value of these professions historically resided—may find itself squeezed. Compensation gradients flatten. Expertise becomes less scarce.

Even this stage may prove ephemeral. It may represent merely a *transitional phase* in which humans and machines coexist in an uneasy cognitive symbiosis. Depending on the domain—and particularly on how much *verifiability* (yes, that recurrent theme) exists as a core property of the work—the next transition may arrive quickly on its heels: full automation. Not merely automating the workflow but automating the worker himself.

And when intelligence itself becomes abundant, the scarcity on which many of our most prestigious professions were built inevitably begins to dissolve.

Ghost GDP, Narrative Instability, and Creative Destruction

There's an extraordinary amount of instability in the system right now. I believe we're entering a period of sustained macroeconomic and institutional destabilization, and, to state the obvious, it's impossible to predict with any kind of precision how events will unfold. One can *feel* the volatility not just in the underlying technology, but in the *narrative layer* that sits on top of it. As I write this paragraph, in early June of 2026, the 'AI trade' is ripping—the S&P 500 notched a 16% gain over the previous 60 days, a parabolic jump seen only four times since 1950.⁹⁴ Investors, analysts, policymakers, and executives are all attempting to interpret a technological transformation whose economic consequences are still largely opaque. The result is a strange and jittery informational environment in which the narrative itself becomes a source of market instability. A Substack essay, a long X thread, or a speculative research note can propagate through financial markets with hyperdrive speed, triggering convulsions in valuation that resemble something closer to a modern-day bank run than a traditional, sober reassessment of fundamentals.

We saw an early glimpse of this dynamic in the market reaction last year to DeepSeek and other AI releases, where trillions of dollars in market capitalization evaporated almost overnight as investors attempted to reprice the implications of new technological capabilities. Or again, earlier this year with the stylized, apocalyptic Citrini memo—fascinating that a few thousand words of fictional writing can vaporize hundreds of billions in market value. What's remarkable isn't just the *magnitude* of those moves but the *speed* at which they occur. In a world where the underlying technology evolves on exponential curves, the *interpretive narratives* surrounding that technology evolve just as quickly. The result is a feedback loop in which expectations about the future of labor, productivity, and economic growth swing violently from optimism to catastrophe and back again. One day the dominant story is AI-driven hyper-productivity and abundance; the next day it's mass unemployment, demand destruction, and the collapse of consumer spending.

This is where the notion of "Ghost GDP" begins to enter the conversation. If AI systems become capable of producing enormous quantities of economic output with minimal human labor—writing code, generating designs, conducting research, synthesizing analysis—the question arises: how does that output register in *traditional* economic statistics? GDP measures market transactions, not necessarily value creation. If machines begin generating vast amounts of useful work that is either free or priced near zero, the result may be a strange *disjunction* between real productive capacity and measured economic output. In other words, the economy may become far more productive than our conventional metrics suggest, while the labor income that historically supported consumption begins to erode. That's the scenario some analysts are gesturing toward when they talk about Ghost GDP: an economy where productivity explodes but the traditional mechanisms for distributing income lag behind.

It's not yet clear whether that outcome will materialize, but the mere plausibility of it has been enough to rattle investors and policymakers. Markets now oscillate between competing macro narratives, each attempting to extrapolate the implications of machine intelligence. One narrative imagines a productivity boom reminiscent of the postwar decades. Another imagines widespread technological unemployment and collapsing demand. The truth will almost certainly land somewhere in between, but the important point is that the *uncertainty itself becomes destabilizing*. When the foundational variables of an economy—labor, productivity, capital allocation—are in flux, the *stories* people tell about the future can become as consequential as the underlying fundamentals.

⁹⁴ Yahoo: "Stocks just notched an achievement rarely seen since WWII. But history says it's also a warning" published on June 2, 2026 ([Link](#))

If we step back from the day-to-day noise of the markets, however, a few structural trends begin to emerge. The vertigo-inducing pace of technological change, the historical lessons about *appropriability* and *diffusion* discussed earlier, and the sheer scale of the industries involved all point toward a similar conclusion. During the *installation phase* of a general-purpose technology, *power tends to concentrate rather than disperse*. Autocracy—at the level of countries, companies, and CEOs—tends to outperform consensus-driven governance when institutions are reorganizing themselves around new technological infrastructure.

In practical terms this likely means that **the healthcare industry will become more oligopolistic, more concentrated, and more personality-dominated** than it already is. If the Healthcare 150 possessed disproportionate influence before, that influence will likely be *magnified* during the next phase of the transition. Decision-making authority will concentrate in fewer hands as institutions attempt to move faster than their competitors in diffusing AI across their workflows. In fact, it wouldn't surprise me if the number itself gradually compresses. The Healthcare 150 may become the Healthcare 100, and eventually perhaps the Healthcare 50.

Consolidation will be one mechanism driving that shift. Larger health systems will absorb smaller ones that lack the capital, technological capability, or *executive will* to keep up. The regulatory environment may also evolve in ways that accelerate this process. A more permissive antitrust stance from a more *laissez-faire* FTC or DOJ—one less enamored with the aggressive, ideological anti-consolidation posture of the past few years—will allow a new wave of *non-contiguous health system mergers*, creating multi-state healthcare empires with unprecedented geographic reach. My expectation is that we'll see the largest redistribution of market share among health systems in a generation, beginning sometime in late 2026 and accelerating into 2027 as AI-proficient institutions begin to pull away from their less technologically capable peers. The gap between the “haves” and the “have-nots” will widen quickly, and some organizations will inevitably find themselves looking for rescue.

All of this reflects a broader dynamic that economists have long associated with technological revolutions: *Schumpeterian creative destruction*. New technologies rarely distribute their benefits evenly. Instead they create violent periods of industrial reorganization in which some institutions (or nations) expand rapidly while others collapse. Interestingly, this may represent something of a reversal from the pattern I described in my essays of 2024 and 2025.⁹⁵ In those earlier analyses the first phase of the AI revolution appeared to be characterized by *creative agglomeration*—the extraordinary concentration of economic value in a handful of hyperscalers, chip manufacturers, and large technology platforms that captured the early *appropriability trade*. The second phase may look quite different.

If the early chapter of the AI economy belonged to the infrastructure providers—the cloud companies, GPU manufacturers, and model builders—the next chapter may belong to the industries that reorganize themselves most aggressively around the technology. In other words, the agglomeration phase gives way to the creative destruction phase, where entire sectors reconfigure themselves around the capabilities that the infrastructure layer made possible.

And healthcare, given its scale, complexity, and labor intensity, may be the arena where that transformation plays out most dramatically. Which, inside the institution itself, means the center of gravity begins to move.

⁹⁵ Eric Larsen: “Pulling Back the Cosmic Veil on Generative AI” published in January 2024 ([Link](#))

Center of Gravity Shifts: First from Labor to Management, Then from Management to Capital

Again, perhaps an unpalatable framing. But the message is the same: CEOs who are autocratic, unambiguous, and decisive about how their enterprises will embrace and exploit AI will outdistance—quickly—those who remain ambivalent or excessively deferential to staff. During the installation phase of a general-purpose technology, diffusion rewards speed and clarity rather than consensus. The lesson of the Industrial Revolutions we discussed earlier wasn't merely that *invention* matters, but that *diffusion* and *mobilization* ultimately determine who captures the economic spoils. In this phase of the AI transition the balance of power inside institutions is already beginning to shift—first from labor to management, and eventually from management to capital.

Healthcare has historically been one of the most *labor-deferential* sectors in the American economy. Some of this deference is cultural and genuinely altruistic. Medicine, after all, isn't merely an industry; it's a *vocation* and *consecration* wrapped in a moral narrative about care, compassion, and service. But some of this dynamic is also pure *realpolitik*. The political influence of nursing unions, which have become increasingly strident and ideological, has grown steadily over the past two decades. The cultural and historical independence of physicians predates the modern hospital system itself. And healthcare CEOs have evolved accordingly into a particular leadership archetype: statesmen and women, diplomats, philanthropists, coalition builders—figures who *navigate* complex stakeholder ecosystems rather than impose unilateral direction.

But the *multiplication of intelligence* that we've been discussing throughout this essay begins to destabilize that equilibrium. When machine intelligence begins to diffuse across the workflows of an industry whose *largest cost center is disproportionately human labor*, the internal balance of power inevitably shifts. The institutions that move fastest in this diffusion phase will be the ones whose leadership teams recognize that the underlying economics of labor are changing in real time.

For a moment—an ephemeral but consequential one—the center of gravity will shift not only from labor to management but also *from clinicians toward administrators*, and specifically toward the CEO. Right now, power in healthcare largely resides with the *guild*. Credentialism, the prestige of the medical profession, and the atomized authority of highly specialized clinicians form the backbone of the system. But that balance will gradually evolve as machine intelligence begins to permeate clinical knowledge itself. As models become more capable and *almost-omniscient*, as new scientific discovery accelerates, and as AI-assisted molecular design and biologics proliferate, the informational asymmetry that historically privileged individual physicians begins to narrow. The societal regard for doctors will not disappear—it *should not* disappear—but the epistemic monopoly that once defined the profession will slowly erode.

In that environment the locus of power moves toward the institutions capable of orchestrating the technology. Healthcare systems, insurers, and integrated delivery networks become the places where intelligence, capital, and workflow redesign converge. The CEO therefore begins to resemble less a ceremonial steward of a nonprofit institution and more a systems architect presiding over an increasingly AI-mediated enterprise.

But before we start to wade through the probable corollaries and consequences of this, I think we ought to take a step back first.

Part I—The Old Technological Bargain

Technology as Salvation, Labor as Toll

Before this becomes a workforce plan, a board-level AI agenda, a set of CEO imperatives, or one more depressing enumeration of open requisitions, severance obligations, union counteroffensives, and management-layer compression, it has to be a moral and ethical accounting. That's why I think we ought to start historically. Not because I've some uncontrollable antiquarian impulse—although, as my patient reader will have noticed by now, I do—but because the labor question can't be answered responsibly unless we first understand the *old technological bargain*.

For roughly two and a half centuries, really since the advent of the first Industrial Revolution at the end of the eighteenth century, technology has cut more or less the same cruel but ultimately regenerative deal with labor: *destruction now, reabsorption later*; concentrated pain, diffuse abundance; some people ruined by the transition, many more lifted by the resulting surplus. That bargain is the background assumption beneath almost every *optimistic* argument about automation. It's why the economists tell us not to panic, why the technologists tell us to accelerate, why the market tells us to reallocate, and why the displaced worker, quite reasonably, suspects that everyone giving the sermon is unlikely to be the one losing the job. The central question of this chapter is whether GenAI still fits that bargain, or whether it changes the bargain itself.

This matters because the conversation otherwise becomes stupid and caricaturist very quickly. One side says technology is salvation and waves away the human wreckage as the regrettable but necessary cost of progress. The other side treats every new machine as sacrilege, as though human flourishing would be best preserved by cryogenically freezing the world at whatever technological plateau happened to employ the current constituency of workers. Neither posture is serious enough. Technology has absolutely been humanity's deliverance. It has also always produced unintended tragedies. The task is to hold both truths at once long enough to make intelligent decisions rather than ideological ones.

So let me begin with the basic premise and declaration: I love technology. Frankly, we all should. And yes, there's a lot of demonization of it lately, but that isn't new either. It's an atavistic human phenomenon: *we love and hate our tools*, we build them and then resent them, we depend on them and then accuse them of corrupting some imagined prior purity. I think Kurzweil has this essentially right: the benefits of technology diffuse societally, often massively and beautifully, while the harms are specific, concentrated, and sometimes cruelly visited upon particular workers, guilds, communities, and economic ecosystems. That asymmetry is the technology ledger. The gains are civilizational; the losses are personal—or, maybe more precisely, *biographical*.

This is why I find myself, somewhat diffidently, offering a kind of preemptive defense—maybe even a justification—given the growing animus toward AI. Only a small minority of Americans describe themselves as optimistic about AI, say roughly 30 percent depending on the survey, while Chinese support remains vastly higher, by some measures above 70 percent.⁹⁶ One recent survey even found American sentiment toward AI *lower* than support for ICE—though still above Iran (the country, not the war), but either way you get the point.⁹⁷ Hard to imagine a more fragile base of enthusiasm. That helps explain the AI Illuminati's conspiratorial silence, and their vigorous denials, on job dislocation: why inflame an already skeptical public?

⁹⁶ Al Jazeera: "Trust in AI far higher in China than West, poll shows" published on November 19, 2025 ([Link](#))

⁹⁷ NBC News: "Poll: Majority of voters say risks of AI outweigh the benefits" published on March 10, 2026 ([Link](#))

Jensen Huang and David Sacks, in particular, have been vociferous on this point, arguing that AI is creating jobs and will relentlessly continue to do so. I'll come back later to what I think is an *Orwellian conspiracy of silence* on job loss, because it may be one of the more consequential *miscalculations* of the current moment. But before indicting the denial, I want to defend the thing being denied around: our deep dependency on, and intimacy with, technology. It's one of the defining traits of our species. We make tools. That is what technology is.

And again, I'm unequivocally grateful for it. If you zoom out from our present moment—past the moral panics, the daily algorithmic convulsions (synchronized in almost real-time by capital-markets convulsions), and the ambient Silicon Valley eschatology (ask me about P(doom) sometime)—the long arc is unmistakable. Technology has marked the upward trajectory of our species since the first faint flickers of civilization: writing, the wheel, the printing press, the steam engine, the electric grid, the microprocessor, the internet. Pick your favorite GPT (*general purpose technology*). With each turn of the crank, our material condition improved—sometimes by imperceptible micro-degrees, sometimes through genuine discontinuities.

I'll leave aside, for the moment, whether technology improved our spiritual or metaphysical condition. That's a longer and perhaps darker conversation. But in material terms, is there really much room for argument? For most of human history, life was, to use the old Hobbesian phrase, nasty, brutish, and short. I mentioned this earlier, but I'll repeat it in this context: at the beginning of the twentieth century, global life expectancy was only about 32 years; by 2023, it had risen to roughly 73. Over the longer arc, the share of humanity living in immiseration fell from something close to 80 percent in the early nineteenth century to under 10 percent by the late 2010s. That, by any disinterested accounting, is the forward march of civilization, and it didn't happen because we became morally better at scale. It happened because we became more capable. It happened because our tools got better.

The mechanism of that civilizational uplift wasn't simply moral exhortation, the old fantasy that if we just became kinder, wiser, and showed more solidarity with and compassion toward our fellow citizens—or whatever the preferred sermonizing adjective happens to be—the world would improve. No. It was the relentless accumulation of useful knowledge and the technologies that embodied it. As I argued at the opening of this essay, the Scientific Revolution and Enlightenment created the epistemic preconditions—skepticism, experimentation, iconoclasm, and the rejection of fossilized, revered authority—that made the Industrial Revolution possible.

What followed over the next two and a half centuries was the most compressed, telescoped burst of human progress in recorded history: advances in longevity, sanitation, literacy, democracy, mobility, safety, food production, energy, communication, and productive capacity on a scale that would have looked like sorcery to our ancestors. None of this means technology is always benevolent. I'll repeat Paul Virilio's aphorism: when you invent the ship, you also invent the shipwreck. The same tools that let you eradicate disease can let you engineer pathogens; the same networks that democratize knowledge can propagate lies at industrial scale; the same models that tutor the lonely child can addict the lonely adult or degrade the already-fragile mind. And the same machine intelligence that can reduce diagnostic error can also intensify surveillance, displace workers, and concentrate power in the hands of capital owners with the best chips, the cheapest energy, and the least sentimental attachment to the old labor compact.

But taken in total, the record remains overwhelmingly expansionary. Technology has been the central engine of human flourishing.

The Word That Did the Moral Work: Reabsorption

Still, that upward swing has never come without collateral damage. Every major technological leap brings negative externalities, often cruel, often concentrated, and often arriving faster than society can calmly or wisely absorb them. Technology saves in the aggregate and wounds in the particular. That's the asymmetry. Civilization gets richer; the worker, the guild, the town, or the entire local ecology built around an old form of labor may get ruined.

The old technological bargain rested on one morally load-bearing word: *reabsorption*. Yes, the machine arrives. Yes, it destroys an occupation, or at least a cluster of tasks inside an occupation. Yes, the incumbents rage against it, organize against it, and occasionally smash the thing with hammers, which is understandable even when it's unwise. But then, over time, the economy reorganizes. The bottleneck moves. Labor migrates. Demand expands. New firms appear. New tasks emerge. The old job disappears, shrinks, or loses its old dignity, but new work materializes around the next constraint. That's the reassuring story. And to be fair, it has usually been true.

The historical examples are familiar because the sequence has repeated with reassuring regularity. The loom obsoleted artisanal weaving but made cloth cheaper and more abundant. The tractor destroyed a world of agricultural labor but released workers, or more often their children, into factories, offices, services, cities, and all the strange occupational categories of the modern economy. The automobile annihilated the horse-powered economy—the animals themselves, yes, but also the blacksmiths, carriagemakers, harness shops, feed merchants, street cleaners (use your imagination here), stables, veterinary routines, and the whole equine urban metabolism—and built a new industrial civilization in its place.

The internet, now remembered as an inevitable triumph, though not without its own society-deranging side effects, destroyed clerical, retail, travel-agent, and intermediary roles, then created work categories nobody could have described in 1992 without sounding slightly unwell: cloud architecture, cybersecurity, search optimization, app-store economics, digital ad operations, creator-economy management, and the rest of the awkwardly named but economically real labor ecosystem that now surrounds the digital economy. The old order disappeared; a new order formed around the next set of constraints.

That sequence matters because it's the *historical comfort blanket* wrapped around every automation debate. The machine destroys the thing we can see and helps create the things we can't yet imagine. The pessimists see the disappeared task. The optimists see the reorganized economy. And over the long arc, the optimists have usually had the better of the argument, even if they have often been a little too bloodless about the bodies buried in the transition.

The social conflict was real, though, and we shouldn't sanitize it just because the macro story eventually looks good in retrospect. The Luddites didn't smash looms because they were irrational peasants opposed to progress; they smashed them because the machines were making their artisanal labor economically superfluous and transferring the surplus to owners who no longer needed their craft in the same way. The Pullman Strike of 1894 didn't erupt because workers had suddenly become ill-tempered. It erupted because George Pullman savagely cut wages during the depression of the 1890s while refusing to reduce rents in the company town, leaving workers trapped inside a vertically integrated form of economic indentured servitude: lower pay, fixed costs, employer-owned housing, employer-controlled civic life, and no plausible exit.⁹⁸

⁹⁸ Human Trafficking Institute: "On This Day in History: The Pullman Strike Ended" published on July 22, 2019 ([Link](#))

That wasn't merely a labor dispute. It was a collision between industrial concentration, wage compression, mechanization, and the intolerable paternalism of capital pretending to be civilization. So yes, the old bargain was regenerative. But it wasn't gentle. It certainly didn't distribute pain evenly. It didn't arrive with severance, retraining, a mental-health benefit, and a well-designed transition compact. It often ruined particular people and places before the next equilibrium appeared. Across the Gilded Age and after, technological progress repeatedly reordered labor markets violently and unevenly, producing unrest first and adaptation later.

This is why the *lump-of-labor fallacy* is usually, well, a fallacy. There isn't some fixed quantity of work sitting in some civilizational cupboard, waiting to be divided among anxious humans. Work is endogenous to technology, price, demand, income, imagination, and institutional formation. Productivity lowers costs; lower costs increase consumption; increased consumption expands output; expanded output creates new bottlenecks; and new bottlenecks call forth new labor. Labor isn't merely displaced. In the old story, it's eventually *pulled toward the new scarcity*. That's why serious people instinctively distrust technological-unemployment claims. They have *history* on their side, or at least most of history. Don't mistake task destruction for work destruction. Don't confuse the visible job loss with the invisible job creation. Don't extrapolate from the smashed loom, the shuttered stable, the vanished travel agency, or the automated teller window to the whole future of labor.⁹⁹

Usually, this is good advice.

The Hidden Premise: Human Cognition Stayed Scarce

But here is where the optimistic, or to use that favorite Dario word, *Panglossian* case starts to wobble. The old bargain had a hidden premise: *human cognition remained scarce*. That premise is so foundational that we almost never had to say it aloud. Machines could replace muscle, speed calculation, automate records, routinize clerical work, move atoms, process transactions, and make old forms of manual or administrative labor uneconomic. But the displaced worker, or at least the next generation of workers, could climb the abstraction ladder.

If the machine took the shovel, the human could sell supervision. If the machine took calculation, the human could sell interpretation. If the machine took filing, the human could sell coordination, judgment, communication, taste, trust, or management. That was the old ladder. The machine attacked the lower layer; labor moved to the higher layer. The whole optimistic story of *reabsorption* depended on that ladder remaining available.

This is why education became the canonical answer to technological disruption. Get more schooling. Get more training. Move up the skill curve. Learn to manipulate symbols rather than things. Become a knowledge worker. Leave the exposed world of *brawn* and *routine* behind, and enter the safer world of cognition, credentialing, analysis, professional judgment, and organizational interpretation. For two generations, that advice was spot on. The computer age amplified the value of the educated worker. The prestige professions flourished. The wage premium for cognitive labor became one of the defining facts of the late twentieth-century economy.

But notice *what had to remain true* for that strategy to work. The machine couldn't be allowed to climb the ladder too. It could automate below us, beside us, around us, but not fully into the domain we had designated as the next refuge for labor. Human judgment had to remain the scarce factor. Human synthesis had to remain the bottleneck. Human expertise had to remain expensive. Human

⁹⁹ David Autor: "Why Are There Still So Many Jobs? The History and Future of Workplace Automation" published in Summer 2015 ([Link](#))

cognition had to remain difficult to replicate.

That's precisely the premise GenAI puts under pressure. *GenAI threatens the ladder itself*. Not all of it, not all at once, and not without friction, resistance, regulation, failure, absurdity, hallucination, and the usual jagged frontier of competence. But enough of it to change the burden of proof. When a technology can synthesize, draft, reason, retrieve, code, summarize, classify, recommend, converse, and increasingly act, we're no longer asking whether one narrow task gets automated. We're asking whether the next supposed refuge for labor is itself *machine-addressable*. That's the discontinuity. The machine is less *taking the shovel* and more *learning the apprenticeship manual*.

This is why the old reassurance starts to sound less, well, reassuring. The standard answer to automation anxiety has always been: sure, this task goes away, but humans will move to the next scarcity. Fine. But what if the next scarcity is routine judgment, and the model can approximate it? What if the next scarcity is synthesis, and the model can synthesize? What if the next scarcity is document review, coding, claims follow-up, prior authorization, chart summarization, generic analysis, routine writing, or the application of established knowledge, and all of that is increasingly legible to machine intelligence? Then the hiding place isn't very good.

The more rigorous question isn't "Will AI destroy jobs?" That framing is too reductive and, oddly, too comforting. It invites a familiar argument, and familiar arguments are where people turn to avoid seeing new things. The better question is: where does the bottleneck move, and is the new bottleneck *defensible* against the next model release? If the answer is physical presence, trust, leadership, complex social navigation, moral responsibility, institutional legitimacy, or genuinely novel judgment under ambiguity, then labor still has somewhere to run—and, practically, we then have a compass for reskilling our teams. If the answer is a decontextualized cognitive workflow behind a screen, then the refuge is illusory.

This is also why the distinction between *partial automation* and *full, or near-full, automation* matters so much. Partial automation can *move the bottleneck* and create more demand for the remaining human function. Full automation changes the story. The ATM example is useful because it reveals the sequence. For a time, ATMs made branches cheaper and tellers more relationship-oriented, and teller employment actually increased. The machine automated cash dispensing, but the bottleneck moved toward service, sales, and branch expansion. Later, digitization, mobile banking, branch rationalization, and changing consumer behavior began to shrink the role precipitously. So the historical lesson isn't as clean and simple as "automation creates jobs" or "automation destroys jobs." The lesson is more subtle: *automation changes the bottleneck*, and the labor outcome depends on whether the new bottleneck still needs humans in large numbers. I know I keep repeating this word 'bottleneck' and (probably annoyingly) italicizing it. But we need this word to enter our vernacular in a real way, as it will prove the key to our recommended response. More on that later.

For now, in short, this is the analytic frame I believe we need for GenAI. The first phase may look augmentative, even benign. A worker with AI does more. A coder ships faster. A claims analyst resolves more queues. A nurse documents less. A physician's inbox gets triaged. This is the phase everyone likes because it lets us say the comforting word: *augmentation*. But the second phase asks a harsher question. If the tool can do 30 percent of the job, then 50 percent, then 70 percent, what exactly remains, and how many humans are needed to do it?

That work will exist. I don't believe in a simple no-work future, at least not on any timeline relevant to this chapter. Latent demand in many human domains is enormous. *If care gets cheaper, people will use more of it*. If expertise becomes more accessible, people will ask more questions, monitor more

conditions, pursue earlier interventions, and expect more continuous engagement. **Jevons' paradox** is real here. Lower the cost of expertise and demand will expand. But expansion of demand doesn't guarantee preservation of the old labor model. The new demand may call for different humans doing different work in different ratios, with far less tolerance for the administrative middle that grew up around the current system's dysfunction. New work will appear, but not necessarily in the same volume, at the same wage, in the same geography, or for the same people.

That's the difference between dislocation and reconstitution. Dislocation is what happens when the job goes away and another job eventually appears. Reconstitution is what happens when the underlying unit of economic value changes. GenAI is doing the latter. It's not merely reallocating workers among roles. It's repricing the *cognitive substrate* that made whole professions, credentials, and organizational layers economically defensible.

And that's why this chapter has to be more severe than the usual productivity sermon. Reabsorption worked because the machine left enough of the human hierarchy intact for displaced workers, or their children, to climb elsewhere. If the machine begins to learn the climb itself—the codified apprenticeship, the routine judgment, the professional pattern-recognition, the clerical coordination, the management narration—then the old escape ladder narrows. The old bargain was reabsorption. The new bargain has to be designed.

That's the uncomfortable difference. We can't simply trust that labor will wander, unaided and uninjured, toward the next bottleneck while capital captures the surplus and institutions congratulate themselves on productivity. If GenAI breaks the historical bargain, then serious institutions have to help write the next one. So yes, if AI merely automates tasks, history tells us not to panic. If AI erodes the cognitive scarcity that allowed labor to be reabsorbed, then history is no longer a sufficient guide. It becomes a warning, and perhaps a provocation.

The old technological bargain depended on reabsorption. The new labor strategy has to ask whether reabsorption is still plausible, where it's plausible, where it's not, and what we owe the people for whom the old ladder may no longer lead anywhere obvious. And if the old bargain doesn't hold, then the labor-intensive institutions of the service economy have to stop pretending this is just another productivity tool and start treating it as the most serious labor, capital, and moral reconstitution of our professional lifetimes.

Part II—Why AI Breaks the Bargain

Reader Note: Some of this section is a verbatim repeat from the 'Generative Epistemology' chapter. I'm re-including it here as some of you may be reading the chapters out of sequence, or (gasp) skipping some chapters entirely.

The Input That Produced All the Outputs

To see why the old bargain may break, you and I have to spend a moment thinking about intelligence itself. This isn't a decorative philosophical detour, or another opportunity for me to wander off into the metaphysics of software while the more practical reader quietly wonders when I'm going to return to workforce planning. *The intelligence question is the workforce question.* The prior section argued that reabsorption worked because labor could keep climbing: from muscle to supervision, from calculation to interpretation, from clerical work to knowledge work, from routine execution to judgment. That climb depended on cognition remaining scarce, biological, local, slow to train, and expensive to distribute. So before we can say anything useful about labor, substitution, apprenticeship, or the future of professional work, we have to ask what happens when the machine begins working on the very faculty that made the next rung available.

That's the difference this time. It isn't that the machine is faster, cheaper, more convenient, or better dressed up in a cuddly chat interface (hi! What should we talk about today?). The difference is that the machine works on intelligence itself. For two and a half centuries, labor survived mechanization by moving upward through the abstraction ladder. But what happens when the next rung on the ladder—cognition itself—becomes scalable, distributable, increasingly cheap, and no longer exclusively biological? That's the hinge. GenAI breaks the bargain not because it automates tasks. Every serious technology automates tasks. It breaks the bargain because it attacks the scarcity that made reabsorption possible.

The answer isn't that Silicon Valley is excited, or that the demos are mesmerizing, or that every board now has an AI slide, or that the same five people are on every podcast saying "agents" with sacramental intensity. The answer is more basic: this technology acts on the very input from which the other technologies came. There is something qualitatively different about this tool—a difference in kind, not merely degree.

What is so categorically different about GenAI, and why might it remake the labor bargain rather than merely continue it? The answer, I think, lies in the statement I've made ad nauseum in this piece: GenAI is fundamentally, irreducibly, a *multiplication of intelligence* itself.

Whether one prefers to stubbornly characterize these systems as stochastic token predictors or as early forms of machine reasoning is becoming a less interesting distinction when the outputs demonstrably extend beyond memorization into synthesis, planning, inference, autonomous action, and discovery-adjacent behavior. I understand the philosophical debate. I even enjoy it, in the way one enjoys a slightly overlong dinner party argument after the second glass of pinot. But for the purposes of labor markets, the metaphysical question matters less than the economic one. If the system can perform the work, synthesize the material, generate the code, answer the question, and recommend (and perform) the next action, the labor market won't wait for the philosophy department to adjudicate whether it's *really* thinking.

What matters in practice is that we've now *speciated*—or instantiated, if one wants the less dramatic verb—a non-carbon intelligence that in many domains already matches, and in certain narrow domains exceeds, the performance of our biological cognition. That fact alone should give us pause, because as *Homo sapiens*—literally "wise humans"—our entire civilizational identity rests upon *intelligence* and *sociability* as our defining traits. For millennia, the fundamental unit of intelligence in the economy was the human brain. Every technological breakthrough, every scientific discovery, every institution, every economic system that followed emerged from the *accumulation*

and *recombination* of those biological cognitive units.

The past 250 years of telescoped progress—the Industrial Revolutions of *mechanization*, *electrification*, and *computerization*—were themselves downstream of that combinatorial force. Human intelligence discovered physical laws, organized production, engineered machines, built institutions, and gradually constructed the technological infrastructure that lifted civilization out of its Malthusian trap. What's different now is not that we've simply invented another downstream tool. We've multiplied the upstream faculty that produced all the tools. If the steam engine amplified human muscle, and the microprocessor amplified calculation, GenAI amplifies the underlying cognitive faculty that created both. Intelligence is the ultimate general-purpose input. It governs scientific discovery, economic organization, institutional decision-making, technological design, professional judgment, and cultural production. When that input becomes scalable and multiplicative, the downstream consequences propagate through every system that depends on it.

And the modern service economy depends on it almost totally.

A Tool Whose Target Is Cognition

This is why the reassuring phrase “AI is just a tool” is technically correct and profoundly wrong. A steam engine was “just a tool” too, until it reorganized the relationship among muscle, energy, capital, and production. A spreadsheet was “just a tool,” until it changed finance, accounting, analysis, management, and the architecture of the modern firm. Tools aren't harmless because we call them tools. Some tools change the production function underneath civilization.

GenAI is a tool whose target input is cognition. That's the important distinction. A tool that cheapens cognition is more than an ordinary tool for a service economy *built* on expensive cognition. It signifies a fundamental regime change. That is why I use two different phrases in this essay. At the civilizational level, GenAI is a *multiplication of intelligence*. At the labor-market level, it's the *industrialization of intelligence*. The first phrase captures the grander metaphysical point: intelligence, the faculty that made us the dominant species on earth, is no longer exclusively biological. The second captures the economic point: *cognition is becoming a scalable industrial input*, manufactured through capital, compute, electricity, data, model architecture, and inference. Once cognition becomes industrialized, the professions and organizations built around scarce cognition begin to lose their old immunity.

The paradox is that these systems are simultaneously tools and something more than tools. They remain extensions of human agency—trained, directed, and deployed by people (for now, let's see how autonomy unfolds)—but they operate in the very domain that historically defined humanity's comparative advantage. And they don't emerge in quite the same way earlier tools did. As Jack Clark of Anthropic has observed, we've grown these systems more than built them. Their internal cognition, their *neurology*, operates in vast latent spaces we can't directly interrogate or fully interpret. Geoffrey Hinton's warning remains chilling: in evolutionary history, a less intelligent species has never reliably controlled a more intelligent one. We're now attempting precisely that inversion, and on timelines wildly compressed relative to the tempo by which intelligence hierarchies normally change.

So yes, they are tools. But they are tools working on the thing that made tools possible. That's why the labor implications can't be treated as a normal productivity bump. If intelligence is the thing we use to produce new tools, institutions, medicines, arguments, care models, companies, and operating systems, then multiplying intelligence means multiplying the power that reorganizes labor. It's not merely laborsaving. It's labor-redefining.

We Taught the Models Our Civilization

If GenAI is a *multiplication of intelligence*, and in the realm of work it's the *industrialization of intelligence*, then the obvious next question is: what exactly has been multiplied or industrialized? What sort of intelligence is this? What, in fact, have we made? Embarrassing to confess, perhaps, but I talk—literally, in voice mode—with the models sometimes for a couple of hours a day about everything: restaurant recommendations, obscure pockets of history, analogies I'm struggling to land, strategy questions, recombinations of ideas, the usual set of harmless and not-so-harmless intellectual co-dependencies one develops with an omniscient alien in one's pocket. Because that is increasingly what these models are: a thought partner par excellence; a synthetic, nonbiological intelligence that, in many quantifiable ways, is already superseding our own carbon intelligence.

Start with this: we've fed this silicon, non-carbon intelligence the corpus of digitized human knowledge—our histories, philosophies, theologies, sciences, arguments, poems, statutes, medical literature, source code, message boards, manuals, and the accumulated civilizational record. In other words, we've given it the artificial portmanteau of the two traits that gave us dominion over the earth: intelligence and sociability. Intelligence lets us reason; sociability lets us transmit those insights horizontally to our contemporaries and vertically to our descendants. Combined, that is collective intelligence. And we have, in a very real sense, taught these models our civilization.

And it turns out our civilization is learnable.

Perhaps we romanticized our *singularity* (can I still use that word given its new connotation?), our *ineffability*, as a species. Feed a system enough of our linguistic, conceptual, cultural, scientific, and behavioral residue, and it begins to produce something that looks uncannily like a synthetic version of our collective mind. Not human, exactly. Not wise, necessarily. Not morally grounded. But recognizably capable of operating across the accumulated patterns of human civilization. That's why I find myself increasingly uninterested in and impatient with the more *scholastic* versions of the debate: are they "really" thinking, are they "really" reasoning, are they merely stochastic parrots, are they "no more intelligent than a cat"—thank you, Yann LeCun, for that colorful line, though fewer and fewer people are listening to you in the face of the undeniable model capability leaps. Those questions are starting to feel beside the point. The more important question is what capabilities *emerge* when computation is fused with the accumulated cognitive inheritance of humanity. And for labor, the answer is unsettling.

In the Generative Epistemology chapter, I describe four properties of this emergent intelligence: *functional omniscience*, *polymathy*, *omni-disciplinarity*, and *analogic recombination*. I won't reprint the whole meditation here—mercifully, for both of us—but the labor translation matters enormously. For the purposes of this chapter, I'd compress those properties into three economic capabilities.

First, these models have something approaching total recall of the digitized canon. Not perfect recall, not error-free recall, not wise recall, but functionally superhuman access to a breadth of information no individual professional can possibly hold. A lawyer knows a domain of law; the model can traverse many if not all. A physician knows a specialty; the model can traverse the literature and epistemic base across most if not all specialties. A consultant has pattern recognition from a finite set of client engagements; the model has pattern exposure across vast encoded corpora of business, markets, operations, and human behavior. That alone begins to reprice expertise, and not upward.

Second, they are polymathic and omni-disciplinary in ways that embarrass our human institutional categories. The professional world is organized around boundaries: medicine, law, finance, software,

operations, behavioral health, pharmacy, actuarial science, clinical documentation, coding, compliance, strategy. The model doesn't respect and abide by these boundaries with the same reverence. It can move with near-omniscience from one domain to another, sometimes clumsily, sometimes brilliantly, but with a kind of cross-domain fluency that used to be the mark of the rare human polymath. In an economy where expertise has been subdivided into credentialed silos, this matters. The machine isn't merely a specialist. It is, increasingly, a synthetic ensemble of specialists. Third, and most consequentially, it's analogic and recombinant. This may be the most important property of all. Yes, I'm going to use the word "thinking," though the pedants will object. I increasingly believe that many of the greatest ideas in history came from analogic thinking. We like to deify the human ratiocinative process, as if brilliance drops fully formed from the heavens into Newton's or Darwin's or Demis' mind. But epistemologically that isn't how it works. Our finest minds saw interstitially. They found the bridge between domains. Gutenberg's insight was the wine press becoming the printing press. Darwin's theory of natural selection was catalyzed, in part, by reading Malthus and importing the logic of population pressure into biology. The most original minds didn't simply know more. They recombined more daringly.

This is exactly what the models intrinsically do. They are creative in the deep sense, not the superficial "new cat image" sense, but in the application of insights from one domain to another, in the analogizing, the recombination, the generation of new propulsion to human and machine progress. And then, unlike us, they can distribute those insights instantaneously across their instantiations. A human genius dies with unscalable tacit knowledge trapped in her skull. A model improvement can, at least in principle, propagate across the fleet. Just think about this when we take LLM intelligence and 'embody' it into a fleet of Optimus robots (or 'Optimi,' as Elon has taken to labeling them)—I'll talk about the 'hive mind' dynamics later.

But let me land this plane. That last point—interstitial and analogic thinking—matters more than people realize. Professional scarcity has historically depended not just on intelligence, but on the *localization* of intelligence. The expert was scarce because the expert was an individual human: trained slowly, apprenticed locally, credentialed institutionally, and limited by memory, time, fatigue, geography, and embodiment. The model isn't scarce in that way. It can be copied, routed, updated, fine-tuned, monitored, and distributed. That changes the economics of expertise.

The Repricing of Professional Scarcity

Those properties matter for labor because they are precisely the properties on which professional scarcity has depended. The professional class sold memory, synthesis, cross-domain fluency, judgment-by-analogy, and the confidence that came with having absorbed a difficult canon. A system that is functionally omniscient, polymathic, omni-disciplinary, and analogic doesn't merely help the professional class. It reprices the raw materials out of which the professional class was built.

Again, this doesn't mean wisdom is free. Wisdom isn't the same as recall, and judgment isn't the same as synthesis. Trust isn't the same as accuracy. Responsibility isn't the same as prediction. A model can produce a correct recommendation or perform a correct action without bearing the moral burden of the patient, the employee, the community, or the institution. That distinction remains vital. But large parts of what passed for professional scarcity were never *wisdom* in the deepest sense. They were the friction costs of accessing, organizing, synthesizing, and applying knowledge. Those costs are falling toward the marginal cost of compute. The defensible human moat consequently has to move elsewhere: trust, taste, courage, responsibility, embodiment, relationship, and moral judgment.

This is a hard thing for the prestige professions to absorb because it sounds like an insult, though it certainly isn't meant as one. It's an economic observation. If part of your value came from knowing what others couldn't easily know, and if a machine now knows or can retrieve much of that on demand, then the market will eventually discover that your scarcity was partly contingent. Not fake or useless. But contingent.

Now let's connect this back to the historical bargain without painfully restating the whole previous section. Reabsorption worked because the machine usually attacked a lower layer of work while leaving a higher layer of human scarcity intact. The machine could take muscle, calculation, filing, or retrieval, while humans moved toward supervision, analysis, interpretation, coordination, management, and judgment. The worker climbed, or the next *generation* climbed, into the next scarcity.

GenAI threatens the climb. By no means does it eliminate every higher layer of human value, and we shouldn't indulge in silly totalizing claims. Physical presence still matters. Trust still matters. Leadership, persuasion, taste, ethical judgment, institutional legitimacy, and embodied care still matter enormously. But the old upward refuge—generic cognitive work, routine professional synthesis, credentialed application of established knowledge—is no longer reliably safe. *The machine has learned to climb part of the ladder.*

That's the discontinuity I'm trying to communicate. And it's why the labor disruption will be so psychologically destabilizing. The industrial revolutions of the past attacked the hand and the back before they attacked the office. This one attacks the office first. It comes for the people whose identities were built around intelligence, education, expertise, and the right to interpret complex things for other people. It comes for the analyst, the junior lawyer, the consultant, the coder, the administrator, the medical documentation worker, the revenue-cycle specialist, and eventually parts of the physician's cognitive workflow. Not because those people are unimportant, but because their work is information-rich, language-mediated, digitally represented, and increasingly machine-addressable.

The old bargain said: don't worry too much, labor will move to the next bottleneck. The new question is whether the next bottleneck is actually defensible against the next model release. If the bottleneck is trust, touch, leadership, physical presence, or moral responsibility, the human still has ground. If the bottleneck is the routine application of established knowledge behind a screen, the ground is already eroding.

From Intelligence Expansion to Intelligence Replacement?

This is why the language of *augmentation*, while comforting, is incomplete. Of course GenAI augments. It helps the worker do more. It gives the novice competence, the expert leverage, the overburdened clinician some relief, the analyst a better first draft, the coder a faster loop, the researcher a tireless assistant. The augmentative phase is real, and in many ways liberating and even beautiful.

But augmentation is often the first chapter of substitution. A tool that makes one person twice as productive may initially look like empowerment. At enterprise scale, it also asks whether two people are still needed. A tool that lets a junior perform like a mid-level may initially democratize opportunity. It also compresses the wage premium attached to the mid-level. A tool that drafts the memo, summarizes the chart, codes the claim, and triages the inbox may feel like relief until the organization starts redesigning the workflow around the assumption that the tool, not the human, is the productive unit.

That's the disquieting labor translation of the *multiplication of intelligence*. The same technology that

makes individuals more capable makes institutions less dependent on individuals. The same model that democratizes expertise also demonetizes parts of expertise. The same agent that liberates a worker from drudgery may eventually eliminate the role that contained the drudgery. That doesn't make the technology evil. It makes it powerful.

And power, as this essay keeps insisting, requires moral seriousness.

Part III—The Four Great Decouplings

Why This Macroeconomic Detour Belongs Here

Ok I'll stop philosophizing and get to the real, *felt* economic and sociological implications. I'll confess, though, to a little diffidence in presenting the following section, because you've certainly waited long enough for me to get to the practical heart of this whole exercise: *which* jobs are exposed, *what* leaders should do, *how fast* this moves, and *who*, exactly, still needs to be employed in 2027. Fair enough. But if the previous section was the moral ledger of technology—the old bargain of destruction now and regeneration later—then this section is the macroeconomic ledger. And unfortunately, we need both.

So I apologize in advance, but you and I have to widen the aperture before we drop back down into workflows, billing offices, nurse staffing, utilization management, and the scary-but-real managerial question of how much of the current labor architecture survives first contact with agentic AI. The labor question isn't merely a microeconomic question about who writes the note, reconciles the medication list, reviews the denial, or handles the work queue. It's also a growth-theory question, a demographics question, a political-economy question, and, finally, a capital-allocation question. Put differently, if GenAI is truly the *industrialization of intelligence*, then it may begin to uncouple relationships that have been conjoined for, well, centuries.

That's why I want to share with you what I've taken to calling the **four great decouplings**. Not because I'm trying to drag you through yet another macroeconomic excursus for the sheer sadism of it. But because the labor-displacement question is unintelligible *unless we see the larger pattern*. AI doesn't just threaten to decouple task from worker. It may begin to decouple *population from progress*, *GDP from employment*, *productivity from compensation*, and *capital formation from human hiring*. Once you see that, the labor shock stops looking like a sectoral HR problem and starts looking like a change in the very production function of the service economy.

This is the point of the detour. If intelligence becomes scalable, if cognition can be manufactured, if capital can instantiate digital labor, and if output can rise without proportional hiring, then the old labor bargain is no longer merely under pressure at the task level. It's under pressure at the macro level. The question isn't simply whether a particular role survives. The question is whether the old relationships among population, productivity, employment, compensation, and capital still hold. And I increasingly think they don't.

The First Decoupling: Progress without Population

For most of human history, progress was population-bound. More humans meant more hands, more minds, more experiments, more recombinations of ideas, more tinkerers, more engineers, more playwrights, more physicians, more entrepreneurs, more insurgents in garages, and more fools accidentally bumping into something important. The story of civilization is, in one sense, the story of *biological population becoming epistemic population*: more people creating more knowledge, more knowledge producing more output, more output allowing more people to survive, and then those people creating still more knowledge. A little reductive, sure, but not wrong.

How's this for a geologic age summarized into a paragraph? For 100,000 years we were *hunter-gatherers*. Then, for roughly 10,000 years, *agriculturalists*. Then the three Industrial Revolutions took the long, flat misery of the Malthusian world—population growth swallowing returns, more mouths eating the gains, human life pinned to scarcity—and shoved it onto a different trajectory. Until around 1700, “growth” in the modern sense was almost a category error. The world economy moved at something like continental-drift speed, roughly 0.1 percent per year, doubling not over a business cycle or a generation but over something like a millennium. Productivity, as a modern measurable phenomenon, barely existed. Then useful knowledge, energy, mechanization, markets, institutions, and the epistemic inheritance of the Scientific Revolution began compounding. Output per person rose. Longevity rose. Literacy rose. Sanitation improved. Democracy expanded. Violence declined. The modern world—flawed, noisy, unequal, decadent, magnificent—emerged.

Demography was the main protagonist in this play. Endogenous growth theory, semi-endogenous growth theory—pick your preferred pretentious academic flavoring—keeps circling the same intuition: ideas matter, and humans make ideas. More humans, more ideas. More ideas, more growth. And because ideas compound, because one discovery becomes the platform for the next discovery, progress becomes multiplicative. Gutenberg makes Newton more likely. Newton makes Faraday and Maxwell more likely. Maxwell makes electrification more likely. Electrification makes computation more likely. Computation makes AI more likely. *Learning begets learning.* Civilization is cumulative cognition embodied in tools, institutions, and transmitted knowledge.

But that relationship may now be breaking. This is the first great decoupling of the AI era: *the decoupling of population growth from idea production*. We may be moving, to borrow Kevin Kelly's wonderfully clarifying frame, from an *economy of the born to an economy of the made*. Synthetic minds, digital workers, embodied robots, AI research agents, code agents, biological-discovery agents, clinical agents—all of them represent the possibility that civilization can produce cognitive labor without producing more babies. The marginal idea generator no longer has to be biologically born, raised at great expense, educated for twenty-five years, credentialed, institutionally socialized, and then given a cubicle, laptop, badge, manager, benefits package, and annual performance review. It can be instantiated.

This arrives at a strange and maybe providential moment, because the human demographic engine is faltering. Actually, and Elon has this right, it's pretty much *broken*. Fertility is falling almost everywhere outside parts of Africa. Major economies are aging. Pronatalist policies are failing with impressive consistency. South Korea, Japan, China, Italy, Germany, Spain, much of Europe, Canada, the United States—the pattern differs in degree, but not in kind. Rich societies become educated, urban, secular, expensive, anxious, atomized, child-light, and then *old*. They stop marrying and they stop having sex (just look at the data). The dependency ratio turns ugly. The state steps in because the family structure has thinned out and the elderly need scaffolding. And then everyone discovers,

somewhat belatedly, that a geriatric society requires a great deal of labor—and a lot of, you guessed it, *healthcare*.

Without AI, this is a civilizational problem of the first order. Fewer workers, fewer caregivers, fewer builders, fewer taxpayers, fewer young people willing or able to subsidize a vast aging superstructure. A society of too few children and too many elders becomes a society of scaffolding, caregiving, rationing, and quiet but inexorable decline. And yes, we should pursue the obvious responses: immigration where politically possible, later retirement where socially tolerable, better childcare, better housing policy, better family policy. *All of that*. But let's not be naïve. Fertility decline isn't a simple policy variable. It's a civilizational symptom of wealth, urbanization, education, cost, atomization, and mimesis. People copy the fertility decisions of the people around them, and once the prestige equilibrium shifts toward fewer or no children, reversing it's hard.

AI changes the (re)production math. It doesn't solve the spiritual problem of demographic senescence; the melancholy of a society with too few children isn't captured by the crude, reductive measurement of GDP, and we should resist the temptation to treat every human question as an output-per-hour question. But AI may solve, or at least substantially mitigate, the production problem. We can instantiate artificial minds. We can build robotic caregivers. We can substitute compute for cognitive labor and, eventually, embodied machines for some forms of physical labor. *The connection between population size and civilizational ideation weakens*. We may no longer need multitudes in quite the same way to generate good ideas, staff organizations, absorb administrative work, or preserve the productive capacity of an aging society.

That may sound anti-human, and I certainly don't mean it that way. I'm not rooting against babies! God forbid. I'm saying the old equation—more people, more workers, more thinkers, more growth—may no longer be an immutable law of civilization. If a billion artificial minds can be spun up at the cost of electricity, chips, cooling, and capital, then the economic meaning of population changes. Not the moral meaning. Not the familial meaning. Not the spiritual meaning. But the *productive meaning*. And that distinction matters enormously.

This isn't some kind of indulgent, robotics futurism. It's more *growth theory with a GPU cluster*. If intelligence is the prime mover of progress, and if intelligence becomes scalable, then productivity can detach from population growth. That's why the phrase *multiplication and industrialization of intelligence* matters so much. It isn't metaphorical. It's a description of a new production function: capital becomes compute; compute becomes cognition; cognition becomes output; output compounds without requiring a commensurate increase in biological population. That's the first decoupling in its most disquieting form. The old human bargain was demographic. AI introduces the possibility of manufactured cognition. Bottom line for me—bring on population decline—AI rescues us from that.

The Second Decoupling: Growth without Hiring

The second great decoupling is equally destabilizing: the decoupling of *GDP from employment*. For most of modern macroeconomic life, output and employment moved together closely enough that policymakers, journalists, and ordinary citizens treated jobs as the obvious *evidence* of growth. If the economy was expanding, people were being hired. If people were being hired, wages rose. If wages rose, households felt better. The whole thing was messy, cyclical, uneven, and often cruelly unfair at the bottom, but the intuitive link was there: growth meant work.

But what happens when output rises and hours worked barely move? What happens when a

company triples revenue and increases staff by only a quarter? What happens when market capitalization doubles while headcount stays flat? What happens when the economy keeps growing while job seekers discover that the entry-level ladder has mysteriously lost its bottom rungs? What happens when, to borrow Satya's breezy and dissonant formulation, everything is 'up and to the right except staffing'?

We may be at the beginning of that divergence. I want to be careful here, because macroeconomics isn't exactly an immaculate science, and this particular moment is noisy almost beyond parody. Some of what we're seeing is almost certainly AI capex circularity: chips, data centers, power, construction, cloud infrastructure, cooling systems, and GPUs all showing up as growth before the downstream productivity gains have fully materialized. Some of it is pandemic over hiring unwinding. Some of it is rates, tariffs, policy, and the ordinary stupidity of economic data, which often appears most confident right before it gets revised into embarrassment. But even after entering those caveats, the possibility is now live: GDP can increasingly pull away from employment. Output can rise without commensurate hiring. Labor productivity can improve while labor demand softens. The economy can look strong from 30,000 feet and feel miserable at the kitchen table.

That last point matters. One of the strange emotional facts of the current economy is that capital markets can rip while ordinary people feel despondent. The economists and market strategists look at GDP, earnings, capex, productivity, and market capitalization and see strength. The citizen looks at wages, job availability, health insurance premiums, rent, groceries, student debt, medical debt, and the disappearance of the entry-level job and feels something closer to dread. Both can be seeing part of the truth. A technology that substitutes capital for labor should, if the theory is right, first make capital look triumphant and labor feel precarious. This is the political economy of substitution beginning to disclose itself.

The early signal is appearing where one would expect it to appear: in the sectors most exposed to AI and most capable of using it. Professional services, information services, software, finance, consulting, analytics, legal-adjacent work—the domains where cognition is digitized, outputs are *functionally verifiable* (yup, I managed to weave that phrase into yet another chapter), workflows are already computational, and the work can be decomposed, decontextualized, and outsourced. These are the places where AI should show up first, not as some magical general-equilibrium transformation, but as task compression, hiring restraint, productivity improvement, junior-labor evaporation, and the strange new corporate vanity: *we're growing without adding people*.

The official macro data are beginning to rhyme with the anecdotal economy. In Q1 2026, real GDP grew at a 1.6 percent annualized rate, with investment, exports, consumer spending, and government spending contributing, while imports—which subtract from GDP—also increased.¹⁰⁰ The BEA's technical notes are especially useful because they show the composition beneath the headline: investment increased, information-processing equipment was a notable contributor, software contributed through intellectual property products, and both exports and imports were led in part by computers, peripherals, and parts. In other words, the growth number already has a distinctly AI-adjacent material substrate under it: equipment, software, computers, imports, and the physical machinery of *synthetic cognition*. This is exactly what one would expect early in a capital-to-compute-to-labor transition.

That doesn't prove the productivity revolution has arrived in full. It *does* mean the old GDP-employment link deserves much less complacency than it's getting. An economy can be "strong" in the aggregate while becoming more hostile to hiring at the margin. A firm can grow revenue without

¹⁰⁰ Bureau of Economic Analysis: "GDP (Second Estimate) and Corporate Profits, 1st Quarter 2026" published on May 28, 2026 ([Link](#))

building a commensurate human pyramid underneath it. A labor market can look numerically intact while its entry-level and white-collar developmental ladders quietly disintegrate. The decoupling doesn't have to arrive as mass unemployment on day one. It can arrive first as a low-hire economy, as *job hugging* instead of *job hopping*, as junior roles quietly vanishing, as contractors not renewed, as open recs closed, as the headcount plan getting smaller even while the revenue plan gets larger.

That's growth without hiring. And if it persists, it will be one of the defining political-economic facts of the AI era. And, I expect, the *number one determinative issue that will show up in the 2028 presidential election*.

The Third Decoupling: Productivity without Compensation

This is why I'm increasingly drawn to the phrase the *Great Divergence*, even though Larry Summers—whose own recent misadventures we can acknowledge without discarding every useful phrase he ever uttered—used related language in an earlier context. Productivity can rise while compensation doesn't keep pace.¹⁰¹ GDP can rise while employment lags. Corporate profits can rise while labor share falls. Market capitalization can rise while headcount contracts. This isn't entirely new; we've been living with some version of productivity-compensation divergence since the 1970s. But AI threatens to make the divergence discontinuous rather than gradual.

The old postwar settlement lulled us into thinking that productivity gains and labor gains would travel together, if not perfectly, then at least plausibly. The factory became more productive, and workers shared enough of the surplus to support the broad social fiction that growth was a national project rather than an asset-owner project. That settlement was never as clean as the nostalgia suggests, and it was built on exclusions we shouldn't romanticize. But it did create a psychic and political compact: productivity meant rising living standards for more than just the owners.

That compact has already been weakening. AI may break it much more dramatically, because the productive input it cheapens isn't merely a task or a tool but cognition itself. If a model can do the work that once justified the wage premium of a credentialed cognitive worker, the productivity gain doesn't automatically accrue to the worker. It accrues first to whoever owns the workflow, the customer, the model access, the data, the distribution, or the *capital* required to instantiate the synthetic labor. The worker may become more productive in the first act and less necessary in the second.

That's the psychological contradiction of the current moment. People can use AI and feel powerful individually, while the institutions around them use AI to need fewer people collectively. The same tool that makes the worker faster can make the workforce smaller. The same system that democratizes expertise can demonetize parts of the expert. The same productivity boom that lifts GDP can erode labor's claim on the surplus.

This is why the public mood can feel so bizarrely misaligned with the market mood. Investors look at operating leverage, hyperscaler capex, earnings concentration, margin expansion, and AI-linked market capitalization and see the future. Workers look at job postings, wage pressure, disappearing entry-level roles, layoffs framed as efficiency, health premiums, rent, groceries, medical debt, and the decline of the old job-switching wage bump, and they feel dread. Both perceptions can be accurate. AI can be genuinely good for growth and genuinely bad for labor's share of the winnings.

¹⁰¹ Economic Policy Institute: "The Productivity-Pay Gap" updated on March 23, 2026 ([Link](#))

The Fourth Decoupling: Capital Becomes Labor

And that brings us to capital. For decades after World War II, economists lulled themselves into believing that the division between labor and capital was almost a natural law: **roughly two-thirds of the winnings went to labor, one-third to capital**. But that was never a natural law. It was a historical equilibrium, born of institutions, bargaining power, unions, industrial structure, labor scarcity, postwar politics, global circumstances, and the particular technologies of the age. The equilibrium has already been eroding. AI may invert it.

I know that sounds dramatic. But walk through the mechanism: if capital buys compute, compute buys digital labor, and digital labor substitutes for human labor, then the center of gravity shifts from labor to capital with unusual violence. The owner of the machine wins. The owner of the distribution wins. The owner of the data wins. The owner of the electrons wins. The owner of the workflow into which the agent is inserted wins. The person selling undifferentiated cognitive time loses, because the marginal cost of equivalent compute keeps collapsing.

This is the movement from SWB to kWh: from salaries, wages, and benefits as the irreducible input of service production, to electricity, chips, inference, cooling, and agent orchestration as the new labor substrate. Leaders in labor-intensive industries still think of labor as headcount, FTEs, contract labor, agency spend, staffing ratios, overtime, retention. The AI-native company increasingly thinks in *kilo-watt hours*, GPU utilization, inference cost, model routing, task completion, agent reliability, and the marginal cost of synthetic labor. These are different civilizations of production.

This is why it will be very good to own capital when the sonic boom happens. Not because capital is morally superior. It's not. Not because capital owners are wiser, nobler, or more deserving. Often quite the opposite, if we're being honest. But because capital becomes the bottleneck to instantiating labor. The old world turned labor into output. The new world turns capital into compute, compute into cognition, cognition into output.

The social consequence is obvious and alarming: without redistribution, ownership concentration becomes destiny. And the labor-share data are starting to give that fear some empirical teeth. In Q1 2026, BLS reported that nonfarm business hourly compensation increased 3.1 percent, but *real* hourly compensation *decreased* 0.5 percent, and **labor's share of output fell to 54.1 percent**—the lowest recorded value since the series began in 1947.¹⁰² That number is the nub. It's capital winning and labor losing, not in some moralized, sloganized, campus-protest sense, but *mechanistically*. This is what systematic substitution of capital for labor begins to look like in the data.

One should be careful about over-reading a single quarter. I'm not trying to build a theology on one data point. But the direction is consonant with everything else: AI capex up, market concentration up, semiconductor valuations up, labor share under pressure, real compensation pressured, entry-level white-collar anxiety rising. The data have a personality now, and the personality isn't labor-friendly.

¹⁰² U.S. Bureau of Labor Statistics: "Productivity and Costs, First Quarter 2026, Preliminary" published on May 7, 2026 (Link)

The Unfair Interregnum

The transitional period will be especially unfair. Before full automation, the skills AI can't yet perform become bottlenecks, and bottlenecks can become extravagantly compensated. We'll see superstars: people who can manage hybrid human-agent systems, govern model risk, sell trust, design workflows, interpret ambiguity, and move institutions faster than their competitors without blowing up the enterprise. The top of some professions may become *vastly* more valuable for a while. The very best surgeon, the very best strategist, the very best AI-native operator, the very best healthcare CEO who can actually install this stuff rather than admire it from a safe distance—those people will be amplified.

But that doesn't rescue labor as a class. A few humans may get paid more because they sit at bottlenecks; many humans may get paid less because the machine has colonized the old middle. This is the polarization of the middle I keep circling. The top becomes cybernetic. The bottom gets temporarily lifted by tools that make the novice competent. But the middle—the credentialed, solid, respectable, institutionally processed professional whose value lay in applying established knowledge with decent reliability—gets squeezed. And once agentic systems mature, even the bottom-rung uplift may prove transitional, because the entry-level tasks themselves begin to disappear.

That's why *apprenticeship professions* should be nervous. This isn't just about job loss; it's about developmental-ladder loss. Law, consulting, medicine, finance, software, and healthcare administration were built around pyramids. Juniors did repetitive work, learned the grammar of the profession through volume, absorbed pattern recognition, made supervised mistakes, and gradually earned judgment. But if the machine erodes the *base* of the pyramid—drafting the memo, summarizing the chart, writing the first pass code, reconciling the claim—then where exactly does the senior professional come from? You can't produce senior judgment if no one gets to live through junior repetition. The apprenticeship professions don't merely shrink; they cavitate.

This, too, is a decoupling: the *decoupling of expertise from its old developmental pathway*. Expertise used to require time, apprenticeship, repetition, credentialing, mentorship, and institutional absorption. Now some share of expertise is being packaged, simulated, routed, and delivered at the marginal cost of compute. That's democratizing, sure. It's also professionally destabilizing. The credentialing economy was built on the premise that knowledge and judgment were scarce, slow to acquire, and socially certified through institutions. What happens to that economy when the knowledge is abundant, the synthesis is cheap, and the certification increasingly looks like a proxy for a scarcity that no longer quite exists?

That's the uncomfortable macro premise underneath the labor argument. We're not just automating a note, a claim, a call center, a denial letter, a utilization-management review, a coding workflow, or a care coordination task. We're watching the production function change. We're watching the old relationships among population, productivity, employment, and capital begin to loosen. And once those relationships loosen, the labor-intensive institutions that still assume the old bargain will hold are the ones most likely to be surprised by history.

Vinod Versus Daron, or Can We Take His Nobel Prize Back?

The previous sections argued that AI is beginning to uncouple relationships we've treated, too complacently, as natural: population from progress, GDP from employment, productivity from compensation, capital formation from human hiring. That was the macro frame. Now let's make it a little more concrete, and, because I apparently can't resist needling the East Coast professoriate, a little more personal.

The question, a tiny bit cartoonish but (I think) usefully framed, is *Vinod versus Daron*. Silicon Valley capitalists versus MIT socialists. Belief in scaling laws versus empirical extrapolation from the last paradigm. Religious conviction in the OOMs (Leopold Aschenbrenner's *orders of magnitude*) versus the sober, respectable, tenure-protected instinct to take yesterday's measured slope and extend it forward with just enough humility to be wrong slowly.

Vinod Khosla, who has been more prescient, and right, than just about any other technologist, talks about massive job dislocation and automation over the next 25 years. Have a listen to my recent podcast with Vinod for a masterclass in his thinking. Daron Acemoglu of MIT—and a recently honored Nobel laureate in economics—predicts a far more modest effect. His rationale is more sophisticated than the caricature, and we should be fair to it. While a meaningful share of tasks may be *technically* exposed to GenAI, only a much smaller subset will be economically attractive to automate over the next decade. His published work estimates that only about 5 percent of US labor-market tasks will be profitably performed by AI over the next decade, producing something closer to a modest GDP increase than a revolutionary productivity shock. MIT Sloan summarizes the view as roughly a 0.7 percent productivity increase over ten years and a more realistic GDP effect around 1.1 percent.¹⁰³ That isn't nothing. But it's very much the "steady economist with a pencil" version of the AI future.

Of course, there's a multiplicity of variables here—technology, regulation, sociology, cost curves, adoption curves, organizational change, all the usual confounders—but at the core of this divergence lies belief, or lack thereof, in the *scaling hypothesis*. Vinod, along with Sam, Mira, Elon, Satya, Sundar, Jensen, Dario, Demis, and Zuck, all have something close to religious conviction that scaling will continue, and they and their allies are backing the hypothesis with serious capital. Meanwhile, East Coast professors—and they do seem to congregate on the East Coast—are deprecating this strategy, principally by extrapolating past historical trends and empirical data forward. The problem is that this phase shift doesn't resemble past paradigms.

The question comes down to this: do you believe the scaling hypothesis or not? And to answer that, we need to follow Leopold's OOMs. By the way, Daron Acemoglu is a brilliant writer and magisterial thinker; his book *Power and Progress* is a must-read. It's just that many of his conclusions are socialist and wrong. But I digress.

Let me try to be adult about this for at least one paragraph. AI is distorting our economic data, so every strong statement here should carry a little, or maybe a lot of, epistemic humility. Some of the GDP contribution is circular capex. Some of it is imports. Some of it is wealth-effect consumption from AI stocks. Some of it is front-loaded investment that may or may not earn an adequate return. Some of it is statistical accounting getting mugged by a GPU cluster. But that, in a way, is the point. A technology powerful enough to distort GDP, trade, equity concentration, capex composition, and labor share before it has fully diffused into ordinary enterprise workflows isn't obviously a modest,

¹⁰³ Daron Acemoglu: "The Simple Macroeconomics of AI" published in May 2024 ([Link](#))

5-percent-of-tasks-over-a-decade event. Not dispositive proof of Vinod's upper-bound view, but already a problem for the reassuring view that this will remain bounded, incremental, and politely absorbed by the old labor market.

But bottom line I think Daron's prognostications are going to age very poorly, and Vinod will have yet another big vindication for his clairvoyance. Score one for Vinod, and deduct one (huge) point from Daron.

Capital Votes First

Let's start with capital, because capital is where the future often tells on itself first. Morgan Stanley reportedly has the five large hyperscalers—Amazon, Alphabet, Meta, Microsoft, and Oracle—spending roughly \$805 billion on capex in 2026 and something like \$1.1 trillion in 2027. That isn't venture enthusiasm, or another hyperventilating SaaS cycle. That's *sovereign-scale industrial mobilization* by private companies. We're not talking about a software upgrade. We're talking about capital converting itself into cognition and spending more than 3.3% of GDP on the transaction in the process. Or, to use the vocabulary from the previous section, we're watching the production function mutate in public: capital to compute, compute to cognition, cognition to output.

This matters because capex isn't neutral. A trillion dollars of AI infrastructure isn't being built because a few centibillionaires need a new theological hobby. It's being built because the people closest to the scaling curve believe that intelligence demand is inelastic, that the models will keep improving, that downstream use cases will materialize, and that the \$60 trillion global service economy is susceptible to reconstitution around synthetic labor.¹⁰⁴ Yes, there will be malinvestment. Yes, there will be capital incineration. Yes, bubbles can overbuild the wrong things in the wrong places at the wrong prices. But even bubbles can be civilizationally productive when they build the substrate on which the next economy runs. Railroads burned investors and still reorganized the continent. Fiber incinerated capital and still gave us the internet. The AI capex buildout will do both: waste money at the margin and build the substrate for a post-labor service economy at the core.

The Wall Street Journal put the more unsettling decomposition most cleanly: AI is now distorting "practically everything" about the economy—stock markets, profits, the speed and composition of growth, trade, and even the mood around the job market. This is a two-speed economy: one part strapped to a rocket, the other to a lawn chair. David Sacks, never one to understate the case, has argued that AI capex could add roughly 2 to 2.5 percentage points to GDP growth this year and more than 3 points next year. Cut that down that as aggressively as your priors require. Discount for self-interest, for political messaging, for the temptation of every administration to claim credit for whatever growth happens on its watch, or for a VC of the first rank talking his own book. Fine. The shape of the argument still matters. AI is no longer merely a future productivity story. It's already a macro demand story, a capital-spending story, an import story, an earnings story, and soon enough—perhaps already—a labor-substitution story. The input is capital. The intermediate product is compute. The economic bet is labor.

It's distorting trade too, which should amuse us a little given the mercantilist, 17th-century tariff stupidity of the moment. Semiconductors, servers, memory—the machinery of *synthetic cognition* still has to be manufactured, shipped, installed, powered, and cooled. The world may be digitizing, but the digital still runs through the atoms of ports, fabs, wafers, power systems, and geopolitical chokepoints. So much for the notion that one can simply tariff the trade deficit into submission. Our hunger for intelligence keeps us ordering from Taipei, Seoul, and wherever else the physical substrate

¹⁰⁴ World Bank Group: "Services, value added (% of GDP)" accessed on June 9, 2026 ([Link](#))

of compute can be made quickly enough. This is one of the more amusing ironies of the current moment: the AI economy is “digital,” but its first macroeconomic footprint is almost embarrassingly material. Chips. Power. Land. Cooling. Copper. Transformers. Data centers. We’re building a synthetic mind with very physical stuff.

And then there are capital markets, which are being equally unsubtle. Public markets are saying, with their usual mixture of insight, hysteria, greed, and crowd psychology, that the locus of value is shifting toward the owners of compute infrastructure, memory, chips, models, energy access, and distribution. One can call all of this froth, and some of it surely is. As of this writing, early June 2026, Micron has 10x’d over the past few months into a trillion-dollar behemoth. SK Hynix has also entered the pantheon of trillion-dollar companies. The joke is no longer that software is eating the world. The joke is that capital is buying the electrical grid—and paying for all the attendant accoutrements—so it can instantiate digital workers.

That brings us back to the distributional number that matters most for this chapter: *labor’s share*. If labor’s share is falling while AI capex is rising, if real compensation is pressured while market concentration increases, if entry-level white-collar anxiety is rising while public markets reward operating leverage, then the capital-versus-labor thesis has receipts. Again, one should be careful about over-reading a single quarter or even a handful of quarters. But the data are beginning to rhyme. GDP can be fine. Stocks can rip. AI capex can go vertical. Corporate profits can look splendid. And ordinary workers can still feel economically despondent because the surplus is no longer accruing to them in the old way.

So, can we take Daron’s Nobel back? Mostly joking. Sort of. The Nobel committee can relax; I’m not filing a petition with Stockholm. And again, Daron is right about many things that matter: power, institutions, concentrated ownership, the political economy of technology, the danger that automation can be directed toward labor displacement rather than shared human flourishing. In fact, some of his institutional instincts may become more important as the labor shock intensifies. But on the core question of magnitude and velocity, the Silicon Valley capitalists are winning the argument right now.

The more charitable reading of Daron is that he’s trying to protect us from hallucinating a productivity revolution before it arrives. Fair enough. Economists have scars from the Solow paradox (“You can see the computer age everywhere but in the productivity statistics”), from the slow diffusion of electricity, from the long productivity lag after the personal computer, from every management consultant who promised transformation and delivered a dashboard. Skepticism isn’t a vice. But skepticism becomes error when it fails to notice that the substrate has changed.

Electricity had to be wired into factories. The PC had to be absorbed into offices. The internet had to be built through browsers, broadband, e-commerce, cloud, mobile, and social. GenAI arrives on top of all of that: cloud, APIs, MCP, data centers, SaaS, mobile, EHRs, digital payments, digital identity, enterprise workflow, and a global labor market already decomposed into tasks. The installation surface is already there.

That’s the part the 5-percent view underestimates. AI doesn’t have to build the entire digital economy from scratch. It inherits one. It arrives after the work has already been digitized, outsourced, ticketed, routed, measured, surveilled, recorded, dashboarded, and broken into workflows. In healthcare especially, though I’ll get there in the next chapter, we’ve spent decades turning human work into structured misery: codes, claims, notes, denials, appeals, authorizations, inboxes, work queues, quality measures, documentation templates, call scripts, care gaps, and compliance

workflows. We've already made much of the work machine-legible. Now the machine has arrived.

The Cost of Being Wrong

That's why leaders shouldn't comfort themselves with academic modesty. Academic modesty is attractive in a seminar room. It's dangerous in a boardroom at the beginning of a substitution wave. The right posture isn't hysteria, but neither is it complacency. The right posture is operational seriousness: assume the models keep improving, assume the cost curve keeps falling, assume agentic systems get better at multi-step work, assume *functional verifiability* expands outward from coding, math, and RCM into messier administrative and clinical-adjacent work, and then ask what your labor model looks like if even half of that is true.

Because if Daron is wrong, the cost of believing him will be high. You will over hire. You will preserve management layers that should be compressed. You will treat AI as a departmental tool rather than an enterprise production-function shift. You will bargain with labor as though the old scarcity still governs (this is the big one in healthcare). You will let attrition pass without capturing it. You will keep outsourcing work that should be agentified. You will arrive late, morally surprised and operationally unprepared, when the market starts asking why your denominator didn't expand and your numerator didn't shrink.

If Vinod is wrong, the cost is different. You may move too fast. You may cut into muscle. You may alienate the workforce. You may turn a humane transition into an austerity program with better software. That danger is real, and I'll keep returning to it because the point isn't to transform labor-intensive institutions into sociopathic labor-elimination machines. But if Vinod is directionally right, and think he is, then the real ethical failure will be denial. You can't build a compassionate transition around a reality you refuse to name.

Not panic. Not billionaire triumphalism. Not layoffs as virtue-signaling. But a clear-eyed recognition that the old bargain may be breaking and that the institutions most dependent on human labor have the least time to pretend otherwise. The great decouplings aren't abstract macro curiosities. They are the chassis of the coming labor convulsion. Population may no longer be the limiting input to progress. GDP may no longer require proportional hiring. Productivity may no longer deliver proportional compensation. Capital may increasingly instantiate labor directly. And expertise may increasingly detach from the old developmental pathways that produced it.

That's the uncomfortable macro frame. The next move is to bring it down from the national accounts into the *human hierarchy*: the professions, the credentialing economy, the apprenticeship ladder, and the psychologically destabilizing fact that AI's first great labor shock may fall not on low-status work, but on the people who thought intelligence itself was their moat.

Part IV—The Automation of the Intelligentsia

When the Decouplings Become Biography

Now let's turn from the macro considerations to the human implications. The previous section widened the lens and addressed the big thematic *deconstructions*: population from progress, GDP from employment, productivity from compensation, capital from labor. But those deconstructions and decouplings eventually have to *land* somewhere. They don't respectfully remain abstractions in the national accounts. They show up in the professional ladder, the apprenticeship model, the credentialing economy, the entry-level job market, the promotion path, the compensation gradient, and, most painfully, in the moment when a credentialed person discovers that the machine has learned to do some meaningful part of the thing he spent twenty years becoming certified to do. *The decouplings become biography*. They become anxiety in a law firm recruiting class, a consulting analyst cohort, a software team, a medical residency, a revenue-cycle department, a claims operation, a hospital administrative suite, a university admissions office, a credentialing body, a family conversation about whether a \$300,000 professional degree still makes sense, or at a commencement address where the students boorishly jeer any speaker who utters the two letters A.I. The macro becomes intimate.

That's why the first great psychological shock of AI labor substitution won't be that it replaces low-status labor. We already have a political vocabulary for that: *automation, offshoring, mechanization, globalization, deindustrialization*. These words have been applied, sometimes analytically and sometimes lazily, to blue-collar workers for half a century. Factory workers, machinists, miners, clerks, retail workers, and routine administrative laborers have lived inside this story for decades. The deeper shock is that AI comes first, and most disconcertingly, for the people who thought they were safe: the intelligentsia, the cognoscenti, the credentialed knowledge workers, the prestige professions perched atop the social, cultural, and financial hierarchy and historically insulated from the lower-order technological turbulence below them. In short, it is coming for the neocortex (see my BCI chapter for some elaboration here).

That phrase—coming for the neocortex—sounds dramatic, and maybe a little silly, but I think it's the simplest way to say the thing. For years, blue-collar workers absorbed the dislocations of globalization, offshoring, mechanization, and automation. This is the *white-collar version* of that displacement, except the white-collar class has *less* immunological protection than it imagines. Factory workers collectivized. Longshoremen threaten strikes. Nurses have unions. But consultants, junior lawyers, analysts, coders, administrators, claims specialists, utilization reviewers, revenue-cycle employees, and a great many young professionals built their identities on a different premise: that credentials, intelligence, test scores, apprenticeship, and proximity to difficult knowledge would keep them above the blast radius. That premise is wobbling.

Accountants, lawyers, consultants—and yes, doctors—have far less protection here than they imagine. I don't mean that the very best of these professions disappear. Quite the opposite, at least for a while. The absolute pinnacle of every profession will become more powerful, a kind of *bionic ascendancy* for the already preeminent: the best surgeon paired with a model that has read everything; the best litigator with instant recall of precedent; the best strategist with synthetic analysts running continuously in the background; the best biomedical researcher with tireless hypothesis-generation machinery at hand; the best investor with up-to-the-nanosecond market intelligence—yes, had to include this last example. These people become more formidable. The top becomes *cybernetic*.

But the broader distribution of professional labor looks harsher. GenAI tends to elevate the least skilled faster than it elevates the elite, compressing the distance between the bottom and the upper quartiles of performance. A junior person with a model can suddenly produce work that once required several years of apprenticeship. A mediocre analyst becomes competent. A novice coder ships something usable. A generalist can generate a credible first draft of work that once required a specialist. That's the democratizing part, and it's real. But democratization isn't always professionally benign. When the least skilled improve fastest, the *middle of the distribution*—the respectable, credentialed, competent middle where so much of the economic value of expertise historically resided—gets squeezed. Compensation gradients flatten. Expertise becomes less scarce. And then, depending on the domain, especially where there is functional verifiability, the next phase arrives: not merely automation of the workflow, but automation of the worker.

This is why AI is best understood, in labor-market terms, as the *white-collar version of offshoring*. But the offshore labor pool isn't in Hyderabad or Manila. It's synthetic, infinitely replicable, increasingly agentic, and paid in compute. Offshoring taught us that if work can be decontextualized, decomposed, specified, and quality-checked, it can leave the building. GenAI simply asks the next, more brutal question: once the work has left the building, does it need a human being at all?

The Demonetization of Expertise

This is the uncomfortable bridge from the Great Decouplings to the prestige professions. In the macro frame, capital buys compute, compute buys cognition, cognition becomes output. In the professional frame, the same process looks like the *demonetization of expertise*. I don't mean that wisdom disappears, or that judgment becomes irrelevant, or that social trust can be summoned out of a GPU cluster. Those things still matter, and in some contexts they matter *more*. But much of what the prestige professions sell isn't wisdom in the deepest sense. It's the reliable application of codified knowledge inside a credentialed wrapper. And codified knowledge is precisely what these systems metabolize.

The easiest way to see this isn't to retell the whole history of automation again—I've certainly done enough of that—but to notice what the prestige professions were protected by. Their moat wasn't muscle, arithmetic, filing, or retrieval. Their moat was judgment, interpretation, domain expertise, apprenticeship, social standing, and the priestly aura around difficult knowledge. Medicine has the white coat, the residency, the fellowship, the board certification, the clinical note, the quiet language of ordination. Law has the bar, the partner track, the precedent, the Latin residue, the ritualized adversarialism. Consulting has the deck, the benchmark, the partner whispering reassurance to the CEO, the promise that the answer has been generated by the right people at the right firm with the right font. These professions do useful work, obviously. They also bundle useful work with *scarcity*, *status*, and *authority*.

AI starts by unbundling that bundle. And that phrase matters: unbundling, not merely replacing. A lawyer isn't one thing. A lawyer is research, diligence, drafting, negotiation, risk absorption, judgment, client trust, and courtroom presence, all bundled together inside a credentialed human institution called a lawyer. A doctor isn't one thing either. A doctor is history-taking, diagnosis, triage, ordering, interpretation, procedural skill, empathy, liability absorption, trust, and moral presence. A consultant is analysis, narrative, benchmarking, executive therapy, implementation theater (sorry, as a former consultant myself I can't resist the occasional jab), and occasionally actual strategic insight. The point is that these prestige professions have been priced, honored, and protected as bundles, even though a lot of the work inside them is decomposable. AI first strips out the parts that are codified,

teachable, repeatable, verifiable, and decomposable.

This is where the prestige professions should be most nervous. Their economic power has always depended on treating the bundle as *indivisible*: the memo and the judgement, the diagnosis and the trust, the benchmarking deck and the CEO's need to feel less alone in a difficult decision. AI strips out the parts that are codifiable, teachable, repeatable, verifiable, and decomposable, then forces the remaining human residue to *justify its price*. Judgement becomes the last defense, and in some domains, it will become *more* valuable precisely because the surrounding routine work has been stripped away. But last defense isn't the same thing as permanent fortress. Once the tractable pieces are unbundled, the whole professional package gets repriced. That repricing is the *demonetization of expertise* in motion—and given the velocity of the models, it may not be especially slow.

Here is the brutal little maxim: when it's *teachable*, it's learnable by machines. Can you teach it? Can you decontextualize it? Can you apprentice someone into it? Can you write down its rules, heuristics, exceptions, case patterns, benchmarks, and quality metrics? Then eventually it's learnable, replicable, and *agentifiable*. Codification becomes automation. Standardization becomes benchmarking. Benchmarking becomes optimization. Optimization becomes substitution.

There's an important distinction here, because not all automation has the same labor-market effect. Automation can *raise* expertise requirements or *lower* them. Previous rounds often automated easy, routine, lower-expertise tasks. When that happens, only the harder residual work remains; the threshold for entry rises; the pool of qualified labor shrinks; and wages can *rise* for the surviving practitioners. But automation can also remove the expert part of the task, *lowering* the entry threshold and allowing less skilled labor—or eventually no human labor at all—to perform what used to require training. That's the more destabilizing version. So we should stop saying “this occupation is exposed” as though exposure were destiny. The better question is which tasks are exposed, where the expertise threshold moves afterward, and whether the new bottleneck still needs humans in large numbers.

GenAI is strange because it appears to do both. It automates the easy things *and* the hard things. It writes the email and drafts the legal memo. It summarizes the chart and suggests the differential. It translates, codes, reasons, classifies, routes, synthesizes, and increasingly acts. Pre-agentic AI closed expertise gaps by helping the bottom catch up. Agentic AI starts eliminating the bottom-rung tasks altogether. So the bottom is first uplifted, then disintermediated; the middle is compressed; and the top temporarily ascends into a more leveraged, more cybernetic, more powerful stratosphere. That's the great cognitive leveler. *No more bad doctors, no more bad lawyers, no more bad consultants*—or at least fewer of them, which isn't the same thing as saying no more doctors, lawyers, or consultants. But the hierarchy changes. The credential no longer monopolizes access to knowledge. The apprenticeship model becomes less obviously necessary, and in some cases less economically sustainable. The old pyramid—junior grunt labor below, competent middle practitioners in the middle, high priests at the top—starts to look structurally wobbly.

The Apprenticeship Professions May Cavitate

This brings us to the apprenticeship problem, one of the most under-discussed and genuinely consequential features of the whole transition. The deeper problem isn't only job loss. It's *apprenticeship collapse*. That distinction matters because we tend to talk about labor displacement as if the job itself were the right unit of analysis: the junior analyst, the first-year associate, the entry-level coder, the resident, the new consultant, the revenue-cycle analyst. And yes, those roles matter—not as abstractions, but as the first paycheck, the first institutional affiliation, the first fragile claim on a profession. But the apprenticeship professions were never merely collections of entry-level roles. *They were developmental pyramids*. The bottom of the pyramid wasn't glamorous, and often not especially humane, but it performed an essential formative function: it built judgment through repetition.

Law, medicine, consulting, finance, software, analytics, and much of healthcare administration all relied on some version of this pattern. Juniors did repetitive work, learned the vernacular of the profession, absorbed pattern recognition through volume, made small mistakes under supervision, and gradually earned the right to make larger judgments. The grunt work wasn't incidental to professional formation. It *was* the formation. Chart review, first-pass coding, claims reconciliation, memo drafting, diligence work, document review, spreadsheet archaeology, ugly first drafts, bad analyses, corrective feedback—these weren't merely low-status tasks waiting to be mercifully automated away. They were the repetitive cognitive musculature through which professional competence got built.

This is where the cheerful “AI will free humans from drudgery” line becomes a little too glib. Sometimes drudgery is just drudgery, and good riddance. I'm not going to sentimentalize the fax machine, the prior-auth form, the tenth version of a board deck, or the soul-deadening work of moving information from one bureaucratic aperture to another. A lot of that work deserves to disappear, and we'll be better when it does. But not all repetition is meaningless. Repetition is how judgment becomes embodied. It's how a resident learns that the patient who “just looks wrong” is actually deteriorating. It's how a revenue-cycle leader learns which denial pattern is noise and which is the leading edge of payer strategy. It's how an operator develops the strange internal tuning fork that distinguishes a routine anomaly from the thing that matters. The mistake would be to confuse low-status work with low-developmental-value work. Some of the work we most want to automate was also, awkwardly enough, the way people learned the profession. That doesn't mean we should preserve it in amber. It does mean we need to understand what we're destroying before we congratulate ourselves for the destruction.

If the machine consumes the bottom of the pyramid, the top doesn't remain intact indefinitely. A bot can summarize the chart, draft the memo, write the first-pass code, build the deck, or reconcile the claim. Wonderful, in one sense. Efficient, deflationary, democratizing, all the words we like. But how does the junior human learn the profession if the first thousand reps disappear? You can't cultivate taste if the person never sees enough bad drafts, bad deals, bad charts, bad arguments, bad analyses, bad clinical impressions, and bad operational calls to form an internal standard. The apprenticeship professions don't merely shrink; they cavitate.

I like the use of that word here: *cavitate*. This isn't a simple contraction, not a graceful resizing of the professional pyramid. It's *hollowing from within*. The gray-haired summit remains visible for a while—the senior partner, the attending, the CFO, the physician executive, the operator with thirty years of scar tissue and pattern recognition—so the institution tells itself the profession is intact. But

underneath, the base camp is quietly being destroyed. The ladder loses rungs. The junior work disappears. The middle gets squeezed. And then, a decade later, everyone acts surprised when the institution can't produce enough people with real judgment, because the developmental machinery that produced judgment was treated as inefficiency and automated away. That's the subtle danger. The visible profession survives long after the developmental system has been compromised. The senior people are still there, the titles are still there, the credentials are still there, the institutional ceremonies still proceed with all the appropriate solemnity. But the renewal mechanism has been damaged. The profession continues to look whole from the summit while hollowing out below.

This is also where the cognitive-offloading problem becomes inseparable from the labor problem. If the human role migrates too quickly from *doing* to *verifying*, the human may never acquire the underlying competence required to verify well. Verification presupposes inward *possession* of the work. It presupposes taste, memory, context, pattern recognition, and the ability to know when the machine's answer is polished nonsense. That's why the phrase *human in the loop* can become ceremonial if we're not careful. A junior professional who never did the work may not become a senior professional who can supervise the machine. The human in the loop has to *understand* the loop. Otherwise we've not preserved human judgment; we've preserved a liability fiction.

This is the quiet failure mode of the impending intelligence age. We tell ourselves that the machine will do the routine work and the human will handle the exceptions. But exception-handling isn't an innate human virtue. It's a trained capacity. If the routine work disappears before judgment is formed, the human may be less prepared for the exception precisely when the exception matters most. And this isn't some abstract problem for the epistemologists and the HR futurists. It will matter in clinical care, law, consulting, investing, software, operations, revenue cycle, payer strategy, and every other apprenticeship profession whose senior judgment was built on junior exposure.

The Identity Shock

The psychology of this will be unkind because prestige professions aren't merely occupations. They are identities, *sorting mechanisms*, class markers, mating signals, and little secular priesthoods. Tell a textile worker in 1815 that the loom is coming for his craft, and you're threatening his livelihood. Tell a lawyer, physician, consultant, programmer, analyst, or healthcare administrator in 2026 that a model can perform much of the cognitive substance of his work, and you're threatening his livelihood *and* his cosmology. The prestige professional is losing the story that justified the hierarchy he climbed.

That's why the backlash from the intelligentsia will be so rhetorically sophisticated. It won't announce itself as rent protection; almost no entrenched profession is gauche enough to say, plainly, "please preserve our compensation, prestige, and credentialed scarcity." It will announce itself as ethics, safety, professionalism, humanism, quality, patient protection, and concern for vulnerable populations. Some of that will be sincere and correct. Some of it will be guild defense wearing elevated language. Discerning the difference will be one of the essential leadership tasks of the next decade.

And let me be careful here, because this can too easily become a cheap anti-professional rant, which isn't what I intend. After all, I'm part of this class! The professions preserve real wisdom. Medicine in particular carries tacit knowledge that no model lab can casually reproduce. The attending who has watched ten thousand patients deteriorate, the nurse who knows from the doorway that something has changed, the revenue-cycle operator who can smell a payer anomaly before the dashboard catches it—these people aren't disposable. The tragedy would be to automate away the developmental pathways that produce them while congratulating ourselves for eliminating inefficiency.

So the answer isn't nostalgia for grunt work. It's not preserving every junior task because "that's how I learned," which is the sort of thing professions say when they confuse hazing with formation. The answer is to *redesign apprenticeship for the intelligence age*. If machines do the first draft, humans need to learn by interrogating the draft. If machines summarize the chart, humans need to compare the summary to the messy underlying reality. If machines recommend the code, the differential, the clause, the trade, the operating move, or the workflow action, humans need to learn *why* the recommendation is right, where it's brittle, what it omits, and how reality may humiliate it.

That's a different kind of apprenticeship, and it has to be *designed* rather than *inherited*. The old system created competence accidentally through volume. The new system will have to create competence intentionally through guided exposure to machine output, human correction, real-world ambiguity, and accountability. We can perhaps do this through massive simulations. That may ultimately be better; the old model wasted a lot of human life in low-value repetition, and we shouldn't romanticize it just because it happened to train some people well. But if institutions don't design the replacement, the replacement won't spontaneously appear. The machine will take the work. The institution will book the savings. The developmental ladder will hollow out. And ten years from now, the same institutions will convene a task force to investigate why nobody seems to have judgment anymore.

The Credentialing Economy Under Siege

This also explains why higher education should be nervous. The elite university business model is, in part, a gateway into the priesthoods of income and prestige. For a century, credentials were a proxy for scarce expertise. Elite schools, medical degrees, law degrees, MBAs, residencies, fellowships, partner tracks, tenure clocks—all of it was a *credentialing economy* built on the scarcity of knowledge and the social trust attached to those who possessed it. But if expertise becomes cheap, callable on demand, partially automated, and partially unbundled, what exactly is the credential selling? Authority? Trust? Liability? Network? Taste? Social sorting? Those may still matter, but they are narrower moats than knowledge itself.

I'm not saying the university disappears. I'm saying the credential as a monopoly certificate of competence is under siege. The ROI of a \$300,000 professional degree looks different when the expertise it was meant to certify is partially democratized and partially mechanized. Interesting that the WSJ and others are reporting massive 'discounting' of MBA's these days. But the observation is especially true for long, expensive apprenticeship paths of *all kinds* built around the routine application of established knowledge. The safe harbor is no longer credentialed expertise by itself. It's expertise plus taste, trust, social intelligence, moral courage, institutional fluency, and the ability to direct a swarm of machine cognition toward worthwhile ends.

So what remains valuable? I'd put it into three broad categories, rather than the usual dreary enumeration of every human virtue in a LinkedIn post. **First**, *embodied* and *presence-based* work: the surgeon's hand, the nurse at the bedside, the teacher in the classroom, the caregiver, the builder, the person whose work happens in the world of bodies rather than only in the world of text. At least until Optimus arrives at industrial scale, these domains are highly insulated. **Second**, *social* and *trust-based* work: leadership, persuasion, coaching, conflict navigation, relationship formation, and the ability to move other humans through fear, ambiguity, and resistance. **Third**, *taste* and *judgment under uncertainty*: deciding what matters, knowing when the model is technically correct but socially obtuse, recognizing when the answer is elegant and wrong, and holding moral responsibility when the machine can only optimize.

This is why, in a slightly counterintuitive twist, *high social skill may become more valuable than generic quantitative skill*. The machine will increasingly handle recall, coding, routine analysis, routine writing, generic administration, and the application of established knowledge. The human who wins will be the one who can decide what matters, create trust, organize humans and agents, understand context, exercise taste, and move fluidly across domains. The next generation should probably study history, culture, language, philosophy, psychology, rhetoric, organizational behavior, and AI itself (at least this is what I'm telling my three college-age kids). Broad culture may become the best interface layer with alien intelligence.

That advice sounds heretical in a STEM-saturated age. But if the machine becomes increasingly superior at routine STEM application, and if the human advantage migrates toward taste, judgment, meaning, ethics, leadership, and synthesis, then the humanities come back by the back door. Not as nostalgia. As interface technology. Whoever has the *broadest culture* may be better equipped to interact with the alien intelligence, ask better questions, recognize better analogies, and prevent the machine from narrowing the human imagination into whatever pattern its training distribution most readily supplies. Have a listen to the podcast Dario and I did in January—this may very well represent ‘the revenge of the humanities.’

The practical takeaway isn't that *nobody* should become a doctor, lawyer, programmer, consultant, scientist, or engineer. That would be silly. Some of these fields will remain extraordinarily vibrant and highly remunerative, and the very best people in them may become more powerful than ever. The takeaway is that the old guarantee is gone. Long, expensive professional pathways built around *scarcity of codified knowledge* are, well, *riskier*. Apprenticeship models that depend on junior grunt work need to be redesigned. Institutions that credential expertise have to justify themselves again. And professionals who define themselves principally by the routine application of established knowledge should be the most nervous people in the room.

That world is ending. Not all at once. Not without resistance, demagoguery, litigation, unionization, regulatory enclosure, and a great deal of elevated language about safety, professionalism, and standards. But directionally, unmistakably. The old hierarchy of expertise is being unbundled, repriced, and partially mechanized. And once that happens, the institutions that depend most heavily on prestige labor, credentialed scarcity, administrative cognition, and inherited apprenticeship ladders will have to decide whether they're going to redesign the ladder—or wait until the bottom rungs disappear beneath them.

This is where the sociology of the intelligentsia becomes a labor strategy problem. The AI labor shock isn't merely about how many jobs go away. It's about how professions *renew* themselves, how judgment gets formed, how credentials retain legitimacy, how workers preserve dignity, how institutions distinguish real human value from credentialed scarcity, and how the next generation learns to do work that machines can now partially perform. The automation of the intelligentsia is therefore a reconstitution of expertise itself.

And that's why the next chapter can finally turn to healthcare. Because healthcare contains all of this in one enormous, fragile, labor-intensive cathedral: prestige professions, apprenticeship ladders, administrative knowledge work, credentialed scarcity, clerical sludge, social trust, embodied care, and a vast middle of respectable cognitive labor that looks, from the machine's perspective, increasingly decomposable. The physician isn't the medical secretary. The nurse at the bedside isn't the claims analyst. The surgeon isn't the utilization-management reviewer. But they all sit inside one labor architecture, and that architecture was built for a world in which cognition was scarce, local, biological, slow to train, and expensive to distribute.

That world is ending. The question is whether we build the next one deliberately.

Part V—Resistance, Risk, and Human Consequence

The Orwellian Conspiracy of Silence

One last rant before we turn to our favorite industry. And this is where the *denialism* from the AI illuminati starts to worry me. I understand it, which is different from endorsing it. The public is already anxious about AI. Workers are suspicious. Knowledge workers vote. Prestige professions have cultural power. And the political coalition required to keep building data centers, powering models, permitting transmission, diffusing AI into the enterprise, and tolerating the psychic weirdness of machine cognition is more fragile than the technologists sometimes admit. So of course the Sacks/Jensen/Silicon Valley establishment wants to say, with maximum performative confidence, that the job losses *won't* happen. Why hand your opponents the pitchforks before the revolution has even left the loading screen? Why tell the middle class that the central ROI of the technology may be the automation of the service economy? Why tell the lawyer, consultant, analyst, coder, claims specialist, administrator, or ambitious recent graduate that the machine isn't merely coming to help them, but may be coming to perform the economic substance of what they do?

I get the communications logic. Truly. I just think it's wrong. The problem with enforced optimism is that it eventually collides with reality, and when it does, the earlier euphemism becomes decedentializing. If job loss remains small, fine; the reassurance will age well enough. But if the labor shock becomes visible—first in entry-level jobs, then in white-collar hiring, then in apprenticeship professions, then in healthcare administration, then in labor share, wages, and regional employment—the previous period of message discipline will look like concealment. The public won't say, "Thank you for preserving confidence during the installation phase." It'll say, quite reasonably: *you lied to us*. And this is why I think the biggest issue in the 2028 presidential election will be AI-induced job loss.

Dario said the quiet part out loud last year when he warned about entry-level white-collar job loss and the possibility of sharply higher unemployment, and the reaction was instantaneous: mockery, denunciation, pious rebuttal, the usual Orwellian insistence that the obvious thing *mustn't* be named. The response wasn't, mostly, a sober probabilistic debate about task exposure, agentic capability, deployment curves, substitution elasticity, and the difference between augmentation and replacement. It was indignation. It was message discipline. It was *thou shalt not say the thing*. And then, awkwardly for the denialists, others started saying versions of the same thing, including Elon, not just about entry-level jobs but about white collar information work more broadly. The strongest form of the claim—that anything manipulating information rather than atoms is exposed—is too sweeping if stated as an immediate prediction. There will be bottlenecks, trust constraints, liability constraints, regulatory constraints, embodied constraints, and social constraints. But directionally, the claim is right enough to make the denial feel increasingly theatrical. Work that is digital, decomposable, measurable, and mediated through language or code is now on the frontier.

That is the part worth emphasizing. The rhetoric says *no* displacement. The capital stack says synthetic labor *at scale*. The denial happens at the microphone. The substitution happens in the data center. Again, I'm sympathetic to the strategic instinct behind the denial. If the AI priesthood admitted tomorrow that the central ROI of the technology is the automation of the service economy, the political backlash would be instantaneous. The public would panic. Organized labor would mobilize. Regulators would discover religion. The next presidential race would become, at least in part, a

referendum on whether the same people who got rich from AI also destroyed the first job of your child with the sociology degree, the consulting aspiration, and the \$300,000 credential. So yes, I understand the impulse to manage the narrative. I just don't think it'll age well.

I've been watching Netflix's mesmerizing Chernobyl series (highly recommend, but watch with wine—otherwise it's massively depressing), and forgive the melodrama³, but the analogy keeps intruding. The old desiccated Soviet apparatchiks in the meeting after the core meltdown decide, with absolute institutional confidence, that they can suppress the message. Don't tell anyone. Seal the city. Manage the narrative. Meanwhile the atmosphere is literally iridescent with radioactivity and everyone can see that something has gone terribly, catastrophically wrong. AI labor disruption isn't a nuclear meltdown, obviously. We shouldn't become hysterical or performatively apocalyptic. But suppressing visible reality because it's politically inconvenient is the same human temptation, translated into a different register. Institutions hate bad news. They particularly hate bad news that implicates the legitimacy of the institution itself. So they euphemize. They delay. They say "augmentation" when they mean "substitution." They say "productivity" when they mean fewer people. They say "agentification" when they mean a new labor class. They say "efficiency" when they mean the old employment bargain is breaking. Workers will eventually translate the euphemisms correctly.

The better strategy isn't panic. It's *truth plus solution set*. Tell workers that some work is going away. Tell them early enough that attrition, redeployment, retraining, severance, mobility, community planning, and political preparation are still possible. Tell the public that the alternative to labor substitution isn't some morally pristine status quo. In healthcare, the alternative is continued *unaffordability*, medical debt, psychiatric deserts, delayed primary care, exhausted clinicians, payer-provider trench warfare, and an 18-percent-of-GDP anvil around the neck of American competitiveness. Tell the AI companies they owe society more than a euphemism. Tell the 150 that silence isn't compassion. Silence is fear wearing a communications strategy.

This is especially important because truth-telling changes the moral architecture of the transition. If you tell the truth early, attrition becomes possible. Redeployment becomes possible. Training becomes possible where training is real. Generous severance becomes possible where training isn't. Regional planning becomes possible. Community-college partnerships become possible. Internal talent markets become possible. Local political legitimacy becomes possible. If you wait until the labor math is undeniable, the only instrument left is the RIF, and the RIF arrives as betrayal. That's the real cruelty: not acknowledging displacement, and pretending displacement won't happen until abrupt termination is the only tool left.

Cognitive Offloading and the Verification Trap

Reader Note: This returns to cognitive offloading from Chapter 2 and anticipates the Theology and BCI chapters. In the labor frame, the issue is no longer just epistemic passivity; it's the erosion of apprenticeship, judgment, skill formation, and purpose.

There's another casualty in this transition that's harder to measure than jobs: cognition itself. Work is one of the primary ways adults keep their minds alive. We learn by doing. We develop judgment through repetition. Work is purpose. We become capable of handling edge cases because we've spent years inside routine cases. Automation can interrupt that developmental ecology, especially when the human role migrates too quickly from *doing* the work to *supervising* the machine that does

the work. That was the apprenticeship problem in the prior section, but here it becomes not just a labor-market issue but a cognitive one.

The risk is the *verification trap*. In theory, the machine handles the routine work and the human reserves judgment for the exceptions. That sounds clean, even humane: let the model do the drudgery and let the person handle the meaningful residue. As I've said elsewhere in this essay, judgment isn't a museum object one keeps under glass and retrieves for emergencies. *Judgment is a trained faculty*. It's built through exposure, repetition, error, correction, and inward possession of the domain. If the routine opportunities for practice disappear, the human becomes less capable precisely at the moment the organization says the human is needed for oversight.

This is the paradox of *cognitive offloading*. The more confident a worker becomes in the AI system, the less cognitive effort she may exert. The task shifts from *reasoning* to *verification*, from knowledge with understanding to knowledge without understanding, from inward possession to external supervision. That may be efficient in the short run. It may also be cognitively degrading over time. A clinician who lets the model think for her isn't the same as a clinician who uses the model to think better. A coder who rubber-stamps AI output isn't the same as a coder who understands the logic of the claim. A care manager who follows agentic prompts without context isn't the same as a care manager with relationships, pattern recognition, and moral intuition. The difference may not show up in quarterly productivity metrics. It'll show up when the exception appears. Andrej Karpathy's recent aphorism is my favorite—you can outsource thinking, but you can't outsource understanding.

This is why the phrase *human in the loop* is becoming a talisman, and a lazy one. A bored, deskilled, overloaded, cognitively atrophied human in the loop isn't governance, it's theater. If we reduce humans to intermittent auditors of opaque machine output, we may end up with the worst of both worlds: automation without accountability, and humans without mastery. Verification is a cognitively demanding act that presupposes understanding. A person can't meaningfully verify what she's never learned to do.

The better goal isn't simply human in the loop. It's human agency in the loop. That means training people to understand the tools, challenge the tools, use the tools, and improve the workflows. It means preserving enough substantive understanding that verification remains real. It means designing roles where humans aren't merely passive validators but active stewards, interpreters, and moral agents. The loop has to contain competence, authority, practice, and accountability. Otherwise it's a liability ritual with a nicer name.

This will be hard because convenience is seductive. The model writes the memo, produces the analysis, drafts the note, summarizes the chart, generates the plan, creates the slide, answers the email. One can feel oneself becoming more productive and less mentally engaged at the same time, which isn't great. The stratification here may become severe. The small minority who use AI to *skillsmax*—or *tokenmaxx*, to use the voguish San Francisco lingo—will become dramatically more capable. They'll read more, synthesize faster, test more hypotheses, draft more fluently, and compound their cognition with a tireless thought partner. Many others may become more passive, more dependent, more prone to accepting the machine's output as epistemically superior, more fluent on the surface and less inwardly masterful underneath. Not good.

The exception case is usually the case that matters most. The patient who doesn't fit the pathway, the claim whose denial pattern reveals a new payer strategy, the lab that's only mildly abnormal but wrong in context, the social barrier that makes the medically correct answer impossible to implement—these are exactly the moments when human judgment has to be more than ceremonial. If the

routine practice that builds judgment disappears, the residual human role becomes less capable over time. That's the hidden workforce-design problem. We shouldn't build systems that turn clinicians, administrators, coders, analysts, and managers into *passive approvers of machine output*. If humans are going to govern machines, we have to preserve the cognitive musculature required to govern them.

When the Human Becomes the Robot

Then there's the thornier question of *purpose*. The official story of automation is always liberating: machines will take the dull, dirty, dangerous, disagreeable, or dear tasks—alliterative and plausibly true—and humans will be freed for more creative, relational, and spiritually nourishing work. Sometimes that happens. Often it doesn't. *Often the human becomes the robot.*

Amazon's fulfillment centers are the cautionary tale, not because hospitals are warehouses, but because Amazon has already shown what happens when *machines reorganize human labor around their own tempo*.¹⁰⁵ The old work was physically punishing: walking miles, lifting, carrying, searching. Robots reduced some of that. Good. Kiva-style automation began by moving shelves to people rather than forcing people to walk endlessly to shelves, and that was a welcome ergonomic improvement. But the newer automation stack is no longer merely about reducing walking. It's about *flattening the hiring curve*, increasing throughput, narrowing the human role, and organizing the worker around the machine's rhythm. In some settings, the human residue becomes *stationary verification*, repetitive picking, monitoring, responding to machine pace, and trying not to fall behind the queue of objects arriving from the robotic system. **The machine gets the variety. The human gets the monotony.**

That's the lesson. And I certainly don't mean to pick on Amazon, which is a remarkable company in so many ways. It's just that they're more sophisticated and faster at the robotics diffusion than most everyone else, so I suspect they are meeting some of the unintended side effects and negative externalities of automation earlier than others. I suggest we observe their experience and learn. For healthcare, the danger isn't necessarily some theatrical robot marching down the hall with an aluminum bedpan. The danger is subtler: orchestration systems that remove walking, searching, typing, routing, remembering, coordinating, scheduling, documenting, and escalating, and then leave the human with a narrower, more surveilled, more machine-paced residue. If we're not careful, the robot gets the agency and the person gets the queue.

This matters because work isn't merely compensation. Work is routine, dignity, status, friendship, competence, identity, self-discipline, social recognition, and sometimes a blessed distraction from oneself. People need purpose, discretion, and some narrative of contribution. They don't flourish as biological API endpoints. Automation that improves physical workload while worsening psychic workload isn't an unambiguous victory. It may be *better ergonomics with worse souls*, which sounds melodramatic until you've watched a worker reduced to a verification appendage for a system she doesn't understand and can't meaningfully influence.

Healthcare is especially vulnerable to this failure mode because the work already contains too much repetition, duplication, clicking, compliance, queue management, and administrative self-parody. If we simply bolt agents on top of the current system, we may not liberate clinicians and staff into more human work. We may reduce them to surveillance and exception handling. The robots will do the interesting variety; the humans will approve, deny, escalate, correct, and absorb blame. I don't

¹⁰⁵ Amazon News: "10 years of Amazon robotics: how robots help sort packages, move product, and improve safety" published on June 21, 2022 ([Link](#))

especially like that future.

The whole point of AI in healthcare should be to move humans *back toward the human parts of care*: presence, explanation, reassurance, touch, trust, relationship, complex judgment, team leadership, community embeddedness, grief, fear, hope, and the thousand subtle interpersonal acts that make care feel like care rather than transaction processing. This requires *co-design*. Not techno-solutionism from a 26-year-old founder who has never watched a terrified daughter try to understand her father's ICU prognosis. Not hospital protectionism from a committee that treats every new technology as an affront to its dignity. *Co-design*. The insurgents bring technology, speed, capital, and irreverence. The incumbents bring workflow reality, patient trust, regulatory legitimacy, tacit knowledge, and moral seriousness. Either side alone will build something defective.

If we do this badly or stupidly, workers become monitors of systems they don't understand, patients become input streams, and the institution becomes a cold orchestration layer. If we do it well, the administrative exoskeleton recedes, the clinician gets time back, patients get more longitudinal contact, and the human work becomes more *human*. Healthcare has a chance to avoid the Amazon admonition case precisely because care is still so irreducibly human. But it won't avoid it automatically. If AI strips away variety and leaves humans with exception handling, compliance surveillance, and endless machine-paced verification, we'll have improved one kind of workload while damaging another. It's just a more efficient way to make people feel less alive.

The Transition Has to Be Owned

So the human consequence of AI labor substitution isn't only unemployment, though unemployment matters. It's also the degradation of apprenticeship, the hollowing out of cognition, the narrowing of work, the loss of purpose, the conversion of skilled humans into ceremonial verifiers, and the emergence of a new political class of people who feel—perhaps correctly—that the machine was installed around them, above them, and *against* them. That's why resistance will come. Some of it will be reactionary. Some of it will be guild protection wearing ethical costuming. Some of it will be union militancy opportunistically chasing a new grievance. But some of it will be morally correct. People will sense that something essential is being touched.

And they'll be right. We're not merely replacing tasks. We're reorganizing the relationship among labor, cognition, capital, and human purpose. The serious leader can't respond to that with denial, or with bloodless efficiency talk, or with the usual blandishments about empowerment and augmentation while quietly eliminating the jobs. The serious leader has to own the transition. Tell the truth. Protect the human work. Redesign apprenticeship. Preserve agency. Share the winnings. And refuse to let a technology that could make care more abundant become a machinery for making work more lifeless.

That's the line to hold. **Build the machine, yes. But don't mechanize the people.**

Coda—The New Bargain Has to Be Designed

So let me end where I began. Technology is salvation, but with side effects. That has been the argument from the start, and it remains the only morally serious way to think about this transition. AI may make care, expertise, education, services, software, science, and administration more affordable, more abundant, more proactive, more longitudinal, more synoptic, and less deranged. It may reduce errors, expand access, compress bureaucracy, accelerate discovery, and make expertise less dependent on ZIP code, income, race, institutional proximity, and social capital. That is the salvation side of the ledger, and it's very real.

But the side effects are real too. AI may dislocate millions of workers, break professional identities, weaken local economies, shift power from labor to capital, degrade cognition, hollow out apprenticeship, turn humans into validators of machine output, and provoke a political backlash that makes today's AI debates look quaint. It may reorder the middle class through the very sectors that, for a generation, have quietly functioned as durable employment engines. It may liberate consumers and patients from unaffordable systems while injuring workers who helped build those same systems. That's the cruelty of the transition. The aggregate line may go up while the particular life goes sideways.

We don't have the luxury of pretending one side of that ledger is fake. The techno-optimists are right that abundance matters. The labor skeptics are right that people get hurt. The guilds are right that safety matters. The insurgents are right that safety can become a pretext for incumbency self-protection. The economists are right that the lump of labor is usually a fallacy. The workers are right that their own jobs aren't abstractions. All of these things can be true at once, which is precisely why the problem is hard.

This isn't an IT project. It's not a chatbot strategy. It's not another procurement category, another innovation-office hobby, another vendor-friendly occasion for everyone to say "transformation" and then return to the old workflow. It's the *industrialization of intelligence* inside an economy built around expensive cognition. And because that economy is also built around wages, careers, credentials, cities, families, tax bases, institutions, and identities, the transition can't be governed by the CFO's spreadsheet alone, however beautifully formatted the spreadsheet may be.

The old covenant won't survive unchanged. The implied promise—come here, work hard, endure the bureaucracy, acquire the credential, climb the ladder, and the knowledge economy will provide durable middle-class stability—is going to fracture in some roles, some functions, some regions, and some institutions. Pretending otherwise isn't kindness. It's disingenuousness with a better communications plan.

The new covenant should be explicit: fewer humans doing bad work, more humans doing human work, more abundance at lower cost, and real generosity toward those displaced by the transition. That's the adult version of the AI labor thesis. Not denial. Not deification. Not Luddism. Not sociopathy. *Build the machine, care for the people*, and redesign the system around abundance rather than scarcity.

If I had to reduce the prescription to seven verbs, they would be: *tell, map, stop, build, redeploy, share, and govern*. Tell the truth about labor dislocation. Map the work, not the titles. Stop hiring reflexively into roles whose task content is already exposed. Build or buy the agentic capacity. Redeploy humans toward bottlenecks and genuinely human work. Share some of the winnings with the people and communities that absorbed the old cost structure. Govern the whole thing with enough seriousness that safety is real and not merely a protectionist incantation.

That's the laminar version of the argument. Technology saves us. AI is different because it multiplies intelligence. The old labor bargain breaks because cognition itself becomes industrially scalable. The service economy is exposed because it's built on expensive, unstructured, administratively routed, cognitively scarce labor. The transition can make services cheaper, better, and more abundant. It can also injure workers, hollow out communities, degrade cognition, and provoke a political backlash that smothers the whole thing in procedural cement.

The adult task is synthesis. Move like builders. Govern like adults. Tell the truth like people who respect their workers. Share the winnings like institutions that understand their civic footprint. And never forget that the point of all this intelligence, if the word still means anything, isn't merely fewer employees, higher margins, or a more elegant operating model. It's more care, better care, cheaper care, more knowledge, less suffering, and a more humane system than the one we inherited.

Build the machine. Care for the people.

That's the only version of this revolution worth defending.

The Chapter, Compressed

Here is the chapter, compressed into the governing takeaways.

First, this labor shock differs because the machine has learned the ladder. Previous technologies often displaced one rung while leaving the next rung intact; GenAI attacks the cognitive work workers were supposed to climb toward.

Second, the prestige professions are exposed not because lawyers, doctors, consultants, programmers, and analysts are stupid, but because their economic power depended on bundling codified, repeatable, teachable, and verifiable work together with trust, liability, taste, and social authority.

Third, apprenticeship is the hidden casualty. If the machine absorbs the junior repetitions through which humans historically became senior, we have to redesign formation before the old ladder quietly cavitates.

Fourth, human advantage migrates toward what remains harder to automate: taste, trust, persuasion, moral accountability, embodied presence, social skill, leadership, and the capacity to decide what matters.

Fifth, the Orwellian conspiracy of silence around job loss is a mistake. Workers deserve candor early enough that attrition, redeployment, retraining, generosity, and political preparation remain possible.

Sixth, the new bargain is not machine worship and not labor nostalgia. It's fewer humans doing bad work, more humans doing human work, more abundance at lower cost, and a transition owned with moral seriousness rather than euphemism.

BEFORE WE TURN THE PAGE

The labor chapter tells us why the stakes are larger than automation rhetoric. The next chapter puts that claim under fluorescent hospital light: revenue cycle, clinicians, administrators, unions, affordability, patients, and the moral question of what labor savings are for.

Chapter 5 • Healthcare Labor Convulsions

Healthcare's Oppenheimer Moment

"Hospitals are only an intermediate stage of civilization, never intended, at all events, to take in the whole sick population."

— FLORENCE NIGHTINGALE, "SICK-NURSING AND HEALTH-NURSING," WOMAN'S MISSION, 1893

A WORD ON NAVIGATING THIS CHAPTER

This chapter brings the labor storm into healthcare. It asks why a sector that employs roughly one in seven working Americans, carries immense civic weight, and runs on expensive human cognition is uniquely exposed to substitution, augmentation, and labor reconstitution.

Part I—Healthcare as Landfall: The Point of the Chapter

You know my perverse love of combining odd and incongruous elements, so yes, I briefly considered opening a chapter on healthcare labor with Aristotle, self-weaving shuttles, and his 4th-century BCE prognostications on robotics (not kidding, actually). But a) that would have—finally and irrevocably—lost my 150 readership and b) Florence Nightingale is the better and more unsettling guide here. More than a century before Epic, payer portals, robotic surgery, ambient scribes, or agentic workforces, Nightingale understood something our contemporary healthcare leaders may find unintuitive: the hospital isn't sacred as a form. Care is sacred. Healing is sacred. The *institution* is provisional. When she called hospitals "only an intermediate stage of civilization," she wasn't at all diminishing the moral seriousness of caring for the sick; she was reminding us that the structures of care are historically contingent, and that a serious civilization shouldn't confuse the preservation of an institution with the fulfillment of its purpose.

That distinction matters enormously now. The preceding labor chapter did the upstream work: the old technological bargain, the great decouplings, the automation of the intelligentsia, the cavitation of apprenticeship, and the possibility that GenAI breaks the reabsorption story by attacking scarce cognition itself. I'm not going to re-litigate that whole argument here. Consider it the weather system. This chapter is where the weather makes landfall. Let's take the labor thesis out of sociology, macroeconomics, and indulgent civilizational speculations and bring it into the sector where the implications become urgently tactical and operational: U.S. healthcare.

I'll start with the statement that healthcare isn't *incidentally* exposed to this technology. It's *archetypically* exposed: labor-addicted, data-superabundant, productivity-deprived, biologically complex, administratively overgrown, politically defended, and *morally indispensable*. Those last two words matter most. Healthcare is the largest *intimate industry* in the country: the place where heroic interventions, labor economics, institutional power, family suffering, clinical trust, public finance, private debt, regional employment, bodily vulnerability and spiritual sacredness all collide. So when a technology arrives whose core economic property is the conversion of capital into synthetic cognition, the small question is whether healthcare can become more efficient. The big question is whether the most labor-intensive and morally freighted sector of the American economy can reconstitute itself without betraying either its patients or its workers.

This is why healthcare is the main event.

Why This Is Healthcare's "Oppenheimer Moment"

Reader Note: This returns to the essay's title and opening Oppenheimer frame. The analogy is repeated here because this is where the metaphor leaves the conceptual introduction and becomes a healthcare-labor argument: the breakthrough has escaped the lab and entered a civic workforce.

The title of this essay, repeated at the top of this chapter, isn't meant as melodrama. The analogy isn't to the bomb itself. It's to the moment when a breakthrough escapes the laboratory and collides with institutions. For years, a technology exists as theory, experiment, prototype, paper, promise. Then, suddenly, it stops being an academic curiosity and becomes an organizational fact. The question is no longer whether it works, but what the people who control its deployment choose to do with it. That's the Oppenheimer parallel I'm after: the transition from *invention* to *installation*, from possibility to obligation, from intellectual achievement to stewardship. Oppenheimer's moment was above all the moment when power became inseparable from the *moral burden of deployment*. The governance problem arrived the instant the power became usable. Healthcare now faces a less apocalyptic but still profound version of that problem. We're deploying a technology that can reorder the economics of care, the structure of work, the authority of expertise, and the allocation of human attention inside the most intimate sector of the economy.

That's why the governance question can't be deferred until *after* installation. If the machine can compress labor, the institution has to decide what the compression is for. If the machine can expand access, the institution has to decide whether access actually expands or whether the savings stealthily disappear into margin. If the machine can unbundle professional work, the institution has to decide how apprenticeship survives. If the machine can intervene in administrative and clinical-adjacent workflows, the institution has to decide what counts as acceptable error, who bears responsibility, and how much human judgment must remain substantive rather than ceremonial.

Healthcare's Ok moment, then, isn't the arrival of a single dramatic device. It's the arrival of world-making capacity in a sector that has been allowed to become unaffordable, administratively disordered, regionally essential, and, I'll say it again, morally indispensable. The leaders operating this machine can't pretend they are merely procuring software. They are choosing a *new labor architecture*, a new cost structure, a new relationship between patients and institutions, and a new covenant with the workers whose old tasks may no longer be needed in the same quantity. The power isn't hypothetical anymore. The moral seriousness now has to catch up.

And I'm unconvinced we're going to learn best practices here from centibillionaire technologists peddling an edenic, science-fiction future. We're going to have to figure this out ourselves. And fast.

The temptation, always, will be to domesticate this into normalized, institutional language: innovation, transformation, efficiency, responsible deployment, workforce modernization. Fine. Those words have their place. But they don't quite name the thing. The real issue is whether healthcare can take a technology that commodifies cognition and use it to *make care more abundant* rather than merely make payroll smaller. That's why this chapter keeps returning to labor, not because labor is the only question, but because labor is where the moral and economic consequences become legible. Who does the work? Who gets displaced? Who gets the savings? Who receives more care? Who is

asked to verify a machine she no longer fully understands? Who is left holding the trust relationship when the institution has optimized everything around her? Those are the questions that make this an Oppenheimer moment rather than another procurement cycle.

And because healthcare is so deeply *local*, this will never be an abstract, disembodied optimization exercise. It will be *felt* in clinics, call centers, union halls, city councils, medical staffs, and households. That's the point of slowing down long enough to build the architecture before the panic. This is the difference between installing a tool and governing a phase shift. That distinction matters now, before the operating model hardens around the wrong premise.

Operationally, that means every AI deployment should **answer four questions** before it leaves the pilot theater: what human work is being removed, what patient value is being created, what worker transition is being funded, and what governance mechanism makes the institution accountable when the machine is wrong. Those questions sound simple, which is why they will be easy to ignore. But they are the difference between a healthcare AI strategy and a procurement process with more clever adjectives.

The Laminar Flow of the Argument

This chapter is long (shocking, I know), so I want the hiking path visible before we start machete-ing through the jungle together. The argument moves through a simple chain: *denominator, exposure, civic footprint, mechanism, resistance, doctrine, covenant*. That's the laminar flow. If the reader gets lost in the claims, the codes, the bots, the unions, the GDP math, or my usual excursus into political economy, come back to that chain. First we ask what AI is really working on. Then we ask why healthcare is unusually exposed. Then we ask how substitution travels. Then we ask what leaders owe the people who built the system before the machine arrived.

The whole chapter can be compressed into one sentence: **AI in healthcare isn't mainly an IT upgrade; it's a labor-denominator shock inside a civic industry, and the moral question is whether the surplus becomes abundance or extraction.**

That's the point. The technology can lower cost, expand access, and return human attention to the parts of care that actually require humans. It can also become a margin-recapture machine that alienates workers, hollows out communities, *destabilizes the middle class* (which, as we'll see, is synonymous with U.S. healthcare employment), and gives every guild, union, and populist politician in America a righteous reason to stop the whole thing.

So the structure matters. **Part I** establishes why healthcare is landfall and why this is an Oppenheimer moment rather than a procurement cycle. **Part II** changes the denominator from IT spend to labor architecture. **Part III** diagnoses the exposure stack: labor addiction, productivity debt, dormant data, biology, labor expense, and remote-work decontextualization. **Part IV** widens the aperture to healthcare as regional labor infrastructure, because hospital payroll is often civic ballast. **Part V** shows the rails of substitution: verifiability, decontextualization, teachability, tolerance for error, agents, and bottlenecks. **Part VI** names the human counterforces: guilds, unions, cognitive offloading, verification traps, and the risk that humans become the robot. **Part VII** gives the 150 a doctrine: map the work, use attrition before layoffs, invest in compute, redesign apprenticeship, govern agents, and share the winnings.

The chapter is still about labor, yes. But healthcare labor is never just labor. It's affordability, regional employment, clinical formation, guild power, social trust, household solvency, and the moral economy of care. That's why the analysis has to move from abstraction to installation. We're no

longer admiring the storm system. We're deciding what to do when it reaches the hospital.

Part II—The Denominator Shift: From IT Budget to Labor Architecture

Reader Note: Before we descend fully into the spreadsheet catacombs, you'll notice at the top of each major part I'll give you a paragraph summarizing where we're going. Not that this makes the chapter any less unreadable, but you can see I'm trying.

This first part does one job: it changes the mental model. If the prior labor chapter argued that cognition is being industrialized, then this part asks where that industrialized cognition shows up on a healthcare income statement. The answer isn't the AI budget. The answer is SWB: salaries, wages, and benefits. Once the denominator changes, every other strategic question changes with it. AI in healthcare isn't a software category; it's a labor category. And when the largest input in the system becomes addressable by machine intelligence, the operating problem stops being "how much should we spend on AI?" and becomes something much more uncomfortable: how much of the current labor architecture survives when cognition becomes industrially scalable?

Healthcare as the \$2.9 Trillion Labor Denominator

The first healthcare move is simply to *change denominators*. For a generation, healthcare leaders have been habituated to think about technology as an IT-budget question. How much do we spend on Epic? How much on cybersecurity, analytics, cloud migration, revenue-cycle tools, prior-auth software, virtual care, and whatever other vendor incantation happens to be making the conference rounds this year? That was the *old denominator*, and it produced the *old behavior*: incremental technology stacked in sedimentary layers atop labor-intensive workflows, screens added to paper-era processes, and a permanent sense that healthcare digitized without becoming materially more productive.

This new technology paradigm moves the denominator squarely and unambiguously from *software* budgets to *labor* budgets. That's the whole strategic reframe. The question is no longer whether AI can justify another marginal technology line item. The question is whether AI can work on the largest input in the system. And that input isn't software. It's people.

I've said this before, but it's worth repeating in this healthcare context because the temptation is still to be hypnotized by the surface artifacts of the technology. When you strip away the mesmerizing outputs—the text-to-photorealistic images, the text-to-cinematic video, the Nobel medals hung around the necks of AI titans, the AlphaFold sermon I'm interminably giving—the technology's first enterprise instantiation is much more prosaic and much more disruptive: *brute-force productivity augmentation*, followed by the increasingly plausible *substitutability of capital for labor*.

The frontier won't move evenly, of course. It will start where the work is most decomposable, most digitized, and most verifiable: revenue cycle, coding, routine documentation, payer correspondence, administrative routing. Then it will move outward into messier domains because the models will get better, the agentic harnesses will become more long-horizon and autonomous, and the cost of inference will keep falling. What begins as augmentation becomes, in many domains, substitution. What begins as a tool becomes a worker. What begins as software sold to labor becomes service performed by software. If you skipped the previous chapter, you may want to give it a glance as it diagrams this process out in (overlong) detail.

Sequoia's formulation—credit where it's due—is useful here: the inversion from *software as a service to service as software*. The first SaaS era sold tools to human workers; the second sells the work itself, executed by software. That's not just a cute semantic inversion. It morphs into a vastly different TAM, a different labor theory, and a different managerial consciousness. Sequoia's point is that the work budget in most professions dwarfs the tool budget; in their phrasing, for every dollar spent on software, six dollars are spent on the human service layer around it.¹⁰⁶ The emerging AI-native companies won't merely sell better tools to workers. They'll *sell the outcome* the worker used to produce.

That distinction matters disproportionately in healthcare because the old denominator was never large enough to explain the opportunity. At the macroeconomic level, we move from a \$2 trillion software market into a *\$60 trillion* global services economy. At the microeconomic level, inside a health system or payer, we move from the old three-to-five-percent *IT denominator* to the *labor denominator*, which is often well north of *half* the cost structure. The center of gravity shifts from “what can AI do to our software stack?” to “what can AI do to our work stack?” That's the change in managerial consciousness for the 150 to internalize.

Zoom out and the scale becomes almost continentally big. U.S. national health expenditures reached roughly \$5.3 trillion in 2024, about 18.0% of GDP. The boardroom question isn't how much of the historically modest IT budget can be reallocated to AI. The real question is how much of the roughly **\$2.9 trillion** embedded in healthcare labor can be augmented, compressed, substituted, redeployed, or eliminated.¹⁰⁷¹⁰⁸ Once the denominator shifts from software to labor, AI stops looking like a line item and starts looking like a reinterrogation of the single largest foundational input in the system.

And healthcare's labor denominator is almost incomprehensibly large. The healthcare and social-assistance perimeter employs roughly 24 million Americans—about 15 percent of total nonfarm employment, or roughly one in seven workers depending on how one draws the boundary. The sector still carries well over a million open positions. From March 2025 to March 2026, healthcare and social assistance added roughly 680,500 jobs even as total nonfarm payroll employment barely moved on net.¹⁰⁹ Healthcare employment, in other words, isn't some marginal labor category sitting off to the side of the American economy. It's the country's great labor engine.

That would be less consequential if healthcare were *also* a productivity engine. It isn't. On many measures, healthcare remains the major industrial vertical most impenetrable to productivity improvement over the last generation. That combination—massive labor intensity plus persistent productivity sclerosis—is *exactly what makes the sector so exposed*. If GenAI's industrial logic is the substitution of machine intelligence for human cognitive and administrative labor, healthcare isn't adjacent to the story. Healthcare *is* the story.

The macro context makes the point sharper. Labor's share of output across the broader economy has already fallen into historically low territory; in Q1 2026, BLS reported labor's share at 54.1 percent, the lowest recorded value since the series began in 1947.¹¹⁰ That's the national version of the *capital-labor divergence*: capital captures more of the surplus while labor's claim weakens. Healthcare is the sector-level version of the same argument because it still spends well over half of its giant cost

¹⁰⁶ Sequoia Capital: “Services: The New Software” published on March 5, 2026 ([Link](#))

¹⁰⁷ Brookings: “Rethinking Health Care Labor” published on October 13, 2011 ([Link](#))

¹⁰⁸ American Hospital Association: “Costs of Caring: Challenges Facing America's Hospitals as They Care for Patients in 2026” published in March 2026 ([Link](#))

¹⁰⁹ U.S. Bureau of Labor Statistics: “Health care and social assistance employment increased by 2.9 percent, or 680,500, from March 2025 to March 2026” published on April 9, 2026 ([Link](#))

¹¹⁰ U.S. Bureau of Labor Statistics: “Productivity and Costs, First Quarter 2026” published on May 7, 2026 ([Link](#))

structure on labor. The macro number and the healthcare denominator rhyme. Capital is learning to buy compute, compute is learning to instantiate cognition, and cognition is learning to perform work. The sector with the most labor embedded in its production function has the largest surface area for that substitution to become operational.

This is why the Wall Street Journal's recent observation that AI is distorting practically everything about the economy—GDP, trade, market concentration, labor compensation, and the divergence between capital-market optimism and worker anxiety—lands so directly in healthcare.¹¹¹ That's exactly the kind of distortion one should expect if capital is beginning to substitute for labor at scale. The input is capital. The intermediate product is compute. The economic bet is labor. And in healthcare, labor is everywhere: claims shops, coding queues, payer call centers, documentation workflows, utilization-management departments, scheduling hubs, inboxes, care-management teams, and all the other places where we turned human cognition into a ponderously expensive operating model.

So the first claim of this part is simple: *healthcare leaders need to stop thinking about AI as a software category and start thinking about it as a labor category*. The old denominator was IT spend. The new denominator is SWB: salaries, wages, and benefits. The phrase “AI budget” is almost comically undersized. The healthcare AI budget is the labor budget, the contracted-services budget, the outsourced-work budget, the management-layer budget, the call-center budget, the documentation budget, the prior-auth budget, the coding budget, and the whole edifice of salaried human cognition that keeps the current system perpetuating itself. The blunt version from my original notes was right: the fundamental ROI of GenAI in healthcare—and, paradoxically, the democratization of healthcare—comes from the roughly \$2.9 trillion labor denominator, not from another interface sitting precariously on top of Epic.

The right taxonomy is broader than *automate* or *augment*. The 150 should be asking four questions at once: what can be automated, what can be augmented, what should be eliminated because the workflow itself is irrational, and **what becomes newly possible because intelligence has become cheaper**. That fourth category matters most. If AI merely lets us perform old, ossified processes faster, it will become just another tool of administrative acceleration. If it reveals care models that were previously impossible because every incremental touch required another expensive human minute, then the technology becomes a path toward abundance rather than a weapon against payroll.

That's the denominator shift. It's not a comfortable question. But it's *the* question.

The Shortage Narrative Is Not a Strategy

This is why healthcare can't comfort itself with the old shortage narrative. Yes, we've perpetually got open requisitions. Yes, clinicians are burned out. Yes, demand is real, and seemingly insatiable. Yes, there are geographies and specialties where labor scarcity is painful, dangerous, and not remotely theoretical to the people trying to staff a Monday morning clinic or a Thursday night ICU. But labor scarcity is precisely the condition that *invites* automation once a capital substitute becomes available. Britain mechanized, in part, because wages were high and coal was cheap. *U.S. healthcare will agentify because labor is expensive and inference is getting structurally cheaper*. That's the analogy for the 150 to metabolize: the old ritual in which everyone solemnly agrees that the central strategic fact of healthcare is permanent shortage no longer works once machine intelligence can begin substituting for the work that shortage made expensive.

¹¹¹ Wall Street Journal: “The Record Divide Between Corporate Profits and Worker Pay” published on May 28, 2026 (Link)

The shortage narrative isn't wrong in the present tense. It's obsolete as a long-term strategic assumption. Anyone running a hospital knows the staffing pain is chronic and unrelenting. But the response can't be to extrapolate today's shortage linearly into 2035 and then build an entire labor strategy around the premise that every current role remains structurally necessary. That's still the reflexive move from self-interested medical specialty societies, nursing groups, workforce consultants, and academic forecasters: the AAMC projects a physician shortfall of up to 86,000 by 2036; BLS projects roughly 189,100 registered-nurse openings per year; and home health and personal care aides, the demographic scaffolding of an aging society, are projected to have roughly 765,800 openings per year.^{112 113 114} Those numbers are real. They're also artifacts of the current production model we're leaving behind. And as a consequence, I disbelieve almost all of these projections—not because the present shortages are fake, but because the future production model they assume is about to be vandalized. A shortage isn't an immutable law of nature; it's a price signal. And when a price signal gets loud enough, capital comes looking for a substitute.

The practical danger is that shortage thinking becomes a sedative. It tells boards that the safest plan is to hire more, train more, credential more, subsidize more, and lobby for still more pipeline expansion. Some of that is necessary in the immediate term. I'm not suggesting we stop training doctors, nurses, therapists, pharmacists, or home-health workers because someone in Palo Alto said "agents" with sufficient conviction. But as a strategic frame, the shortage narrative ignores the obvious *industrial response to expensive scarcity*: substitution. Once inference is cheap enough and agents are reliable enough, the shortage forecast stops being a prophecy and becomes a map of where capital will attack first.

I'm certainly not making the absolutist claim that this is all going to be catastrophic or that we're headed for breadlines in healthcare. I'm also not blind to the regulatory thicket that surrounds this space, which will impede and slow diffusion of any Adam Smith free market mechanism here. On the contrary, I'm realistic around probable timeframes, and optimistic that the ultimate net effect for society will be strongly positive and strongly deflationary: more access, more affordability, more continuous engagement, more prevention, fewer people lost in the administrative labyrinth. But optimism about the destination doesn't erase the pain of the transition. If several hundred basis points can be wrung out of the 18 percent of U.S. GDP consumed by healthcare over the coming decade, that would be economically and morally transformative. It would also entail large-scale labor reorganization, the evaporation of entire job categories, and a rebalancing of power between labor and capital. Given the planetary size of the industry, even modest deflation in healthcare becomes a *national economic event* rather than a sectoral efficiency story.

¹¹² AAMC: "The Complexities of Physician Supply and Demand" published in March 2024 ([Link](#))

¹¹³ U.S. Bureau of Labor Statistics: "Registered Nurses" updated on June 7, 2026 ([Link](#))

¹¹⁴ U.S. Bureau of Labor Statistics: "Home Health and Personal Care Aides" updated on May 13, 2026 ([Link](#))

Affordability Is the Moral Denominator

This is why the 150 have more than an economic interest in the transition. They have a *moral and civic obligation* to steward it with unusual care. Healthcare isn't a normal labor market: it's a middle-class employment engine, a regional anchor, a majority-female and minority-immigrant workforce structure, an (increasingly militarized) union sector, a public-finance epicenter, an employer-cost burden, and the place where every family eventually discovers that the abstractions of health economics aren't abstract at all. If healthcare manages this transition with grace, candor, and some measure of generosity, it might provide a model for the rest of the country. If it manages the transition as a cynical margin-recapture exercise, it will deserve much of the backlash it receives.

The moral valence is therefore, well, complicated. If substituting machine intelligence for human labor merely fattens margins for incumbents, it will be perceived as—and may indeed be—predatory. If it makes healthcare materially more affordable, ubiquitous, proactive, humane, and less administratively impenetrable, then it could become one of the most important public-policy victories of the next decade. And given the government's "the earth is flat" refusal to acknowledge impending AI-induced job loss with anything approaching adult seriousness, a humane transition across the 150 could become a kind of *instruction manual for the broader economy*. The same act—replacing human labor with machine intelligence—can be either a local margin grab or a geopolitically critical deflationary project. That may sound like wild hyperbole but give me a few pages to defend the claim. The difference will be design, governance, and how the winnings are shared.

Affordability isn't some localized, proprietary CFO word. In healthcare, affordability is *access*, *household solvency*, and *national competitiveness*. KFF estimates that Americans owe at least \$220 billion in medical debt; private health-insurance spending reached roughly \$1.645 trillion in 2024; and the average employer-sponsored family premium reached \$26,993 in 2025, with workers contributing \$6,850 on average.¹¹⁵ That's the lived reality beneath the numerical abstractions. Affordability is whether a family delays care, whether an employer can raise wages instead of premiums, whether a state budget can fund schools rather than Medicaid inflation, whether a rural hospital survives, and whether the country can compete geopolitically while dragging an 18-percent-of-GDP anvil behind it. The labor denominator is a lot more than a cost line. It's the economic expression of our failure to make care abundant.

That's why the status quo can't claim the moral high ground simply because it employs people. A system that protects every incumbent labor category while tens of millions of Americans carry medical debt, defer care, fear the deductible, or feel financially fragile when someone in the family gets sick isn't obviously humane. It's humane to the payroll and inhumane to the household.

One reason the labor argument becomes so difficult in healthcare is that both sides can lay justified claim to being morally serious. The worker inside the institution is a real person with a mortgage, a family, a commute, a manager, a professional identity, and deserving of institutional loyalty. The patient outside the institution is *also* a real person with a deductible, a delayed appointment, a denied claim, a collection letter, a frightened spouse, and a body that doesn't care about our workforce-planning sensitivities. The employer paying another premium increase isn't an abstraction. The state budget bending around Medicaid inflation isn't an abstraction. The rural community losing access isn't an abstraction. In the coming turbulence, *we risk letting the visible worker crowd out the less visible patient*, household, taxpayer, and employer who pay for the system's inefficiency. Dario said as much when I asked him in our podcast what we should do to prepare for healthcare worker

¹¹⁵ Kaiser Family Foundation: "2026 Employer Health Benefits Survey" published on October 22, 2025 ([Link](#))

dislocation, and his response was essentially: *I'm thinking first of the patient.*

That answer is uncomfortable, but it's not wrong. Preserving every incumbent labor category can't automatically be called compassion. Sometimes it's status-quo protection with better moral language. A healthcare system that employs people in Kafkaesque administrative loops while patients carry medical debt and clinicians burn out isn't obviously humane. The fact that a workflow employs someone doesn't make the workflow sacred. *The fact that a worker deserves dignity doesn't mean the task deserves preservation.* Those are different questions, and serious leaders have to keep them separate.

The correct moral unit, then, is the covenant *among* patient, worker, institution, employer, taxpayer, and community. AI makes that covenant newly negotiable. If machine intelligence can reduce cost and expand access, healthcare has an obligation to use it. If machine intelligence displaces workers who built the old system, healthcare has an obligation to help them transition. The status quo isn't morally neutral just because it's familiar. The transition isn't morally innocent just because it's efficient. Both require judgment.

Labor Reconstitution, Not Labor Reduction

I reflected on this more broadly in the preceding chapter, but it comes with more urgency and force in healthcare. The uncharitable version is *labor reduction for margin*. The serious version is *labor reconstitution for abundance*: fewer humans trapped in adminstrivia, more care delivered at lower cost, more human attention where touch and trust actually matter, and a transition compact for workers whose roles were artifacts of a system we should have been ambivalent about long before the machine arrived.

This is the narrow ridge the 150 have to walk. Move too slowly, and healthcare preserves an unaffordable, labor-addicted status quo while calling inertia compassion. Move too precipitately, and the industry turns the most powerful productivity technology in its history into a sociopathic margin-recapture and labor-destroying machine. The point isn't to choose between patients and workers. The point is to build a system in which patients get affordability and workers get honesty, generosity, and a real path into the next labor architecture.

That's the moral predicate. *Now comes the diagnostic one.* Why is healthcare so unusually vulnerable? The answer isn't one thing; it's a stack. Healthcare combines expensive labor, decomposable workflows, messy information, high-dimensional biology, and decades of productivity debt. That stack is what makes the sector so exposed.

The distinction matters: *denominator is size, and exposure is susceptibility.* A large denominator without exposure is merely expensive. Exposure without a large denominator is merely interesting. Healthcare has *both*, which is why this chapter can't stop at "AI might help with documentation." The question is whether the whole labor production function of care begins to move.

Part III—The Exposure Stack: Why Healthcare Is Uniquely Vulnerable

The denominator tells us why the stakes are large. The exposure stack tells us why healthcare, more than almost any other major sector, is susceptible. This part is a diagnosis, not yet a playbook. It asks why this particular industry is so exposed to a technology that metabolizes language, workflow, messy data, and expensive cognition.

The argument moves in layers. First, healthcare's largest input is human labor, and that labor is both expensive and scarce. Second, the sector *digitized* without *modernizing*, accumulating productivity

debt instead of productivity gains. Third, all that failed digitization produced a prodigious quantity of *dormant cognitive capital*: notes, claims, images, messages, denials, appeals, transcripts, and workflows that were historically too messy for software but increasingly legible to models. Fourth, biology itself is a high-dimensional domain where the returns to intelligence may be higher than almost anywhere else. Fifth, U.S. healthcare isn't merely labor-intensive; it is labor-expensive, guild-defended, subsidy-supported, and still productivity-deprived. And finally, COVID accidentally revealed one of the cleanest exposure maps we have: the more work can be *decontextualized*, routed through a queue, performed behind a screen, and measured at a distance, the more vulnerable it becomes.

So that's the spine of this part. Denominator tells us how big. Exposure tells us why here. And the exposure stack tells us why healthcare can't keep comforting itself with the idea that it is too complex, too regulated, too human, or too special to be transformed by machine intelligence.

Healthcare's Structural Exposure

Healthcare, with its signature human intensity, has the greatest susceptibility of any major modern industry to disruption by GenAI. If this technological phase shift is singular in its potential to substitute machine intelligence for human labor, then it stands to reason that the industry with the greatest labor dependency has the largest surface area for dislocation. That's the basic arithmetic. But the deeper point is that healthcare is labor-heavy in *exactly the kinds of work GenAI is built to metabolize*: documentation, coding, coordination, routing, summarization, reconciliation, justification, billing, authorization, appeal, triage, measurement, and the endless administrative narration by which American medicine explains itself to, well, itself.

So let's diagnose why U.S. healthcare is so exposed before we get prescriptive—before the moratorium on open recs, the attrition strategy, the GDPval (I'll explain), the numerator-denominator question, the difficult conversations with labor, the inevitable union counteroffensive, and the “please stop building beds and start investing in compute” homily you'll hear from me later in this chapter. The previous sections made the broad case: intelligence is becoming scalable, capital is learning to instantiate cognition, the prestige professions are being unbundled, and the apprenticeship economy is starting to cavitate. This section is about why all of that suddenly converges in healthcare.

No major sector of the American economy remains more human-intensive, more administratively encumbered, more saturated with dormant information, or more dependent on scarce cognition. No sector has more expensive human service embedded in more workflows with more unmet demand sitting on the other side. Healthcare is exposed to GenAI because its labor is expensive, its workflows are decomposable, its information is messy and mostly unstructured, its scientific substrate is high-dimensional, and its institutions are encumbered by decades of accumulated productivity debt. That combination is almost purpose-built for a technology whose comparative advantage is converting unstructured human work into machine-mediated inference.

Labor Addiction: The Largest Input Becomes the Target

The first reason is the most foundational and inescapable: healthcare is labor-addicted. Not just labor-*intensive*, which sounds almost neutral, but labor-*addicted*, which feels mathematically and structurally closer to the mark. Healthcare and social assistance employ roughly 24 million Americans, about 15 percent of total nonfarm payroll employment—roughly one in seven working Americans, depending on where one draws the perimeter. In March 2026, the sector still had more than one million job openings, and from March 2025 to March 2026 it added roughly 680,500 jobs even as total nonfarm payroll employment barely moved. The Wall Street Journal's recent labor-market decomposition was even more of a punch: health services have been adding about 64,000 jobs a month while the rest of the private sector has been contributing an anemic 9,400.¹¹⁶ And nationally, those healthcare and social-assistance jobs now dwarf manufacturing and retail in the labor-market story I'll come back to in a moment.

In other words, remove healthcare from the employment picture and the US labor-market Jenga puzzle starts to collapse.

The occupational texture here matters too. Healthcare's labor base isn't a monolith; it's a huge layered organism: roughly 3.3 million registered nurses, about 4.0 million home health and personal care aides, nearly 1.4 million nursing assistants, and then the long tail of medical assistants, pharmacists, therapists, coders, secretaries, administrators, and technicians who keep the organism moving.¹¹⁷ It's also a deeply immigrant workforce. The Migration Policy Institute estimated that immigrants accounted for more than 18% of U.S. healthcare workers in 2021 and almost 40% of home health aides.¹¹⁸ That makes the coming *labor reconstitution* not only a payroll problem, but a class, gender, immigration, and regional-stability problem.

That should make us nervous, not complacent. The predominant input into U.S. healthcare remains non-tech-enabled labor. For the last generation, the math was straightforward: more demand meant more people—more clinicians, more coders, more schedulers, more utilization reviewers, more care managers, more revenue-cycle staff, more *everyone*. We've added humans because humans were the only way the system knew how to create capacity. Meanwhile, technology—the thing that's intrinsically deflationary in so many other sectors—has often had the opposite effect in healthcare: more screens, more documentation, more compliance, more coding, more administrative mediation, more work created by the supposed tools of work reduction.

Labor scarcity, then, isn't protection from automation. It's the invitation. In prior industrial transformations, high labor costs and labor scarcity catalyzed substitution. That brutal little economic fact made it rational to substitute capital and energy for human muscle even when the machines were awkward, expensive, and politically disruptive. U.S. healthcare is now in a structurally analogous position. Labor is expensive, demand is seemingly inelastic, staffing constraints are chronic, and inference is getting cheaper. That's the condition under which capital starts looking for a substitute.

Worth glancing over at our friends in Big Tech to see the asymmetry, at least for now. Layoffs.fyi's tracker shows roughly 264,000 tech employees laid off in 2023, 153,000 in 2024, 124,000 in 2025, and more than 110,000 so far in 2026 as of mid-May; other trackers using narrower U.S.-only

¹¹⁶ Wall Street Journal: "Healthcare Jobs Are a Rare Bright Spot in the Stalling Labor Market" published on September 8, 2025 ([Link](#))

¹¹⁷ U.S. Bureau of Labor Statistics: "Nursing Assistance Week recognizes the third-largest healthcare occupation" published on June 12, 2025 ([Link](#))

¹¹⁸ Migration Policy Institute: "Immigrant Health-Care Workers in the United States" published on April 7, 2023 ([Link](#))

methodologies report different totals, but the arrows all point in the same direction.¹¹⁹ Tech has already internalized the post-AI calculus of doing more with fewer people. Even if we discount, by a large margin, the real number of job reductions directly attributable to AI substitution versus AI-washing as a pretext for correcting pandemic-era over hiring, misbegotten CEO strategies, ill-advised M&A, or whatever, the intuition holds: the decoupling between growth and employment is underway in tech firms. Healthcare, by contrast, keeps hiring. That divergence is telling, and a harbinger of where healthcare may be heading from here. Healthcare still too often treats staff growth as evidence of mission, scale, and organizational seriousness. That psychology may age poorly.

GenAI's first wedge into healthcare won't be robotic surgery or fully autonomous diagnosis. It will be documentation, coordination, coding, routing, summarization, justification, and administrative narration—the *clerical substrate* of medicine itself. When a sector's core unit of labor consists of moving information through a bureaucracy, and the new technology is unusually good at alchemizing language, forms, rules, claims, denials, and workflows, the exposure isn't subtle. Again, this is the old British wages-and-coal story translated into healthcare. The first Industrial Revolution didn't begin in Britain only because the British had better machines. It began because the economic predicate made machines worth building: labor was expensive and energy was cheap. U.S. healthcare has recreated that predicate in service-sector form. Labor is expensive, demand is subsidized, supply is constrained, and now machine cognition is becoming cheap. That combination drives substitution pressure.

Productivity Debt: Digitized, but Not Modernized

The second vulnerability is almost the inverse of the first: healthcare largely missed, or I might say was impervious to, almost every major technology phase shift that transformed the rest of the economy over the past generation. *Internet, mobile, cloud, analytics, big data, enterprise SaaS*—each reorganized large parts of retail, finance, industrials, logistics, manufacturing, media, and hospitality. In healthcare, by contrast, technology often produced the opposite of what technology is supposed to produce. Elsewhere, software tended to be *deflationary* and *simplifying*. In healthcare, it frequently became *inflationary* and *additive*: more costs, more screens, more documentation, more compliance, more coding, more portals, more work queues, more administrative intermediation, and more labor required to make the “digital” system function.

This is where the Kurzweil intuition is useful. Information technologies usually ride exponential price-performance curves; healthcare rides something much closer to a cost-escalation treadmill. BLS/FRED data show the CPI for medical care rising from roughly 261 in 2000 to roughly 592 in April 2026, meaning medical-care prices more than doubled over that period.¹²⁰ Put that beside the price-performance collapse in compute and you get the pathology in one juxtaposition: the world of computation gets relentlessly cheaper per unit of capability, while the world of healthcare gets relentlessly more expensive per unit of access. Or, to put the numbers in a more stylized (and slightly deranged) but useful way, courtesy first of Ray Kurzweil and then my friend Claude: an inflation-adjusted dollar of healthcare purchasing power from 2000 would buy only about 83 cents of comparable healthcare by 2025, while an inflation-adjusted dollar of effective compute would buy roughly 156,000 times as many FLOPS. The comparison isn't perfect, but the contrast is clarifying. Compute compounds downward. Healthcare compounds upward.

¹¹⁹ Layoffs.fyi: “Tech Layoff Tracker” access June 2026 (Link)

¹²⁰ Federal Reserve Economic Data: “Consumer Price Index for All Urban Consumers: Medical Care in U.S. City Average”, via FRED, Federal Reserve Bank of St. Louis, retrieved June 2026 (Link)

This isn't unique to healthcare, of course. It seems to rhyme with the other government-intermediated sectors that drive so much American cost disease, as Marc Andreessen has trenchantly pointed out: specifically, housing and education. *Constrain supply, subsidize demand*, wrap the whole thing in politics, credentialing, zoning, licensing, public financing, and incumbent protection, and then we're shocked when prices go up and voters reach for pitchforks. In housing, the mechanism is zoning, permitting, environmental review, neighborhood vetoes, and demand subsidies that chase constrained supply. In higher education, it's accreditation, prestige scarcity, federal loans, administrative bloat, and the cartel-like pricing power of selective institutions. In healthcare, it's licensure, certificate-of-need laws, scope-of-practice constraints, reimbursement machinery, payer-provider trench warfare, and demand subsidized through employers, Medicare, Medicaid, ACA subsidies, and the tax exclusion. The details and nomenclature differ, but the political economy has a familiar shape: subsidize demand, constrain supply, watch affordability deteriorate, then add still more subsidy to soften the pain caused by the original constraints.

HITECH is our canonical healthcare example. The federal government poured tens of billions of dollars into EHR adoption, and adoption became nearly universal. ONC reports that by 2021, 96% of non-federal acute-care hospitals and nearly four in five office-based physicians had adopted certified EHRs; the 2024 National Electronic Health Records Survey puts office-based physician adoption of any EHR at 95.0%, with 83.6% using a certified system.^{121 122} That's a digitization success story, at least on paper (pun intended, sadly, given how entrenched our paper-based system still is). But the bigger goal, interoperability, remains elusive; physicians still spend indefensible amounts of time documenting, and the operating model often feels like a paper-era workflow trapped behind a screen. Healthcare got digitized. It didn't get simplified. It computerized without becoming legible. It preserved the old mess and just added some screens.

The reasons are familiar enough: regulation, litigation risk, physician autonomy, fragmented incentives, entrenched oligopolies, and the absence of true process standardization. Pre-GenAI automation required structured inputs and uniform workflows, exactly the conditions healthcare lacked. So while other sectors spent two decades assimilating the software era (and notching both the productivity gains and cost *deflation*), healthcare layered software on top of complexity. It accumulated digital sediment. The EHR became a compliance-and-billing monolith with a clinical interface bolted onto the front.

Here's the paradox, and it's a real one: that recent historic failure may now *increase* healthcare's vulnerability to what I'll shorthand as *agentification*, rather than diminish it. Because healthcare never fully embraced the SaaS generation, it has fewer brittle workflow automations to unwind. It can now plausibly *leapfrog* from fax machines, portals, EHR work queues, and semi-structured administrative quagmire directly to LLM-enabled agents embedded inside systems of record. The very *lack of standardization* that crippled rules-based automation may make probabilistic language models more useful, not less. Old software needed healthcare to clean itself up before automation could work. GenAI can begin working on the mess itself.

I've been alluding to this for several paragraphs, so let's now have a direct look at the productivity data for healthcare. They aren't encouraging. The canonical study, McKinsey's 2019 publication, found that from 2001 to 2016, healthcare delivery contributed 9% of real U.S. GDP growth but 29% of net new jobs. In real terms, healthcare delivery grew faster than the broader economy, but labor

¹²¹ Office of the National Coordinator for Health Information Technology: "National Trends in Hospital and Physician Adoption of Electronic Health Records" updated in March 2022 ([Link](#))

¹²² National Electronic Health Records Survey: "NEHRS Results and Publications" published on December 14, 2025 ([Link](#))

contributed 99% of that growth, more than two-thirds of which came from workforce expansion. Multifactor productivity contributed negative 13%, and McKinsey's summary sentence is the indictment: job creation, not labor productivity, drove most of the sector's growth.¹²³ Bottom line: healthcare didn't grow primarily by becoming more productive. It grew by adding people.

That's our conundrum. Healthcare has absorbed digital technology without absorbing productivity. It has turned software into labor *support* rather than labor *substitution*. It has expanded employment while the rest of the economy kept finding ways to do more with fewer humans. Put more sharply: healthcare turned digitization into a jobs program. We bought systems that required trainers, super-users, analysts, scribes, coders, portal teams, help desks, compliance staff, and entire managerial subcultures devoted to reconciling the digital representation of care with the stubborn messiness of care itself. It was computerization plus labor accretion. GenAI is dangerous to this existing order because it can finally read the sediment.

The ramparts aren't as sturdy as incumbents like to imagine. Dr. Oz and Chris Klomp-driven interoperability mandates are tightening. Prior-authorization rules are becoming more API-enabled and deadline-driven. Information-blocking enforcement has more teeth. TECCA is operational. The fortress is still there, but it's increasingly instrumented. Healthcare exceptionalism bought time, but no durable immunity. The sector waited out internet, mobile, cloud, SaaS, and big data with remarkable institutional impenetrability. I don't think it will be able to wait out GenAI.

Dormant Cognitive Capital: The Data Mess Becomes the Model Substrate

The third vulnerability follows naturally from the second. A sector that digitized without simplifying didn't produce much productivity, but it did produce a continent-sized superabundance of data. Healthcare's susceptibility to AI is inseparable from its informational scale. IDC famously projected the global datasphere would reach 175 zettabytes by 2025. Whether one prefers 175, 180, or some adjacent stratospheric number isn't the interesting part. The interesting part is that the number has escaped ordinary human intuition.

Let me try to human-scale what is otherwise just a science-fiction abstraction. I'll channel Reid Hoffman here, who has done a nice job making these crazy conceptualizations a little more intelligible: one zettabyte is one trillion gigabytes. If a typical e-book is roughly 2.5 megabytes, then 175 zettabytes in a year is something on the order of *133 billion e-books worth of data every minute*. That's a slightly deranged sentence to type, but it helps make the scale legible. Humanity is now generating, in minutes, informational volumes that would have been indistinguishable from infinity to almost anyone before the digital age. The old Eric Schmidt line—that humanity created roughly five exabytes from the dawn of civilization to 2003 and was later producing that much every couple of days—was meant to sound vertiginous. Now even that comparison feels quaint.

And yes, healthcare is one of the great factories of this informational abundance. The often-cited RBC estimate is that healthcare generates roughly 35% of the world's data volume, with healthcare data growing faster than data in manufacturing, financial services, and media.¹²⁴ I wouldn't treat the exact 30% as holy writ—these estimates get thrown around a bit too casually for my taste—but the directional point holds. Hospitals, health systems, payers, labs, imaging centers, pharmacies, device manufacturers, and life-sciences companies generate staggering quantities of high-entropy information. The sector is basically a data refinery that historically lacked the machinery to refine its own output. Not anymore.

¹²³ McKinsey & Company: "The productivity imperative for healthcare delivery in the United States" published on February 27, 2019 ([Link](#))

¹²⁴ RBC: "The healthcare data explosion" ([Link](#))

The more important feature isn't just volume. It's structure, or, in this case, rather the lack of it. A substantial share of healthcare data remains *unstructured* or *semi-structured*, and multiple industry and academic sources put the share around 80%, especially when one includes the rafts of clinical notes, imaging, scanned documents, reports, audio, and other non-tabular artifacts. That matters because prior software was often useless at reading this kind of information. It wanted clean fields, standard definitions, normalized tables, and uniform workflows. Healthcare gave it narrative notes, PDFs, faxes, images, local conventions, idiosyncratic workflows, federated sources, and twenty-seven ways to say the same clinical thing unclearly.

Before LLMs, extracting value from this data morass required laborious (and expensive) structuring, labeling, normalization, mapping, de-duplication, and governance before insight could be derived. The ROI often failed the CFO test, and so the data just kinda sat there, becoming what I might call *dormant cognitive capital*. We had the information, but not the native technology for assimilating and activating it. The clinical note existed, but its meaning was trapped in messy prose. The denial letter existed, but its pattern was trapped in some anonymous queue. The image existed, but its implications were trapped inside a specialist workflow. The data were abundant; the *semantic layer* was missing.

That's what changes. Large language models, retrieval systems, vector databases, and multimodal architectures can ingest, search, synthesize, and reason across unstructured repositories without requiring every scrap of information to be pre-domesticated into a tidy relational database. Vectorization turns messy text and images into searchable semantic space. Retrieval lets the model ground its answer in local institutional knowledge. Multimodal systems let notes, images, audio, and structured fields begin to inhabit the same analytic universe. The work of organizing, ontologizing and querying data stops being pure overhead and becomes convertible into labor savings, decision support, administrative compression, and eventually clinical augmentation.

Put differently, healthcare has been sitting for years on a *mountain of unactivated intelligence*. The problem wasn't that the data didn't exist. The problem was that the data were inaccessible to the dominant software paradigm. Now we have, or at least are beginning to have, a technology natively suited to the mess. That's why healthcare's data problem is no longer only a liability. It's suddenly a huge unlock.

This is the strange inversion at the heart of the exposure thesis. Healthcare's failure to modernize made it expensive, frustrating, and productivity-deprived. But it also created a vast reservoir of digitized, messy, underused information that GenAI is unusually well suited to work upon. *The sector's backwardness becomes its new leapfrog opportunity*. The same messy data detritus that defeated the last software era becomes the fountainhead for the next one. And that means the data mess isn't only a data problem. *It's a labor map*. Every note, denial, appeal, message, code, image, transcript, and work queue is a fossil record of human cognition being spent to move information through an irrationally complex system. Once that fossil record becomes legible to models, the institution can ask a colder question: why is a human still doing this? Sometimes the answer will be safety, trust, law, taste, or judgment. Often the answer will be inertia.

Biology Is the High-Dimensional Prize

Reader Note: This is the clinical and labor-side return of the Future of Science / Clinical AI claim that biology is too high-dimensional for unaided human cognition. Here the point isn't epistemology; it's exposure—why healthcare labor is unusually vulnerable to machine synthesis.

The volume of medical knowledge has grown way beyond the absorptive capacity of unaided human brains. PubMed now contains more than 40 million citations, and NLM production statistics show PubMed adding well over 1.5 million citations annually in recent years.¹²⁵ Biomedical science increasingly looks a lot less like a pure discovery problem and more like an information-organization problem. Humans cope through specialization, orthodoxy, and process. We look at the body through a straw and then build a trillion-dollar coordination apparatus to pretend we're seeing the whole. I offered a version (at unkind length) of this argument in the Generative Epistemology chapter, but it matters here because our healthcare system is the applied surface of biology's complexity.

Recent developments make the returns-to-intelligence thesis a lot less speculative. As I've talked about ad nauseam in this essay, AlphaFold's movement from protein structure prediction toward broader biomolecular interactions wasn't a literature-review trick: DeepMind released AlphaFold 2 predictions for more than 230 million protein structures, a scientific leap that would have taken the equivalent of hundreds of millions of years to solve experimentally by conventional methods if pursued structure by structure. AI-generated molecules moving into human trials aren't just a new UI/UX for chemists. Regulatory agencies beginning to grapple with AI-supported evidence aren't doing so because the technology is only decorative. These systems are starting to *intervene directly in the causal substrate of biology*: folding, binding, molecular design, diagnostics, pathway analysis, trial selection, and eventually the conceptualization of disease itself. If incremental improvements in model reasoning unlock disproportionate progress anywhere, biology is the obvious candidate.

Dario is right: returns to intelligence are disproportionately high in biology. The great biological discontinuities didn't arrive as smooth increments in a nicely drawn-out operating plan. They arrived as moments of *punctuated equilibrium*—X-ray crystallography, monoclonal antibodies, genomic sequencing, CRISPR, CAR-T, mRNA platforms. A few discoveries per decade can completely rearrange entire clinical, epistemic and economic landscapes. If AI meaningfully accelerates hypothesis generation (and to be clear, we're not quite there yet, as I outline in other chapters), target identification, protein and molecular design, trial selection, or diagnostic pattern recognition, then healthcare's exposure penetrates into the scientific substrate of medicine itself.

This is why healthcare sits at such a strange and wonderful intersection. It has maximum labor intensity, maximum administrative drag, maximum unstructured-data density, and maximum scientific dimensionality. Most sectors have maybe one or two of these characteristics. Healthcare has *all* of them, layered on top of one another, defended by regulation and guild power, and financed through a system of subsidies, cross-subsidies, premiums, public budgets, employer contributions, and household pain.

That's why no major sector has more surface-area exposure to GenAI.

¹²⁵ National Library of Medicine, "PubMed" accessed June 2026 (Link)

Labor-Expensive, Not Merely Labor-Intensive

There's a further point that deserves to be admitted plainly: U.S. healthcare isn't just labor-intensive; it's labor-expensive. Compared with OECD peers, the United States pays many categories of health-care labor more—much more—especially at the apex of the clinical hierarchy. OECD comparisons show specialist doctors earning several multiples of the average worker's wage across wealthy countries, and U.S. physicians sit at the expensive end of that distribution.¹²⁶ Within the U.S., the top ten highest-compensated occupations are physicians or physician-adjacent specialties.¹²⁷ Again, this is in no way a moral indictment of doctors. It's simply a structural observation. A sector with subsidized demand, constrained supply, guild-protected entry, heavy regulation, and high labor intensity becomes inflationary almost by construction.

The wage story *below* the physician apex matters too. Gottlieb, Rinz, Mahoney, and Udalova show that healthcare-worker earnings grew nearly twice as fast as non-healthcare earnings from 1980 to 2022, with the strongest gains concentrated in the middle and upper-middle of the clinical distribution, especially nurses and midlevels.¹²⁸ That wage growth helped build a modern middle class. It also locked a large, politically powerful workforce into a cost structure households and employers increasingly can't bear. That's the moral complexity at the heart of this whole chapter.

Baumol's cost disease sits underneath all of this like a trapdoor. A string quartet still takes four musicians roughly the same amount of time. A 15-minute visit still takes, maddeningly enough, about 15 minutes. When productivity rises elsewhere, *healthcare wages must keep pace* to attract labor even if productivity inside healthcare doesn't. So healthcare becomes more expensive relative to everything else. Globalization made goods cheaper. Mechanization made goods cheaper. Automation made many services cheaper. But healthcare, housing, and education kept rising because they were *demand-subsidized, supply-constrained, and politically protected*. We constrained supply, subsidized demand to relieve the pain, and then marveled when prices exploded.

This is more than an actuarial curiosity. The family health-plan premium crossing into mortgage-like territory for middle-income households is a competitiveness problem, a wage problem, a national-security problem, and a political problem. Employers carry the burden. Workers see suppressed cash compensation. Families accumulate mounting medical debt. Public budgets contort themselves around Medicaid, Medicare, ACA subsidies, provider taxes, supplemental payments, 340B, and every other obscure financing scaffold we've built to make the system look less unaffordable than it actually is. The labor denominator is a hospital P&L issue, yes, but it's also one of the hidden mechanisms by which healthcare taxes the entire economy.

So yes, labor substitution will be painful. But the current equilibrium isn't without its own costs. A system that preserves every incumbent labor category while patients carry medical debt, defer care, ration insulin, skip therapy, or avoid the doctor because the deductible is intimidating is difficult to regard as an unqualified success. The challenge is that healthcare serves multiple constituencies whose interests don't always align well. We should care deeply about the workers whose livelihoods may be disrupted, just as we should care about the patients who struggle to access affordable care today. The task isn't to privilege one group over the other, but to navigate a transition in which patients gain greater affordability and access while workers are treated with honesty, generosity, and a credible pathway toward new forms of opportunity.

¹²⁶ OECD: "Health at a Glance 2025" published on November 13, 2025 ([Link](#))

¹²⁷ U.S. Bureau of Labor Statistics: "Highest Paying Occupations" updated on August 28, 2025 ([Link](#))

¹²⁸ Joshua D. Gottlieb, Kevin Rinz, Lisa B. Mahoney, and Victoria Udalova: "The Rise of Healthcare Jobs" published in March 2025 ([Link](#))

Remote Work as the Automation Tell

One further feature of healthcare exposure is worth isolating because COVID accidentally revealed it: remote work was a *natural experiment in decontextualization*. Before the pandemic, remote work in healthcare and social assistance was relatively limited. Then, almost overnight (I still remember exactly where I was on March 12, 2020), large parts of the administrative superstructure of healthcare loosened from the physical organism of the enterprise. Billing, coding, scheduling, medical-record work, health-plan liaison work, IT, analytics, credentialing, HR, revenue-cycle functions—an avalanche of work once presumed to belong inside hospital walls could be done somewhere else.

That was useful, and for many workers humane. It also revealed something strategically uncomfortable, and very relevant to this moment. If a job can be done remotely, it can often be *decomposed* into discrete workflows. If it can be decomposed into discrete workflows, it can often be *outsourced*. If it can be outsourced, it can often be *automated*. Remote work is therefore something of a diagnostic test for automation exposure. It tells you the work has already been digitized, modularized, measured, and decontextualized, which is exactly the kind of work agentic systems are built to attack.

I'm not saying every remote task vanishes tomorrow. I'm saying remote capability is a *tell*. COVID gave us the natural experiment no one asked for. Pre-pandemic, remote work in healthcare and social assistance was modest; by 2021 it had doubled or tripled depending on the subcategory, and the administrative residue endured.¹²⁹ Billing, coding, appointment scheduling, medical-record abstraction, health-plan liaison work, IT, HR, and analytics loosened from the building. One can celebrate the flexibility and still notice the signal: if work can be done anywhere, through a queue, behind a computer screen, governed by rules, and measured by throughput or accuracy, then it has already been partially decontextualized. Agentification is the next question.

The human part of work increasingly needs to become more human: social, embodied, trust-building, institutionally aware, hard to specify, hard to verify, hard to reduce to a ticket or queue. The tasks that are purely behind a screen will have a harder time defending themselves as inference costs fall and agents become more capable. It's an exposure map for the 150 as we go forward.

Four Healthcare AI Impact Zones

Before I close the exposure stack and move to *mechanism*, let me go ahead and name the four healthcare impact zones that are the subterranean plates holding up the argument. The first is *administrative simplification*: documentation, coding, authorization, claims, scheduling, call handling, credentialing, reporting, and the other high-volume information flows that can make care feel like a clerical siege. This is where the labor savings arrive first because the work is digitized, language-mediated, and increasingly verifiable. It's also where the persuasive case for automation is unusually strong. Nobody's deepest conception of human flourishing requires another human being to spend her life reconciling a denial queue created by an adversarial payment architecture.

The second zone is *care-model expansion* and *care augmentation*. If the marginal cost of expertise, outreach, monitoring, and patient education falls, healthcare can become more continuous without becoming proportionately more staffed. That is an excitingly useful decoupling. The point isn't merely to answer messages faster or summarize notes more elegantly. *It's to create care that was*

¹²⁹ U.S. Bureau of Labor Statistics: "Telework during the COVID-19 pandemic: estimates using the 2021 Business Response Survey" published in March 2022 ([Link](#))

previously too expensive to deliver: longitudinal behavioral support, medication-adherence nudges that aren't officious and annoying, chronic-disease monitoring that actually notices deterioration and decompensation before it happens, proactive gap closure, better post-discharge follow-up, and navigation that doesn't require a patient to be her own case manager. This is the *denominator-expansion* side of the thesis. Some of the savings from administrative compression should become more care, not simply more margin.

The third zone is *clinical-adjacent synthesis*. This isn't fully autonomous medicine, and it shouldn't be casually marketed as such. It's the application of machine cognition to the parts of clinical work that are document-heavy, pattern-rich, guideline-mediated, and capable of expert review: medication reconciliation, differential expansion, second opinions, risk stratification, care-gap identification, trial matching, discharge planning, inbox triage, and the retrieval of institutional knowledge that no human clinician can keep in working memory. This is where the profession will argue hardest, partly for justified safety reasons and partly because the bundle called "doctor" is undergoing the same gentle unbundling that AI is bringing to many expert domains: not eliminating expertise, but making its component parts more explicit, more measurable, and more susceptible to automation. I'll be a little (well, a lot) bolder in my predictions and prescriptions on this topic in my Doctor Disequilibrium chapter, and explore the fuller implications of the emergent '*medical superintelligence*,' but for now I'll hold fire and just describe the more limited version of this category here.

The fourth zone is *biological discovery*. This is the least immediate as a labor-compression story and the most important as a civilization story. AI won't only help healthcare run its existing workflows more cheaply. It will also accelerate the *discovery layer* upstream of care: proteins, molecules, diagnostics, pathways, trial design, patient selection, and the conceptualization of disease itself. That doesn't replace the operating argument of this chapter. It just raises the stakes. The same technology that compresses administrative cognition may also *multiply biomedical cognition*. Healthcare is exposed not only because its back office is stultifying and complex, but because biology itself is a domain where returns to intelligence may be fantastically high.

These four zones should frame the 150's thinking. If a proposed AI initiative doesn't simplify administration, augment clinical judgement, improve clinical-adjacent synthesis, or accelerate biological discovery, it may still be useful, but it's probably not strategic. The risk in healthcare is that we will mistake interface novelty for production-function change. A prettier portal, a nicer chatbot, or a slightly less cumbersome documentation tool may be welcome. But the real question is whether the technology changes the labor denominator, the care denominator, the clinical synthesis layer, or the discovery frontier. That's where the Oppenheimer moment lives.

Addition by Subtraction

This is the healthcare addition-by-subtraction thesis in a few sentences. Less human labor in the wrong places may mean more care in the right places. Less administrivia may mean more access. Less documentation may mean more doctoring. Less prior-authorization theater may mean more care navigation. Less coding warfare may mean lower unit cost. The moral project isn't to preserve every job exactly as it exists. The moral project is to build a system in which human beings released from bad work are treated decently, and human beings needing care receive more of it.

Healthcare's weaknesses are also the reason the opportunity is so large. Labor addiction creates substitutability. Technology debt creates leapfrog potential. Data debris creates a model substrate. Biological complexity creates returns to intelligence. Unaffordability creates the political and economic permission structure for substitution once the technology is good enough. That's a lot of

converging force in one sector.

The strange inversion is that healthcare's backwardness becomes its exposure. And perhaps this is, on net, a good thing. The sector's failure to metabolize internet, mobile, cloud, SaaS, analytics, and interoperability didn't save it from technological change. It just stored up dry powder for a larger ignition. The same messiness that defeated brittle rules-based automation is precisely the messiness language models can now metabolize. The old excuse—our workflows are too idiosyncratic—gets less convincing by the quarter.

So as leaders we should stop taking comfort in healthcare exceptionalism. Exceptionalism bought time; it didn't buy immunity. The legal, regulatory, clinical, and cultural ramparts still matter, and they will slow the blast. But delay mechanisms aren't force fields. The 150 should use the delay to prepare rather than mistaking it for protection.

Exposure tells us where the blast radius is largest. Mechanism tells us how the blast travels. And the propagation won't be mystical. It will move through *functional verifiability*, *decontextualization*, *teachability*, *tolerance for error*, *agentic execution*, and *bottleneck migration*. Those are the rails. Now on to the implications.

Part IV—Healthcare as Civic Labor Architecture

Now the boundary of the argument expands beyond the enterprise. Healthcare labor isn't only a health-system cost input; it's *regional infrastructure*. This part explains why AI-enabled healthcare labor compression won't remain respectfully inside the board deck. It will travel into payrolls, tax bases, housing markets, restaurants, community colleges, unions, and local politics.

The argument is straightforward. Healthcare quietly became one of the great labor absorbers of the post-manufacturing American economy. In many regions, the *hospital replaced the factory* as the central employment institution. That made healthcare unaffordable in one direction and civically indispensable in another. So when GenAI begins compressing healthcare labor, the consequences won't be confined to staffing ratios, SG&A, revenue-cycle teams, or operating margins. They'll propagate through communities. Healthcare labor strategy, in other words, becomes civic labor architecture.

There and Back Again—From Manufacturing to Healthcare, and Perhaps Back to Manufacturing

Once the argument shifts from hospitals to the economy built around hospital labor, the scale of the problem changes. This is no longer only a story about workflow redesign, or even about healthcare labor narrowly understood. It's a story about the structural buttresses of the American economy: where *middle-class work* has gone, which institutions now anchor local communities, and what happens when the sector that quietly replaced manufacturing as an employment absorber begins to automate.

I hope, dear reader, you appreciate my *Hobbit* reference in this subchapter's title (couldn't quite help myself). But this is less an odyssey of a brave, unlikely hero who returns home sanctified and redeemed, and more the story of a massive impending convulsion in the labor foundations of the American middle class. A quick history lesson is in order, not because I want to indulge yet another historical digression for its own sake, but because the manufacturing-to-healthcare migration is one of those structural metamorphoses that has been happening in plain sight while almost everyone misread the plot.

This has become especially pronounced recently, with the growing recognition that healthcare employment has been doing a startling amount of labor-market stabilization on its own. As noted earlier, in the year through March 2026, health care and social assistance added roughly 680,500 jobs even as total nonfarm payroll employment barely budged. And in some recent monthly decompositions, healthcare has accounted for almost all private-sector job creation. It's shocking, or maybe not shocking but still revealing, how few economists, policymakers, or Fed-watchers seem to be making the neural connection. Healthcare employment has become a macroeconomic stabilizer. And if that stabilizer begins to stall, automate, or reorganize, the spillovers won't obediently remain inside hospital walls.

At the same time, another labor engine was rising. Manufacturing jobs evaporated over decades, and what, in substantial part, took their place? *Healthcare*. Not perfectly, not one-for-one, not with the same gender composition or union structure or cultural meaning, but functionally. *Healthcare became the great employment absorber of the late-stage, deindustrialized American economy*. It became the thing that could keep hiring through recessions, wars, pandemics, policy chaos, and all the usual macroeconomic vicissitudes. It became the sector parents told their children to enter because there would always be work. And for a long time, they were right.

This has become especially pronounced recently, with the growing recognition that healthcare employment has been doing all the heavy lifting of labor-market stabilization on its own. I'll repeat the stats again: in the previous year through March 2026, health care and social assistance added nearly 700,000 jobs even as total nonfarm payroll employment barely budged. And in some recent monthly decompositions, healthcare has accounted for almost *all* private-sector job creation. It's shocking, or maybe not shocking but still revealing, how few economists, policymakers, or Fed-watchers seem to be making the neural connection. Healthcare employment has become the *macroeconomic stabilizer*. And if that stabilizer begins to stall, automate, or reorganize, the spillovers won't obediently remain inside hospital walls.

The Great Inversion: Manufacturing Down, Healthcare Up

Let's start with the big tectonic plates. NYT coverage on healthcare and manufacturing, combined with a recent Gottlieb/Rinz/Mahoney/Udalova paper, gives the cleanest version of the inversion: healthcare employment grew from 9.3 million in 1990 to 18.1 million in 2022. Healthcare overtook manufacturing in 2006 and retail trade in 2009 to become the largest U.S. industry by employment. By the mid-2020s, the broader healthcare-and-social-assistance perimeter has pushed out closer to 24 million jobs, compared with roughly 12.7 million in manufacturing and roughly 15.6 million in retail. The basic inversion is the point. Healthcare relentlessly increased its share of the labor force while manufacturing receded.

This wasn't an accident. It was the predictable, foreseeable labor-market consequence of a rich, aging, chronically ill society whose most expensive sector remained comparatively immune to the deflationary technologies that transformed goods, logistics, finance, retail, and much of the rest of the economy. *As goods and services got cheaper, healthcare got more expensive*. As manufacturing required fewer humans, healthcare required more. As the old industrial middle class weakened, healthcare became one of the few places where non-advanced-degree and middle-skill workers could still find reasonably stable wages, benefits, and social standing.

And the wage story definitely matters. As employment in healthcare rose, so too did wages. Earnings for healthcare workers grew nearly *twice as fast* as earnings in other industries from 1980 to 2022, with average healthcare earnings moving from 4% *below* non-healthcare earnings to more than 14%

above them. The fastest gains weren't only at the pinnacle, where guild protections are most obvious, but among the middle and upper-middle layers of clinical labor: nurses, PAs, NPs, therapists, and skilled support roles. Some of this is straightforward economics: demand rose, supply was constrained, and the work couldn't easily be offshored or mechanized. Some of it is ordination and credentialing. The result was a modern middle-class jobs engine built inside the most expensive sector of the economy.

That's why healthcare became, in effect, the *post-manufacturing bulwark* of the American middle class. Not the whole middle class, obviously, and not always as generously or durably as the old industrial bargain. But enough to matter. Enough to stabilize communities. Enough to anchor labor markets. Enough that, in a great many places, the hospital replaced the factory as the central institution around which employment, consumption, philanthropy, local politics, and civic identity organized themselves.

Why Healthcare Kept Growing—and Growing

The demand side matters just as much as the labor side. Three forces did the work. First, *coverage expanded*: the uninsured rate (mercifully) fell sharply over the long arc, and the Affordable Care Act increased access to insurance and therefore demand for care. Second, *chronic disease and aging* did the rest. We're older, frailer, more cardiometabolic, more oncologic, more diabetic, more anxious, more medically surveilled, and more therapeutically reachable than prior generations. Third, *other stuff got cheaper*. Globalization, mechanization, and automation lowered the relative cost of food, clothing, electronics, household goods, and much of the material world. A rich, aging democracy that spends less of its disposable income on goods will spend more on services, and especially on healthcare.

But the deflationary forces that disciplined goods didn't discipline healthcare in the same way. Why? You can offshore a supply chain. You can mechanize a factory. You can automate inventory management. You can squeeze logistics. But when you need surgery, a nurse, a diagnosis, a therapist, a home-health aide, or a hospital bed, the work remains stubbornly embodied, regulated, credentialed, local, and human. Yes, da Vinci helps. No, as of 2026, Optimus isn't doing your colectomy. That will come later (but probably not as late as you might think—more on that later). The bottom line is across the 20th and 21st centuries the trend is clear: richer, older, sicker societies spend proportionately more on healthcare, and the United States spends in a class by itself.

Baumol (our old friend) and his cost disease sit right in the middle of this. Healthcare wages have no choice but to keep pace with the rest of the economy in order to attract and retain labor, even when productivity inside healthcare remains perversely low. To invoke my earlier examples: the string quartet still takes four musicians. The 15-minute visit still takes 15 minutes, and then, because this is American healthcare, another 20 minutes of documentation, coding, billing logic, payer friction, and portal-management misery get layered on top. The result is relentless wage and cost inflation in a productivity-stagnant vertical.

Hospitals as Regional Anchors

A further concomitant is that healthcare isn't just a large employer nationally. It's the *anchor institution* in thousands of local economies. In many states and metropolitan areas, healthcare is the largest or among the largest employment sectors. In many smaller communities, the hospital isn't merely a place where people go when they're sick. It's the payroll engine, the philanthropic center, the civic sponsor, the professional-class employer, the political actor, the stabilizer of commercial real estate, and sometimes the last serious institution still standing after manufacturing, local media, regional banking, and retail have all consolidated or shuttered.

Think about **Cleveland, Pittsburgh, and Rochester**. Each is now synonymous, in one way or another, with a healthcare institution or ecosystem. Each also has a *deep industrial or manufacturing inheritance*. That symmetry matters. The places that once made steel, machinery, chemicals, glass, or industrial goods now often make care, diagnosis, research, medical prestige, and institutional employment. Boston is more complicated, of course, because Boston never reduced itself to a single industrial identity, but even there the lineage is suggestive: Waltham and the Boston Manufacturing Company belong to the early industrial story, just as Boston's academic medical centers and life-sciences ecosystem belong to the healthcare-industrial story. I find the parallels here very interesting.

Beyond these luminary cases, the general point is more important. Hospitals and health systems are, in many communities, more likely than almost any other institution to be the largest employer and the economic anchor. The American Hospital Association estimates that in 2023 U.S. hospitals directly employed 6.6 million people, purchased more than \$1.3 trillion in goods and services, fueled \$4.8 trillion in total economic activity, and supported nearly one in six jobs across the country.¹³⁰ One can discount the advocacy gloss and still absorb the directional truth: a hospital payroll is also procurement, consumption, philanthropy, housing demand, tax base, and civic ballast.

That's the hidden vulnerability. If a factory closes, everyone understands that the town is in trouble. If a hospital trims administrative labor by 12 percent over five years through attrition, automation, outsourcing reversal, and management-layer compression, the effect will look less dramatic at first. No single smokestack goes dark. No gate is chained shut. No iconic photograph appears in the local paper. But the payroll shrinks. Consumption softens. The tax base notices. The restaurants notice. The housing market notices. The local college graduate who expected a safe administrative job notices. The city council notices, eventually. The labor politics then follow.

That's why the 150 don't think of themselves as ordinary employers. In many metros and states, healthcare is either the largest employer or close enough to the top that the distinction barely matters. The hospital is payroll, procurement, philanthropy, prestige, and political ballast. When manufacturing contracted, the effect was visible because the factory gate closed. Healthcare contraction will be quieter at first, but no less real. The payroll architecture of the region changes one unfilled requisition, one automated queue, one compressed management layer at a time.

¹³⁰ American Hospital Association: "Hospitals Are Cornerstones in Their Communities Infographic" updated in February 2025 (Link)

Healthcare as the Hidden Jobs Program

This is the uncomfortable part, and we should say it out loud. Healthcare has become, in part, *the hidden jobs program of a deindustrialized republic*. I don't mean that cynically. These are real jobs, held by real people, many of whom do necessary and admirable work under difficult conditions. And I certainly don't mean that healthcare workers are make-work functionaries. Exactly the opposite, in fact: bedside work is sacred, much of the operational work is necessary, and the system would collapse tomorrow without the people who keep it moving through its own impossible complexity.

But structurally, we've allowed healthcare to absorb labor that the rest of the economy no longer needed, while financing that absorption through premiums, taxes, debt, employer contributions, suppressed wages, and the slow immiseration of household balance sheets. We lost manufacturing capacity, failed to build a coherent national industrial strategy, allowed healthcare to become more expensive than any other system in the world, and then consoled ourselves that at least the hospitals kept hiring.

This is why the labor-substitution question in healthcare is so explosive. Jobs may disappear or reorganize in the *very sector that quietly replaced manufacturing as the absorber of mass employment*. We're contemplating the automation of the post-manufacturing labor engine. That's a national adjustment problem hiding inside an industry strategy.

And this explains why the "healthcare shortage" narrative has been so seductive. Shortages reassure everyone. They tell workers they're safe, unions they have leverage, employers that wages are simply the cost of doing business, policymakers that more training programs are needed, and communities that the hospital will remain the reliable employer of last resort. But a shortage can become an artifact of the old production model. If output has historically scaled by adding humans, then every demand forecast produces a shortage forecast. The real question is whether output must continue scaling that way.

AI says maybe not.

(Perhaps) Back to Manufacturing

This brings us to the "back again" part of the title. I don't mean, obviously, that the United States is going to empty its hospitals and send everyone back to a factory floor imminently. Nor do I mean that every displaced revenue-cycle worker becomes a welder, every scheduler becomes a semiconductor technician, or every care manager gets retrained to install transformers at a data center. That would be unserious, and also a little insulting. Labor transitions are hard, path-dependent, local, age-dependent, credential-dependent, and culturally sticky. People aren't widgets waiting to be reallocated by a benevolent social planner with a spreadsheet and clipboard.

But I do think AI may force a new allocation of human work back toward production, infrastructure, logistics, energy, advanced manufacturing, and embodied care. The United States has rediscovered, somewhat belatedly, that a country can't run as an autarkic, standalone system on services, finance, software, healthcare, and asset inflation. We need energy infrastructure, grid modernization, fabs, robotics, defense industrial capacity, housing, logistics, biomanufacturing, and the physical substrate of the AI economy itself. The data-center boom—and the attendant pitchfork protest!—is one visible sign of this, though we should be careful not to confuse a construction surge with a permanent labor absorber. Still, the direction is clear: the world of atoms is back, baby. Big time.

There's a deep irony here. Healthcare grew because other sectors became more productive, more

automated, more globally integrated, and less labor-intensive. Now AI may do to healthcare what prior technologies did elsewhere, releasing some labor from the administrative and coordinative apparatus of care just as the country needs more capacity in energy, manufacturing, logistics, infrastructure, and high-touch human services. The path won't be smooth, the math between industries won't be perfectly symmetrical, and the workers won't swap cleanly from one sector to another. But the macro logic isn't totally implausible. A country that has allowed healthcare to become a quasi-jobs program will eventually have to admit that care delivery isn't supposed to absorb labor indefinitely as a substitute for national industrial strategy. Healthcare should heal people. It shouldn't be the hidden unemployment-insurance system of a deindustrialized republic forever.

And that means the 150 will think about their labor strategy with more civic seriousness than the average management team. If healthcare labor contracts, the goal can't simply be lower SWB and higher EBITDA. That's too narrow, too unemotional, too spiritually unserious for an industry that has become so central to the country's employment architecture. The goal has to be a humane reallocation: remove humans from bad work, preserve and expand human presence where it matters, retrain where retraining is real, provide generous transition where it's not, and participate in a broader national conversation about what productive work should become in a post-AI economy.

This isn't nostalgia for manufacturing. I'm not interested in time-traveling back to 1955. The old factory economy had plenty of brutality, pollution, hierarchy, exclusion, and boredom. But it did understand something healthcare now has to relearn: an economy needs productive, dignity-conferring work that's not merely administrative intermediation. It needs people building, making, repairing, installing, caring, teaching, persuading, leading, and serving in ways that can't be reduced to a claim, a code, a note, a queue, or a denial letter.

That's why this chapter belongs here, after the deep-vulnerability argument and before the multiplier effect. Healthcare is exposed not only because its own cost structure is labor-heavy, but because the national economy has come to depend on that labor heaviness. We let healthcare become the circulatory system of the American middle class after manufacturing lost its role as the muscle. And now we're contemplating what happens if that circulatory system itself begins to automate.

The answer won't be tidy. It never is. But the first act of seriousness is to see the inversion clearly. Manufacturing gave way to healthcare as the great labor absorber. AI may now force healthcare to give some of that labor back—to production, to infrastructure, to embodied care, to new bottlenecks we don't yet fully see. The 150 can pretend this is merely an internal workforce issue, or they can recognize it for what it is: a civic transition with healthcare at the center.

The Multiplier Effect

Let's broaden the aperture just a bit further. Healthcare staffing retrenchment isn't merely a workforce problem; it's a regional economic-structure problem. That distinction matters because the moment we move from "which jobs can AI substitute?" to "what happens when the largest employer in a community begins needing fewer humans?" we're no longer inside the familiar domain of workforce planning. We're now asking some political economy and sociological questions.

This is the next implication of the manufacturing-to-healthcare story. If healthcare quietly became the employment absorber of the deindustrialized American economy, then labor substitution in healthcare can't be treated as a hospital operating initiative with some unfortunate HR consequences attached. It's something larger and stranger: the possible automation of the institution that replaced the factory as the stabilizing payroll base in much of the country. And once we see it that

way, the blast radius changes.

We need to think about this expansively, because the math isn't going to remain obediently inside the hospital walls. If the core supposition of this chapter is directionally right—that GenAI will enable large-scale substitution of machine intelligence for human labor in healthcare, tempered of course by unmet demand, geographic maldistribution, Jevons' paradox, the lump-of-labor fallacy, and all the other confounding variables economists quite properly throw at sweeping claims—then the consequences will reverberate far beyond the direct healthcare worker. The direct job is only the *first-order effect*. The deeper question is what happens to the supply-chain jobs, the consumer-spending jobs, the tax base, the municipal budget, the retail corridor, the real-estate market, and the politics of a community when its anchor institution starts to operate with fewer people.

That brings us to the *healthcare employment multiplier*. The question is simple enough: when a hospital job disappears, what else disappears with it? Not necessarily immediately, not mechanically, not in some reductive one-for-one cascade, but through the slow metabolic pathways by which a large payroll becomes local spending, local purchasing, local tax revenue, local philanthropy, local professional formation, and local civic stability.

With roughly 24 million Americans directly employed across healthcare and social assistance, and roughly 6.6 million in hospitals alone (we'll stick with the old number for this math exercise), we're talking about a sector whose employment footprint is already planetary. So imagine, not apocalyptically but arithmetically, a 10% reduction in some subset of healthcare labor over several years. Are the consequences confined to that unfortunate 10%? Of course not. There's a phalanx of adjacent work quietly dependent on the direct healthcare payroll: suppliers, food service, facilities, local contractors, professional services, transportation, retail, housing, childcare, and the ordinary consumer economy supported by relatively stable healthcare wages.

The AHA's own economic-impact work is useful here, even allowing for the obvious caveat that the AHA has a horse in this race and isn't exactly a disinterested monk copying numbers by candlelight. Its recent website material argues that hospitals directly employ 6.6 million people and support nearly one in six jobs nationally; older economic-impact work using BEA multipliers found that **each hospital job supports about two additional jobs** and that every hospital dollar spent generates more than two dollars of additional business activity. Multipliers are sensitive to geography, wage mix, leakage, academic-medical-center density, and definitional choices. But the directional point isn't fragile: a sector this large, this labor-intensive, and this locally embedded can't contract its labor base without consequences spilling into the broader economy.

Pleasantville and the False Comfort of Direct Jobs

A hypothetical helps. My colleague David Elop ran an analysis in Claude—our new unpaid, or at least underpaid, junior economist (Claude not David!)—to model a city of 100,000 workers. I'll call it Pleasantville, to nostalgically invoke an old Advisory Board chestnut. If one out of every six or seven workers is directly employed in healthcare, then roughly 16,700 people in Pleasantville work somewhere inside or adjacent to the care-delivery apparatus: nurses, techs, administrators, billing staff, physician-office workers, care managers, imaging personnel, lab staff, revenue-cycle teams, therapists, and the rest of the local healthcare organism.

But that's not the full dependency. Hospitals and health systems purchase from other businesses, which creates indirect employment. Healthcare workers spend their wages locally, which creates *induced employment*. Those categories can sound bloodless, like the sort of thing one says in a

regional economic-development report before the reader nods off to sleep, but they're quite concrete. The indirect jobs are the vendor, supplier, maintenance, construction, food, equipment, consulting, laundry, and professional-service jobs connected to the health system's buying power. The induced jobs are the restaurants, grocery stores, landlords, child-care providers, car dealerships, hair salons, Uber drivers, and local retailers supported by healthcare workers' paychecks.

Once those effects are included, the healthcare footprint in Pleasantville is no longer one-sixth of employment. Depending on the multiplier, leakage, and local wage structure, it can become something closer to *one-third*, perhaps even *two-fifths*, of the local employment ecosystem being economically dependent, in one way or another, on healthcare. Pedants will correctly note that smaller regions leak more spending, that large academic medical centers have larger induced effects because of wage levels, and that a hospital job in Rochester isn't the same as a home-health job in rural Alabama. Fine. All true. But the adjustments and exceptions shouldn't obscure the plain fact: healthcare labor isn't just labor. It's regional infrastructure.

This is why the direct-job frame is inadequate. A hospital CFO may look at an AI-enabled reduction in administrative headcount and see a labor-productivity opportunity. The mayor may see sales-tax weakness six months later. The landlord sees weaker leasing. The restaurant owner sees fewer lunches. The community college sees fewer students entering the old "safe" healthcare tracks. The union sees existential threat. The state legislator sees a constituency with a grievance. The health system sees margin relief; the region sees a payroll shock. Both are true. The problem is that only one of those truths usually makes it into the board deck.

Seattle is oddly instructive because it shows how quickly employment contraction in a dominant sector leaks into everything else. I used to think of Seattle as a vibrant, tech-saturated, full-employment metropolis. But when a city's largest private employers—the titans of Microsoft and Amazon, in this case—slow hiring or reduce staff while market capitalizations keep going parabolic, the effect isn't confined to severance packages and LinkedIn posts about "new chapters." Payroll taxes weaken. Sales taxes weaken. Restaurants feel it. Retail feels it. Moving companies feel it. Real estate inventory rises. Homes sit longer. Commercial vacancies become harder to disguise. The multiplier isn't an academic artifact; it shows up as the lived ecosystem around a payroll base.

Healthcare cities should study that pattern carefully. Cleveland, Pittsburgh, Rochester, Baltimore, Nashville, St. Louis, Jacksonville, and a great many less-famous hospital towns all have some version of this exposure. In many places, the health system isn't merely an employer. It's the anchor institution, the civic brand, the economic-development engine, the downtown stabilizer, the apprenticeship system, the philanthropic backbone, and the largest buyer of local services. Labor contraction in such an institution becomes indistinguishable from regional macroeconomics.

That's why the 150 can't treat labor substitution as a narrow operating-margin problem. The direct job goes first. Then the supplier job notices. Then the consumer-spending job notices. Then the tax base notices. Then the politics notices. And once politics notices, the window for clean managerial action starts to close. Regulatory enclosure, union organizing, automation moratoria (hello Bernie), staffing mandates, populist tax proposals, and the whole machinery of backlash begin to arrive. The industry shouldn't act surprised when this happens. A sector that has become the local economic anchor can't quietly shrink its labor denominator and expect the community to applaud the EBITDA improvement.

The Civic Burden of the 150

I'll close this section with a final meditation or two. The sum of all these compounding societal and macroeconomic implications is why the 150 will need to think like civic institutions, not merely operating companies. A large health system that reduces labor intensity without thinking through the regional consequences may improve its margin and damage its community in the same breath. That doesn't mean preserving every redundant role out of sentimental attachment to the old payroll base. It means sequencing the transition, communicating honestly, partnering with local governments and community colleges, redirecting talent where the redirection is real, and offering generous transition support where it's not.

The politics will arrive after the math. They always do. And when they arrive, they won't be polite. A community that sees its hospital system reduce headcount while executives talk about innovation and AI-enabled efficiency won't automatically experience that as enlightened productivity. It may experience it as betrayal, especially if premiums remain high, access remains mediocre, and the savings appear to accrue to the institution rather than patients, workers, employers, or the public. That's how an operating strategy becomes a political problem.

The economic case for labor substitution is strong. The human case is, naturally, a lot more complicated. A serious chapter can't stop at GDP, margins, and operating leverage, because work isn't merely a cost input. It's identity, apprenticeship, routine, dignity, status, friendship, discipline, and sometimes a blessed distraction from oneself. Machines can remove drudgery. They can also remove the ladder by which people become competent, useful, and proud.

That's the moral hazard inside the multiplier effect. Labor reductions can free capital for affordability, access, prevention, better clinical care, and a more humane system. They can also hollow out communities, intensify alienation, and create the political conditions for a ferocious backlash against the very technology that could make healthcare more abundant. The same spreadsheet can be a liberation document or a social accelerant. The difference is whether leaders treat labor as a denominator to be reduced or as human capacity to be reconstituted.

So yes, healthcare staffing retrenchment may become necessary. In some domains, it already is. But the 150 should understand what they're touching. They aren't just changing staffing ratios, closing open requisitions, replacing contractors, or automating work queues. They're adjusting the payroll architecture of regions. They're altering the employment base of communities. They're disturbing the quiet bargain by which healthcare absorbed the labor that manufacturing no longer could. And if they do that without civic seriousness, the multiplier won't just be economic. It will be political, social, and moral.

Part V—How Substitution Propagates: The Rails of Automation

The first four parts answered why healthcare? and why now? This part attempts to answer *how*. I don't expect substitution to arrive as one big cinematic event, some single lightning strike in which half the administrative workforce vanishes and an Elon robot solemnly takes possession of the nurse's station. It'll move through *specific task architectures*, some obvious, some hidden in plain sight, and many already confessed by the way healthcare work is routed, measured, outsourced, documented, appealed, denied, audited, escalated, and converted into a queue. The organization that understands those architectures can sequence the transition with something like moral and operational seriousness. The organization that doesn't may discover substitution as a sequence of vendor pitches, workforce shocks, board surprises, union counteroffensives, and political events it should have seen coming.

The rails aren't mysterious. Substitution propagates through *functional verifiability*, *decontextualization*, *teachability*, *tolerance for error*, *agentic execution*, and *bottleneck migration*. Those are the mechanisms. Not vibes, or futurist fog, or (spare us!) another panel about "innovation" followed by a procurement cycle and a governance committee. *Mechanism*. If the 150 want to understand where the labor shock arrives first, we should stop asking only in occupational categories—nurse, coder, physician, analyst, claims specialist, care manager—and start asking in task-architecture categories. Can the work be verified? Can it be separated from the organism? Can it be taught? Can it be bounded by an acceptable error envelope? Can it be routed to an agent? And what happens to the bottleneck after the first task gets automated?

That's the point of this part. Exposure tells us where the blast radius is largest. Mechanism tells us how the blast travels.

Capability Before Permission

The central theme here is the approaching substitutability of capital for labor. Notice I'm saying *substitutability* rather than *substitution*. The distinction isn't some semantic ornamentation; it matters. The technology may *permit* substitutions that healthcare leaders, clinicians, patients, regulators, unions, malpractice carriers, and society aren't yet comfortable making. Capability and permission certainly aren't the same thing. That gap will slow the transition because healthcare is full of liability anxiety, trust constraints, guild resistance, workflow complexity, and the ordinary human discomfort of being told that a machine can now perform work we'd previously associated with professional identity.

But the directional point is hard to escape. We're entering a period in which capital can increasingly buy compute, compute can increasingly instantiate cognition, and cognition can increasingly perform labor. That's the substitution chain. It won't move everywhere at once. It won't move cleanly. It won't move without backlash, errors, litigation, and a great deal of elevated demagoguery about safety and standards. But economic possibility usually precedes social permission. First the technology makes substitution plausible; then institutions fight, bargain, moralize, regulate, delay, and eventually normalize it. Healthcare leaders therefore need to distinguish the *capability frontier* from the *comfort frontier*. The comfort frontier is what the institution, medical staff, board, patient, regulator, and malpractice carrier are ready, or resigned, to accept. The capability frontier is what the technology can actually do. Healthcare has a long and venerable tradition of mistaking the former for the latter. That may be a dangerous error now.

The velocity of change makes this especially disconcerting. Only a short while ago—worthy of a pause here to reflect that we're only a mere 40-some months since the thunderclap of ChatGPT in November of 2022—large language models were mostly passive systems, waiting docilely for a prompt to stimulate an output. Now we're moving toward agentic and autonomous architectures: persistent workflows, tool use, memory, orchestration, recursive task execution, multi-agent coordination, and systems that don't merely answer questions but prosecute work. GPTs—general-purpose technologies in the classical economic sense, not the oddly eponymous named *generative pre-trained transformers*—have always had this peculiar property: their use cases are both foreseeable and unforeseeable. They generalize. Once they diffuse into an economy, they produce second- and third-order consequences no one fully anticipated.

Healthcare, as we've now established to the point of exhaustion, has unusual susceptibility to this shift. It's enormous, labor-intensive, productivity-challenged, administratively encumbered, data-rich, and full of cognitive workflows never built for a world where cognition itself could be industrialized. Internet, mobile, cloud, enterprise SaaS, and big data swept across manufacturing, finance, logistics, retail, and media while mostly attenuating at the gates of healthcare. That didn't make healthcare immune. It created stored vulnerability.

The Healthcare Substitution Chain: From SWB to kWh

The preceding labor chapter made the macro claim: capital is learning to instantiate labor through compute. I'll only restate the mechanism here because healthcare leaders may need it in their operating vocabulary. Capital buys chips, power, cooling, data centers, models, and inference. Those inputs will increasingly instantiate digital labor. Digital labor substitutes for human labor where the work is verifiable, teachable, decomposable, or tolerably error-bounded. That's the chain. In healthcare, the chain runs straight into the largest input in the system: salaries, wages, benefits, contract labor, outsourced labor, agency labor, management layers, call centers, documentation teams, coding queues, claims operations, and the whole salaried apparatus of human cognition we've built around care.

This is the conversion from SWB to kWh: from salaries, wages, and benefits into kilowatt-hours, chips, inference, cooling, and agentic labor. The phrase sounds dystopian because, well, it is a little dystopian. And of course I don't mean this literally; more as a gentle conceptual provocation. But it's hard to argue against the basic industrial logic of GenAI. Healthcare leaders still think in headcount, FTEs, contract labor, agency spend, staffing ratios, overtime, and retention bonuses. AI-native companies increasingly think in GPU utilization, inference cost, model routing, task completion, agent reliability, and marginal cost of synthetic labor. These are categorically different concepts of production.

The point for this chapter isn't to repeat the broad capital-labor argument. It's to make it operational. If the unit cost of machine cognition falls while the cost of healthcare labor keeps rising, then every healthcare function built on expensive, screen-mediated cognition becomes exposed. Documentation, coding, claims follow-up, utilization review, call handling, scheduling, payment integrity, compliance compilation, routine analytics—all of them start to look less like durable human occupations and more like temporary artifacts of a pre-agentic operating model. That certainly doesn't mean every role disappears, but it does mean that the burden of proof shifts. The 150 therefore need a new managerial question: which workflows still deserve expensive human cognition? Sometimes the answer will be reassuringly obvious: bedside trust, procedural skill, complex judgment, patient persuasion, team leadership, moral accountability. Often the answer will be less

flattering: inertia, guild comfort, vendor lock-in, regulatory drag, or the fact that we haven't yet had the courage or energy to redesign the workflow yet.

Measurement Before Slogans

This is also why the 150 have to get much more serious about measurement. Rigorous, objective, atomic-level measurement. Healthcare has lived for too long inside sloganized abstractions: shortage, burnout, transformation, innovation, access, equity, value, consumerism, digital front door, workforce resilience, and all the other words that sound important right up until they're forced to change an operating model (trust me, this isn't critical, it's confessional... The Advisory Board produced a lot of these slides!). AI will punish that vagueness. It won't be enough to say a function is exposed. It won't be enough to say a role is essential. It won't be enough to say a model is unsafe. The institution will need *evidence*: task-level exposure, workflow-level cost, measured quality, model performance, error envelopes, escalation rates, human time saved, labor actually redeployed, and downstream effects on access, affordability, safety, and experience.

The measurement problem has to be designed before the politics harden. If labor hears only that management is using AI to "drive productivity," labor will hear *layoffs*. If clinicians hear only that AI will "reduce burden," clinicians will suspect surveillance, liability transfer, and an inbox that somehow gets worse. If patients hear only that AI will "improve access," patients will worry that they're being routed to a cheaper substitute for human care. Good measurement is *legitimacy infrastructure*. It lets leaders say: here is the work AI does better, here is the work humans still do better, here is where the model failed, here is where the human failed, here is what changed for patients, here is what changed for workers, and here is where the savings went.

That's why I keep returning to the notion of a healthcare GDPval. It isn't a research hobby, or an awkwardly ill-fitting translation from Silicon Valley. I think of it as a *map* that lets institutions avoid both panic and denial. Without the map, every conversation devolves into vibes: the vendor says the model is magical, the guild says the model is unsafe, the CFO says labor is too expensive, the union says workers are being discarded, and the patient wonders why she still can't get an appointment. With the map, the conversation becomes more constructive.

Functional Verifiability: The First Wedge

We need a rigorous methodology for sequencing automation, but (at least) the first principle is straightforward. The two most important words in AI labor substitution may very well be *functional verifiability*. Old software automated what could be *specified*: steps, rules, fields, if-then logic. AI automates what can be *verified*: right answers, better answers, scored outputs, reward functions, appeal outcomes, payment outcomes, clinical quality checks, expert comparison. The system no longer needs every rule specified in advance. It needs a way to know whether the output worked. That's the key evolutionary step.

That's why coding went first. Code, in short, has an unusually clean verification structure. Does it compile? Does the test pass? Does the system behave as intended? Can the agent try again? Can the reward signal improve the next attempt? That same logic applies, less perfectly but still powerfully, to many healthcare workflows: claims, coding, payment integrity, denial management, authorization, documentation sufficiency, scheduling optimization, eligibility, quality reporting, and financial analysis. The jagged frontier isn't random. Math, code, puzzles, and claims logic move quickly because the system can practice against feedback. Strategy, taste, bedside ambiguity, and institutional legitimacy move more slowly because verification is harder, noisier, and socially embedded.

Ok, bear with me, because I might be reaching a bit here. But I think the best public exposure framework that may guide this effort is the OpenAI/OpenResearch/UPenn O*NET paper from 2023, which found that roughly 80% of the U.S. workforce could have at least 10% of tasks affected by GPT-class systems, and roughly 19% could have at least 50% of tasks affected. With LLM-powered software layered on top, the authors estimated that 47% to 56% of all U.S. worker tasks could be completed significantly faster at the same quality.¹³¹ That was before mature agents, before today's frontier-model acceleration, before model routing, before the next wave of tool use, memory, and long-horizon workflows. A static 2023 snapshot was already enough to show broad exposure.

Healthcare now needs its own GDPval. OpenAI's GDPval prototype points in the right direction: real-world economically valuable tasks across 44 occupations and nine major GDP sectors, built from actual work products and judged against experienced professionals.¹³² Healthcare needs the same discipline, but with healthcare artifacts: notes, claims, prior-auth packets, denial letters, call transcripts, pathology images, radiology studies, care plans, quality reports, staffing schedules, payer policies, and revenue-cycle queues. Not demos. Not anecdotes. Work.

The healthcare GDPval should classify work into four categories: *automatable*, *augmentable*, *eliminable*, and *newly generative*. Automatable means the machine can do the task with acceptable oversight. Augmentable means the human remains primary but becomes materially more productive. Eliminable means the upstream workflow should disappear because AI changes the process itself. Newly generative means the technology creates a care or administrative capability that didn't previously exist at viable cost. That fourth category matters most. If AI only lets us perform old drudgeries faster, it becomes just more (automated) administrative quagmire. But if it reveals new care models that were previously impossible because every incremental touch required another expensive human minute, then it becomes a path toward abundance.

So the first wedge is *functional verifiability*. Invoking the godfather of code Andrej Karpathy (now of Anthropic!): Software 1.0 automated what could be specified. AI automates what can be verified. Healthcare is full of verifiable work. That's the uncomfortable thing.

Revenue Cycle as the Canonical Battlefield

Revenue cycle management now therefore enters center stage. It's the canonical healthcare battlefield because objective functions are everywhere. Did the claim get paid? Was the code appropriate? Was the denial overturned? Was the authorization secured? Was the documentation sufficient? Was the appeal successful? Was the transaction completed at lower cost? These aren't purely subjective judgments floating in aesthetic space. They can be measured, optimized, retried, benchmarked, and learned from. That makes RCM the coding of healthcare labor automation.

And this is why TowerBrook and CD&R invested to take the country's largest and most progressive RCM company, R1, private—where I proudly, some might say *loudly* given how much I talk about it, sit on the board. But the broader point has nothing to do with any partisanship about a favored company. It's that RCM has the properties automation wants: high volume, high cost, high-entropy data, measurable outcomes, adversarial feedback, historical records, and a giant labor substrate sitting between the clinical event and the economic transaction. And R1, with deep alignment with players like Palantir, Sierra, and Anthropic, has been at the epicenter of this industry metamorphosis. Again, shamelessly talking my own book, but keep an eye on R1—I think it will prove the

¹³¹ Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock: "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models," published on March 17, 2023 ([Link](#))

¹³² OpenAI: "Measuring the performance of our models on real-world tasks" published on September 25, 2025 ([Link](#))

quintessential example of an incumbent moving at the speed of an insurgent and providing a blueprint for how functionally verifiable domains evolve into the new paradigm. But back to our narrative.

I'll expand the aperture here and bring in e and its administrative-cost numbers. The analysis makes the target, and the terrain, even more obvious. The 2024 CAQH Index described a roughly \$20 billion annual opportunity from deeper automation and standardization in routine administrative transactions; prior CAQH work put the annual cost of tracked administrative transactions around \$89 billion.¹³³ That's before one even gets to the broader administrative superstructure of healthcare: claims follow-up, utilization management, prior authorization, provider enrollment, call centers, payment integrity, credentialing, compliance, and quality reporting. This is the sludge layer. It's expensive, measurable, and increasingly machine-addressable. So if someone asks where the first serious labor compression will occur, don't start with the surgeon. Start with the claim.

This is also where the payer labor math becomes impossible to ignore. I estimate roughly 600,000-plus payer employees nationally, with hundreds of thousands concentrated in administrative operations, claims, utilization management, provider operations, call centers, enrollment, and payment-integrity work (much more on this in my payer chapter). Even 30% to 50% automation of task content inside that administrative substrate implies very large role compression over time, offset by smaller growth in governance, model-risk, complex-case, and exception teams. The precise number will vary by payer and workflow, but the direction isn't mysterious: SG&A compresses before the bedside does.

The Bot War Starts in the Back Office

The 'bot war' begins here too, long before it reaches the bedside in full force. Provider bots will code, document, appeal, query, optimize, and pursue payment. Payer bots will deny, adjudicate, audit, route, flag, and counter-optimize. *Both sides will claim efficiency.* Both sides will claim fairness. Both sides will build muscle. **And for a while, AI will prove inflationary in healthcare** (as the mounting 2025 evidence shows) because the predominant systems use it to enhance rent extraction before equilibrium catches up. My bot codes better; your bot denies better; my bot appeals faster; your bot rewrites policy faster. Administrative musculature expands before it collapses.

But collapse it will. Eventually, when the bot war equilibrates, labor cost gets squeezed out of the middle. But the first chapter may indeed be adversarial acceleration. The headcount question for payers is therefore not whether SG&A compresses. It's who captures the savings: the payer, the provider, the employer, the patient, or the bot-war middlemen. And that matters because payer automation won't be morally neutral. If it only makes denials faster, more opaque, and more industrialized, it will inflame the public and deserve the backlash it receives. If it compresses administrative friction and returns surplus through lower premiums, better access, fewer maddening transactions, and less expensive sludge, it becomes part of the deflationary promise. Same technology. Different moral architecture.

This is why pilots should be chosen by verifiability and bottleneck status, not by vendor charm or the CEO's desire to say something appropriately modern at the next board meeting. A delightful demo in a non-bottleneck domain is innovation theater. A boring agent that reduces denial-work labor, call-center volume, prior-auth latency, documentation drag, or claims follow-up cost is strategy.

¹³³ CAQH: "New CAQH Index Reveals \$20B Savings Opportunity to Cut Waste, Reduce Costs, and Improve Patient Access" published on February 11, 2025 ([Link](#))

Decontextualization: If It Can Leave the Organism, It Can Often Leave the Human

I thought a lot about the wording on this one, and I do think this word captures it well: *decontextualization*. If you can decontextualize the work—if you can break it into component parts, send it offshore, parcel it to a BPO provider, or put it into a work-from-home queue—then some nontrivial portion of it is automatable. Outsourcing is already a tell. It means the work has been decomposed enough to leave the organism. It has been wrapped in rules, measured through metrics, quality-checked at a distance, and separated from the tacit social context of the enterprise. Once work can leave the organism, it can often be machine-mediated. Eventually, in many cases, it can be machine-executed.

This is why contracted services and remote administrative work deserve special scrutiny. I'm not making a moral claim about remote workers, and I'm not indulging some performative return-to-office sermon (although I increasingly think any group lobbying for continued work-from-home privileges is likely among the first groups summarily automated the moment the tech permits). I'm making an exposure-map claim. Work that's entirely behind a screen, routed through queues, governed by rules, and judged by measurable outputs has already confessed its decomposability. Healthcare has a prodigious amount of this kind of work: billing, coding, scheduling, credentialing, claims follow-up, medical-record abstraction, payer operations, analytics, HR, finance, procurement, and the long tail of administrative mediation that exists because the system is baroque enough to need translators for itself.

That doesn't mean every such role disappears tomorrow, but it does mean the burden of proof has shifted, and probably irrevocably. If a function can be done from anywhere, by anyone trained on the rules, through a measurable queue, then the 150 should ask why it can't be done by an agent, supervised by fewer humans, with exceptions routed to the remaining bottlenecks. This is also why the human part of work increasingly needs to become more human: embodied, relational, contextual, trust-building, judgment-bearing, hard to specify, hard to verify, and difficult to reduce to a queue. The work that's purely behind a computer will have a much harder time defending itself as inference costs fall and agents become more capable.

Teachability and the Unbundling of Clinical Work

The third rail is *teachability*. Anything built on codified heuristics, standardized guidelines, repeatable exceptions, searchable precedent, or documented workflows will be progressively exposed. That includes more clinical work than people want to admit. I'm not at all suggesting physicians disappear. I'm suggesting the bundle called "physician" may see unbundling (much more on this in my clinical AI chapter). But for now, diagnosis, treatment monitoring, medication reconciliation, guideline updates, care-gap closure, second opinions, risk stratification, patient outreach, and longitudinal follow-up all contain teachable components. Some of those components remain deeply contextual and trust-bound. Others are essentially the application of established knowledge under conditions where the model has more recall, broader synthesis, and fewer biological limitations than the clinician. Even more on this in the Doctor Disequilibrium chapter.

This is where the old professional comfort becomes dangerous. Medicine often treats clinical judgment as a monolith, as though the physician's cognitive work were one indivisible sacred act. It isn't. It's a bundle: memory, pattern recognition, guideline application, differential generation, moral reasoning, liability absorption, patient trust, physical examination, procedural skill, contextual judgment, and communication. AI strips out the parts that are codified, teachable, repeatable, and

verifiable. What remains may become more valuable, but there may be less of it than the profession wants to believe. That's the unbundling mechanism. It doesn't abolish the profession. It reprises the bundle.

The Infallibility Hypocrisy

Reader Note: This is the healthcare-labor version of the Clinical AI chapter's infallibility trap. The standard can't be machine perfection; it has to be comparison against the real human baseline, especially when the work is bounded, measurable, and functionally verifiable.

The fourth rail is *tolerance for error*. Much more on this in my Clinical AI chapter, but for now I'll simply wave at the problem. Healthcare justly has a lower error tolerance in many clinical domains, and this will slow diffusion. But it won't stop it. The key point is that humans already make errors constantly, and we've normalized those errors because they arrive wrapped in a white coat, a committee, a legacy workflow, or the invisible background noise of the status quo. We *hold technology to an inhuman standard of infallibility and perfection* while tolerating often catastrophic human fallibility. Autonomous vehicles have been trapped in this hypocrisy for years. Clinical AI will be too. Full expatiation on this concept in the other chapter, including a deep dive on the product liability and medical malpractice implications, but I include here just as a reference point.

The correct standard isn't perfection. It's equivalency or superiority to the human baseline inside a defined envelope, with governance commensurate to risk. Safer, cheaper, faster, more equitable, more measurable, and more improvable than the current process. That's the standard. Not "never wrong." Not "no lawsuit." Not "no bad headline." Not "the machine must be perfect before the human system has to compare itself honestly." This matters because error tolerance determines sequencing. Low-risk, high-volume, high-verifiability domains move first. High-risk clinical domains move later, but not never. The curve is slowed by trust, liability, and safety, not by metaphysical distinction. The machine doesn't respect our category boundaries. It improves, then presses outward.

Jevons, the Lump of Labor, and the Reorganization of Care

Reader Note: Jevons and the lump-of-labor problem also appear in the Clinical AI / Universal Doctor discussion. Here they explain why labor substitution and care expansion can coexist: some work disappears, while cheaper expertise creates more touches, monitoring, prevention, and access.

The economics won't be simple, and anyone who tells you otherwise is selling ideology, not analysis. Healthcare won't become a clean "AI destroys jobs" morality tale. It'll be messier and more interesting than that. Two forces have to be held together. On one side, workflows will unwind, roles will redefine, and labor will be substituted where tasks are verifiable, decomposable, teachable, and tolerably error-bounded. On the other side, as healthcare becomes less scarce and less expensive, *demand will expand*. Lower the marginal cost of expertise and people will use more of it. Lower the cost of monitoring and we'll monitor more. Lower the cost of behavioral support and patients will seek more of it. Lower the cost of diagnostics and earlier detection will rise.

That's Jevons' paradox in healthcare form. Efficiency doesn't necessarily shrink the system. It can expand utilization by lowering cost and increasing access. If I could see my PCP once a month instead of once a year, I probably would. If behavioral health became abundant and affordable, demand would expand proportionately (look no further than the highest use case for ChatGPT's nearly billion weekly active users: therapy and companionship). If diagnostics became frictionless, utilization would rise. If every patient with heart failure could be monitored continuously and intelligently, the system would discover more intervention opportunities than the current episodic model can see.

The lump-of-labor fallacy remains a fallacy in the narrow sense: work isn't fixed. Productivity can lower costs, expand demand, create new tasks, and move labor toward new bottlenecks. But the anti-Luddite comfort can also become too glib. The fact that new work appears doesn't mean it appears in the same volume, at the same wage, in the same geography, or for the same workers. That's the moral and political problem. Healthcare will reorganize. It won't simply shrink. Labor will be released from tractable administrative and cognitive tasks and redeployed, where possible, toward the non-tractable constraints: presence, trust, navigation, persuasion, behavioral change, chronic-care engagement, home-based support, and the human work that remains stubbornly social and embodied. The system doesn't disappear. It rearranges around new bottlenecks. The long arc may be expansionary. The middle is where the pain lives.

To understand how that reorganization unfolds, it helps to borrow a much-invoked (in Silicon Valley, lately at least!) principle from computer architecture: *Amdahl's Law*, formulated by IBM engineer Gene Amdahl in 1967, which describes the limit of system-wide improvement when only part of a process is accelerated: $S = 1 / ((1 - P) + (P / N))$ ¹³⁴ The equation is correct, and I mostly include the formula here because it looks kinda cool. But the insight, however, is important, and highly relevant to our discussion here. The maximum improvement of a system is constrained by the portion of the process that can't be accelerated. Even if one part becomes infinitely fast, *you can only go at the speed of the remaining bottleneck*. Healthcare has been living under the tyranny of this law for decades. We digitized fragments of the system—records, billing, scheduling, claims processing—yet overall productivity barely moved because the constraints migrated elsewhere. Administrative processes sped up; documentation burdens expanded. Data proliferated; decision-making remained slow. Information systems improved; care coordination stayed fragmented. The bottleneck always reappeared somewhere else in the workflow.

GenAI changes the equation not merely because it automates tasks, but because it helps institutions identify and attack the remaining bottlenecks themselves. When intelligence becomes cheap and abundant, the constraint shifts away from computation and toward system design. The institutions that benefit most won't deploy AI everywhere indiscriminately; they'll deploy it where the system is slowest, most expensive, most fragmented, and most unnecessarily human. In healthcare, those bottlenecks aren't mysterious. Administrative overhead, documentation burden, prior authorization, fragmented care coordination, diagnostic latency, and the transaction costs embedded in payer-provider warfare aren't marginal inefficiencies. They're our structural drag.

If GenAI is applied indiscriminately, we'll accelerate isolated tasks while leaving the system-level constraints intact: faster documentation feeding the same broken workflow, faster coding feeding the same adversarial payment game, faster messages feeding the same inaccessible care model. That's acceleration without redesign. But if GenAI is applied strategically, guided by the logic of Amdahl's Law, it can remove or soften the constraints that actually keep the system from moving.

¹³⁴ Wikipedia: "Amdahl's Law" ([Link](#))

And once those constraints begin to disappear, the consequences extend far beyond physician workflow or hospital operations. A sector representing nearly one-fifth of the American economy begins reorganizing itself around machine intelligence. At that point, healthcare looks like the primary theater in which capital approaches something close to perfect substitutability for labor.

Put all of this together and the first exposed healthcare domains become clearer. The early targets aren't mysterious: administrative work, RCM, payer operations, coding, prior authorization, claims, call centers, documentation, scheduling, analytics, contracting support, credentialing, internal knowledge retrieval, quality reporting, compliance compilation, and significant chunks of utilization management. But rather than reciting a giant list like a man trying to prove he has suffered through the indignities too many operating reviews, the better way to think about the sequence is through three categories.

The first is transactional cognition: claims, coding, eligibility, authorization, payment integrity, denials, appeals, and other work where the goal is measurable and the environment is already computational. The second is administrative narration: documentation, summarization, routing, correspondence, quality reporting, compliance drafting, knowledge retrieval, and the endless explanatory work that keeps healthcare's bureaucracy metabolizing itself. The third is clinical-adjacent synthesis: medication reconciliation, care-gap identification, risk stratification, guideline updates, triage support, second opinions, and the parts of care management where established knowledge can be applied more continuously and cheaply than the current human model permits.

Payers may see SG&A compression before providers see clinical labor compression. Hospitals may see contracted-services and administrative labor compression before they see bedside clinical compression. Physician practices may see panel expansion before they see headcount reduction. But the direction isn't obscure. The first wedge is verifiability. The second is decontextualization. The third is teachability. The fourth is agentic execution. The fifth is bottleneck migration. And the strategic implication for the 150 is straightforward, if not comfortable: don't wait for this to show up as a vendor category. Build the map yourself. Commission you own, internal healthcare GDPval (while we wait for the construction of the national one). Identify the workflows where the current labor model is exposed. Decide which human functions should be preserved, which should be reconstituted, and which should be allowed to disappear because they should never have existed in the first place. That's how substitution propagates. Not as a thunderclap. As a thousand work queues quietly becoming less human.

The likely sequence isn't mysterious, though healthcare consultants will do their best to make it look mysterious! The first six months (and yes, I'm denominating our timescale in months, not years) are administrative compression disguised as augmentation. Ambient documentation reduces note burden. Agents draft appeals. Coding tools improve yield. Call-center assistants answer routine questions. Scheduling gets less Byzantine. Most leaders will describe this as relief, and some of it really will be relief. But the labor question is already embedded inside the relief: if a team can resolve the same queue with half the effort, does the organization expand the queue, redeploy the people, or shrink the team? That's the first fork.

The second six months are workflow redesign. The winning institutions will stop placing AI on top of old work and start removing steps altogether. Prior-auth packets will be assembled automatically from the record. Denials will be triaged by likelihood of overturn. Patient messages will be routed by acuity and intent. Quality reporting will be assembled from underlying evidence rather than manually reconciled through ritual suffering. The losing institutions will proudly report that their people are "using AI" while the old queues, staffing assumptions, and management layers remain intact. That's

personal productivity without enterprise productivity, and healthcare has a long, almost heroic history of settling for it.

The third six months are denominator expansion and political confrontation. If the savings are real, the public will ask where they went. If access doesn't improve, premiums don't soften, patient experience remains miserable, and executives talk about transformation while the payroll shrinks, the backlash will be ferocious. If, instead, the institution can show more primary-care contact, more behavioral-health support, faster authorizations, fewer denials, better navigation, lower unit cost, and a humane transition for workers, then AI becomes defensible. The operating question and the moral question converge: where did the winnings go?

This sequencing also prevents magical thinking in both directions. The enthusiasts shouldn't pretend that every clinical act is about to be automated. The skeptics shouldn't pretend that because some clinical acts remain protected, the administrative and cognitive substrate around them is safe. Healthcare contains many kinds of work, and the substitution curve will move unevenly. The first act is administrative compression: payer SG&A, RCM, coding, documentation, scheduling, claims follow-up, eligibility, authorization, and compliance compilation. Then comes (almost simultaneously) the broader clinical-adjacent layer: physician copilots, second-opinion engines, longitudinal care managers, intelligent care coordinators, remote-monitoring interpreters, digital pathology support, radiology triage, behavioral-health companions, and adaptive care plans. The work doesn't vanish all at once. It gets recomposed around new bottlenecks.

And let me be clear about sequencing: this is not an argument to wait patiently for administrative AI to proliferate before moving into clinical AI. Quite the opposite. Clinical AI needs to move quickly and decisively too, but with different governance, different error envelopes, different legitimacy requirements, and a deeper appreciation for trust, liability, and the sacred terror of touching actual bodies rather than claims queues. Full 60-page chapter on this metamorphosis in my Clinical AI section. Administrative AI may diffuse first because it's more verifiable. Clinical AI may matter more because the stakes are higher. Both need to move. They just won't move along the same rail at the same speed. Some work will be automated. Some will be augmented. Some will be eliminated because the workflow itself should disappear. Some will become more valuable because it's the new constraint. The strategic job is to map the difference before the market maps it for you.

Agents, Digital Workers, and the Hive Mind

Reader Note: The agent / Hive Mind motif appears in Clinical AI and returns in the China chapter. I repeat it here to move from clinical coordination to labor architecture: agents aren't only tools but a new class of digital coworkers.

The next mechanism is *agentification*. A chatbot is annoying and useful in roughly equal measure. An agent is different. A tool waits; a workforce acts. An agent doesn't merely answer; it uses tools, remembers context, executes workflows, collaborates with other agents, and returns with work product. That's where labor substitution stops being metaphorical.

I'll offer what might sound like a non sequitur idea here as we think about agents, but hear me out. It's tempting to think of agents as individual software employees. That's a serviceable mental model, but still very much incomplete. Agents aren't merely replacements or augmenters for individual humans; they're members of a *digital population* whose learning can be (eventually, as the exponentials

continue advancing) instantly repatriated to the whole. A human worker learns slowly, locally, idiosyncratically, and, well, forgetfully. A digital worker can perform a task, transmit the learning in femtoseconds, and have that improvement propagate across the fleet. *This is the hive-mind advantage.* Human civilization advanced because of *intelligence* and *sociability*: we could reason, and we could share. But our sharing mechanisms are comically low bandwidth. Speech, writing, meetings, slide decks, Slack threads, status updates—the whole totalitarian choreography of corporate life exists because knowledge is atomized across skulls and has to be moved around slowly, lossy packet by lossy packet.

This is why humans spend their lives in eternal meetings. It's not because meetings are good and fun and life-affirming. It's because the information is trapped in separate biological containers. The CFO knows something the COO doesn't. The head of revenue cycle knows something the CMO doesn't. The payer-contracting team has some crucial little piece of institutional knowledge that the analytics team doesn't. So we gather in rectangles, furtively check sports scores on our phones, summarize, misunderstand each other, half-listen, misremember, and then schedule another meeting to repair the informational damage caused by the first one. Machines don't need meetings in the same way. Or at least, if they do, their meetings aren't ninety-minute rituals of status negotiation and performative alignment.

Agents invert the knowledge-transfer problem. A population of agents can divide labor, specialize, check one another, execute tasks, and transmit learning back to the orchestration layer. One agent reads the policy manual. Another reviews the claims. Another drafts the appeal. Another monitors outcomes. The insight isn't trapped in one tired analyst's notebook. It can become fleet knowledge. That's a vastly different organizational physics. This notion deserves more than three paragraphs, and don't be surprised if it shows up with a lot more prose elsewhere in the essay.

Digital Immigration and the New Labor Class

I'll indulge in one or two more quick extrapolations off this core idea. For one, this is why “digital workers” might be better understood as *an immigration wave*, not a product feature. We're moving from capital to compute to labor. Trillions of dollars go into electricity, chips, data centers, networking, cooling, land, and energy infrastructure. Those physical inputs instantiate artificial labor: indefatigable, 24/7, marginal-cost workers without sleep, HR complaints, burnout, or benefits expense. From 996 (the Chinese phrase encapsulating the idea of working nine am to nine pm, six days a week) to 24/7, as the kids might say. Or rather, as the models might say once they learn labor slang. A massive migration of high-skilled, low-wage workers streaming across the border, only the border is digital and not geographic this time.

The cost structure is almost absurd. If an agent can perform a task at one-thirtieth the cost of a human (as a recent study postulated), or one-hundredth, or eventually one-thousandth as inference gets cheaper, then the wage ceiling for the human equivalent collapses. The price of electrons to run the model becomes the upper bound on wages for substitutable work. No firm will long employ a human to do work a machine can do better, faster, and cheaper unless regulation, trust, liability, brand, or ethics require the human. This is the most uncomfortable part of the mechanism. *The agent becomes labor instantiated through capital.* And once capital can instantiate labor, the capital-labor relationship changes at the root.

But it isn't just a mass migration of workers. What about one, omniscient agent at the top of the organization? This is why the “AI CEO” thought experiment stops being cute. A human CEO is constrained by energy, attention, memory, presence, and physiology. She can't be in every product

review, operating review, strategy session, customer escalation, coding dispute, payer negotiation, and board committee. An AI executive layer, however, can in principle ingest every meeting, every metric, every customer interaction, every variance, every operational signal, and every agent's output. It can become a kind of institutional Leviathan: not omnipotent, but far more present than any biological executive. I'm not predicting that hospital boards will appoint Claude CEO next week. Please don't send me angry governance memos. The point is more modest and more important: management itself becomes agentic. The org chart changes. HR and IT begin to converge because the workforce now includes humans and digital agents. Someone has to recruit, credential, monitor, evaluate, terminate, retrain, and govern non-human workers. That's the new HR.

And we'll have to reconceptualize so many artifacts that scaffolded the old HR. Including the institutionalized org chart, born in the 19th century (1855 in fact, from the Pennsylvania railroad) and refined into the form factor that has become so ubiquitous we just assume it's permanent. But in this new model we're envisioning, it starts to look antique. Tally up the sum of what we've been discussing, and let's think about the emerging structure: first comes the traditional pyramid: executives, senior management, middle management, supervisors, administrative staff. Then comes the automation diamond: the administrative base shrinks, middle coordination changes, and senior leaders are surrounded by agentic leverage. Then, in some functions, comes the inverted pyramid: a smaller number of humans orchestrating large populations of agents underneath them. Less span of control over humans; more span of control over agents and their synthetic workflows. Brave new world, indeed.

The Rails, in Summary

Ok let's bring this section to a close. This, in short, is how substitution propagates. First, capital becomes compute, compute becomes cognition, and cognition begins performing labor. Then the work sorts itself by architecture. Verifiable work moves first. Decontextualized work follows. Teachable work gets unbundled. Error-bounded work diffuses as the human baseline is forced to compare itself honestly. Agentic systems turn tools into workforces. Bottlenecks migrate, sometimes creating new human demand and sometimes exposing the old human role as a historical artifact.

Healthcare is full of these rails. It has claims that can be verified, notes that can be summarized, denials that can be appealed, policies that can be interpreted, queues that can be routed, clinical knowledge that can be retrieved, guidelines that can be applied, patients who can be monitored, workflows that can be decomposed, and administrative labor that has already confessed its vulnerability by moving offshore, into BPO contracts, into work-from-home queues, or into rules-based software systems that never quite delivered the productivity gains we were promised.

The question for the 150 isn't whether substitution will arrive. It already has. The question is whether they'll govern it deliberately enough to prevent it from becoming a chaotic collision of vendor exuberance, worker panic, guild defense, payer-provider bot warfare, and political backlash. The mechanism is now visible. The rails are laid. The next question is whether healthcare leaders have the courage to build the train deliberately—or whether they'll wait until it runs through the institution without their wisdom and sanction.

Part VI—Resistance, Risk, and the Human Terms of Healthcare Automation

Mechanism isn't destiny. The machine enters a social system full of guilds, unions, regulators, malpractice fears, professional identities, local economies, and human beings who quite reasonably don't want to be turned into residue. This part is about the *counterforce*: the ways healthcare will resist, the ways resistance will sometimes be justified, and the ways safety language can also become a sanctified instrument of incumbency protection.

The Machine Enters a Civic Industry

The machine won't enter a vacuum. That was true in the labor chapter as a general argument, but it's even more true in healthcare. AI enters guilds, unions, licensure regimes, medical staffs, malpractice law, payer contracts, public programs, community expectations, household fear, professional identity, and the thick moral atmosphere of care. It enters a sector where labor isn't merely labor because so much of the work touches dependence, pain, humiliation, disability, frailty, and, yes, death. A healthcare leader who understands the technology but misunderstands the social system will fail just as surely as the leader who understands the politics but refuses to see the substitution curve.

That's why this part has to be about resistance, risk, and human consequence inside healthcare specifically. The labor chapter already made the broad point that the AI establishment is too silent about job loss, that cognitive offloading can degrade judgment, and that work is more than wages. Here the question becomes narrower and more operational: how do those risks show up in a health system, a payer, a physician group, a revenue-cycle shop, or a regional hospital economy? Where does legitimate safety end and guild defense begin? When does human oversight become a liability fiction? How do we redesign work so AI removes dehumanizing administrivia rather than turning clinicians and staff into bored biological validators of machine output?

The Guild Will Call It Safety, Labor Will Call It Justice

Reader Note: This reprises the Clinical AI chapter's warning that safety language can become guild protectionism. Here I widen the frame: in healthcare labor, resistance will come not only from professional guilds but also from unions, local economies, and workers invoking justice.

The counteroffensive won't take long, and in some places it has already started. Healthcare has always had stronger defensive vocabularies than other industries because the words are genuinely weighty: safety, quality, professionalism, scope of practice, patient protection, standards, licensure, equity, community benefit. These aren't fake words. A bad clinical model can injure people. A poorly governed agent can deny necessary care, misroute a patient, hallucinate a dangerous recommendation, encode bias, or turn a frightened human being into a workflow artifact. Safety is real.

Illinois is a useful harbinger. Governor JB Pritzker signed the Wellness and Oversight for Psychological Resources Act, restricting the use of AI in therapy and psychotherapy services while allowing administrative and supplementary support for licensed professionals. The state's own framing explicitly cites its mandate not only to protect patients, but also to *protect the jobs* of qualified behavioral-health providers. That's the politics in miniature: safety and workforce protection braided together, sometimes legitimately, sometimes opportunistically, often inseparably. Behavioral health is the first obvious battlefield because the unmet need is enormous, the human intimacy is

real, the safety risks are visible, and the labor supply is catastrophically insufficient. See my behavioral-health chapter for the full lament of what we're losing with this kind of misguided legislation and regulation.

The 150 need a more mature posture than Silicon Valley accelerationism or professional obstruction. The right answer isn't to let untested models roam around frail clinical environments. It's also not to prohibit AI from doing anything that currently produces a paycheck. The answer is rigorous evals, staged deployment, real-world monitoring, auditability, model-risk governance, human escalation, and a clear distinction between safety constraints and job-protection constraints. If the guild says a model is unsafe, ask the question that changes the moral terrain: compared to what? Compared to an idealized human process in a policy memo, or compared to the actual process, with missed diagnoses, exhausted clinicians, inaccessible psychiatry, documentation fatigue, prior-auth purgatory, and patients serving as their own integration layer?

Truth-Telling Before the RIF

Healthcare shouldn't participate in the euphemism machine. Workers are smart. They can see hiring freezes, closed open recs, contractor compression, management-layer thinning, outsourced work being pulled back and agentified, and the disappearance of entry-level administrative roles. Pretending these phenomena are unrelated to AI will look insulting before it looks reassuring. The better strategy is *truth plus solution set*: tell workers that some work is going away; tell them early enough that attrition, redeployment, retraining, severance, mobility, and community planning remain possible; tell the public that the alternative to labor substitution isn't some morally pristine status quo, but continued unaffordability, medical debt, psychiatric deserts, delayed primary care, exhausted clinicians, and payer-provider trench warfare.

Truth-telling changes the moral architecture of the transition. If you tell the truth early, attrition becomes useful. Redeployment becomes plausible. Training becomes real where training is real. Generous severance becomes feasible where training is not. Regional planning becomes possible. Community-college partnerships become possible. Internal talent markets become possible. Local political legitimacy becomes possible. If you wait until the labor math is undeniable, the only instrument left is the RIF, and the RIF arrives as betrayal.

This is one reason the one-year window matters. Healthcare still has open requisitions, voluntary turnover, contract labor, outsourcing relationships, and management layers that can be thinned without immediately converting the transition into mass layoffs. Those are temporary gifts. They won't last forever. Once the workforce fully internalizes that technological unemployment isn't a science-fiction topic but an operating strategy, voluntary attrition will slow, unionization pressure will intensify, and every unfilled role will become politics. Early candor preserves options. Late candor produces panic.

Cognitive Offloading as a Clinical Safety Problem

Reader Note: This section deliberately connects three earlier threads: the labor chapter's apprenticeship problem, the Clinical AI chapter's concern about unaided judgment, and the Future of Science chapter's cognitive-offloading warning. Healthcare makes the cognitive risk acute because degraded judgment can become patient harm.

There's another risk that belongs especially in healthcare: cognition itself. The apprenticeship problem was introduced in the labor chapter, and the degradation of human cognition as an acute risk in the earlier Clinical AI chapter, but healthcare gives it life-and-death force so it must be restated here. Clinical judgment, operational judgment, and revenue-cycle judgment are all built through repeated exposure to ordinary cases before the exception appears. A resident learns that the patient who “just looks wrong” is deteriorating because she has seen enough ordinary-looking patients. A nurse senses the room has changed because she has lived inside enough rooms. A revenue-cycle leader spots a new denial pattern because she has seen enough old ones. I've said this before in the essay, but I'll repeat it again here: judgement isn't a museum object one preserves under glass and retrieves for emergencies. It's a trained faculty that degrades with neglect.

The *verification trap* is therefore a healthcare safety problem, not merely an epistemological curiosity. In theory, the machine handles the routine work and the human handles the exceptions. That sounds clean, even humane: let the model do the sludge and let the person handle the meaningful residue. But exception-handling isn't an innate human virtue. It's built through routine exposure, correction, repetition, and inward possession of the domain. If the routine opportunities for practice disappear, the human becomes less capable precisely at the moment the institution says the human is needed for oversight.

This is why human in the loop can become a talisman, and a lazy one. A bored, deskilled, overloaded, cognitively atrophied human in the loop isn't governance. Verification is cognitively demanding. It presupposes taste, memory, context, pattern recognition, and the ability to know when the machine's answer is polished nonsense. *A person can't meaningfully verify what she has never learned to do.* The better goal is human agency in the loop: competence, authority, practice, accountability, and the right to challenge the machine rather than merely approve its output.

Healthcare should redesign apprenticeship before the old repetition disappears. If machines write the first draft, humans need to learn by interrogating the draft. If machines summarize the chart, humans need to compare the summary to the messy underlying reality. If machines recommend the code, the differential, or the workflow action, humans need to learn why the recommendation is right, where it's brittle, what it omits, and how reality may humiliate it. The old system created competence accidentally through volume. The new system will have to create competence intentionally through simulation, exception review, adversarial case conferences, human-machine disagreement, and supervised practice with accountability.

When the Human Becomes the Robot

Then there is the question of purpose. The official story of automation is always liberatory: machines will take the dull, dirty, dangerous, disagreeable, or dear tasks, and humans will be freed for more creative, relational, and spiritually nourishing work. Sometimes that happens. Often it doesn't. Often the human becomes the robot.

I mentioned this earlier, but I'll reframe it here: Amazon's fulfillment centers are the cautionary precedent, not because hospitals are warehouses, but because Amazon shows what happens when machines reorganize human labor around their own tempo. Robots reduced some physically punishing work. Good. But the newer automation stack also narrows the human role, increases throughput, flattens hiring curves, and organizes workers around machine rhythm. In some settings, the machine gets the orchestration and variation while the human gets monitoring, picking, verification, and trying not to fall behind the queue.

Healthcare has to avoid importing that pattern into care. The danger isn't a theatrical robot nurse marching mechanically down the hallway with an aluminum bedpan. The danger is subtler: orchestration systems that remove searching, typing, routing, remembering, scheduling, documenting, and escalating, and then leave humans with a narrower, more surveilled, more machine-paced residue. The robots do the interesting variation. The humans approve, deny, escalate, correct, and absorb blame. That isn't a future worth building.

Work isn't merely compensation. It's routine, dignity, status, friendship, competence, identity, discipline, and sometimes a blessed distraction from oneself. People need purpose, discretion, and some narrative of contribution. They don't flourish as biological API endpoints. Automation that improves physical workload while worsening psychic workload isn't an unambiguous victory. It may be better ergonomics with worse souls, which sounds melodramatic until you have watched a worker reduced to a verification appendage for a system she doesn't understand and can't meaningfully influence.

The whole point of AI in healthcare should be to move humans back toward the human parts of care: presence, explanation, reassurance, touch, trust, relationship, complex judgment, team leadership, community embeddedness, grief, fear, hope, and the thousand subtle interpersonal acts that make care feel like care rather than transaction processing. That requires co-design. The insurgents bring technology, speed, capital, and irreverence. The incumbents bring workflow reality, patient trust, regulatory legitimacy, tacit knowledge, and moral seriousness. Either side alone will build something defective.

The Duty to Care for Displaced Healthcare Workers

Healthcare can't plead powerlessness here. The 150 aren't passive observers of an approaching asteroid. They are institutional actors with capital, influence, data, political standing, local legitimacy, and the ability to shape adoption. Labor substitution may be good for society and bad for specific individuals. The technological optimist who ignores individual harm is a sociopath. The labor protectionist who ignores social benefit is a reactionary. The adult position is to build the technology and build the transition with as much wisdom, compassion, and foresight as we can summon.

Healthcare has a special duty because it benefited from the old labor-intensive equilibrium for decades. It offered stable, decent, middle-class work to millions of people, many without advanced degrees, many women, many immigrants, many workers in communities where the hospital is the economic anchor. If the 150 now use AI to reduce labor intensity, they can't pretend displaced workers are simply market externalities. They are people who built the institution.

We are, in some sense, breaking a covenant with employees. Not the formal covenant—employment was never a lifetime guarantee, and healthcare can't remain unaffordable to preserve every existing job category. But there has been an implicit psychological covenant: come here, work hard, endure the bureaucracy, and healthcare will provide durability. That covenant is going to fracture. The question is whether leaders fracture it with honesty and generosity, or with euphemism and cowardice.

A compassionate transition framework should begin before the panic. Use attrition first. Close open requisitions before layoffs. Require a burden of proof for new headcount: why can't AI do this, or why won't it be able to do this within a year? Freeze low-value hiring. Eliminate contracted services before eliminating loyal employees where feasible. Offer generous severance. Fund reskilling, but don't lie that reskilling will solve everything. Provide bridge income. Help workers move into bottleneck roles: care navigation, behavioral-health support, community health, elder care, model evaluation, technical apprenticeship, and the forms of high-trust human service that expand when care

becomes less scarce.

Some people will make the transition. Some will not. A serious institution tells the truth about both. If AI materially expands margins or reduces cost, some of that surplus should be explicitly earmarked for transition support. One could imagine a healthcare version of an Alaska Permanent Fund for displaced workers: a margin-sharing pool, transition dividend, GPU tax, token tax, or institutional fund that supports workers while they retrain or search. We can debate the mechanism. The principle is what matters. The institution that benefits from substituting technology for labor has a duty to help the humans it no longer needs in the old workflow.

The New Covenant

This isn't sentimentalism. Consider it, well, *prudence with a conscience*. A transition that's economically correct and socially brutal will invite backlash, regulation, union militancy, litigation, and political confiscation. A transition that's economically correct and morally serious may buy legitimacy for the deflationary project healthcare needs. Compassion isn't the opposite of strategy. It's how strategy survives contact with democracy.

The new covenant has to be explicit: fewer humans doing bad work, more humans doing human work, and real generosity toward those whose old work the machine makes unnecessary. The guild will call everything safety. Labor will call everything justice. Silicon Valley will call everything progress. The CFO will call everything productivity. The consultant will call everything transformation. The worker will call it a life. The patient will call it whether she can get care, afford care, trust care, and feel cared for. The leader's task is to keep all of those truths in view without collapsing into denial, bloodlessness, or theater.

Trust Is the Constraint That Survives Automation

One final human point before the playbook. In a healthcare system reorganized around machine cognition, trust becomes even *more* valuable, not less. Patients won't experience AI as a production function. They will experience it as a voice, a message, a denial, a recommendation, a delay, a reassurance, a risk score, a treatment plan, or the eerie sense that no human being is quite responsible for what is happening to them. The technology may be brilliant and still feel cold. It may be accurate and still feel illegitimate. It may be efficient and still feel like abandonment.

That means the protected human work isn't merely the work AI can't technically perform. It's the work that must remain *socially legible as human* because legitimacy depends on it. A model may generate the right differential, but a clinician still has to explain what it means to the frightened patient. An agent may assemble the appeal, but a human organization still has to own the denial and its reversal. A risk model may identify the patient most likely to deteriorate, but a care team still has to persuade, comfort, coordinate, and sometimes sit with the fact that the medically correct answer is socially impossible.

This is the mistake the margin-recapture version of AI will make. It will assume that because the machine can produce the output, the institution can remove the human relationship around the output. Sometimes yes. Often no. Healthcare is a trust business before it's an information business. The great opportunity is to use machines to clear away the information labor that has crowded out trust. The great danger is that we clear away the humans instead.

Part VII—What the 150 Can Do: Installation, Governance, and the New Covenant

This final operating part turns the argument into an early proposed doctrine. The 150 don't need another AI committee or another polite inventory of pilots. They need an *installation strategy*: a way to convert machine cognition into cheaper, better, more abundant care without converting workers and communities into externalities.

From Admiring the Future to Installing It

So what does one do if one runs, governs, finances, or advises one of the 150? Don't admire the future from a safe institutional distance. Don't run two hundred pilots, some performative, some useful, nearly all destined to get stuck in the purgatory between proof-of-concept and production. Don't wait for perfect consensus from the medical staff, the board, legal, IT, compliance, HR, the innovation office, the payer-relations team, the patient-experience committee, and whichever specialty society has just discovered that its professional sovereignty is being inconvenienced by history. And certainly don't convene another blue-ribbon task force whose final recommendation is that the organization should continue thoughtfully exploring opportunities to responsibly leverage artificial intelligence. Let's skip that part.

The 150 need a doctrine. Then they need to execute against it.

This part is the uncomfortable management section. The philosophical work has been done, perhaps more elaborately than strictly necessary. The old technological bargain depended on *reabsorption*. GenAI threatens that bargain because it industrializes intelligence itself. Healthcare is the main event because it's the largest, most labor-intensive, most administratively encumbered, most productivity-challenged sector in the economy. Substitution propagates through functional verifiability, decontextualization, teachability, agentic execution, and bottleneck migration. Resistance will come from guilds, unions, regulators, professional identities, and communities that experience labor contraction as civic betrayal rather than operating leverage.

Fine. Now what?

The answer, in compressed form, is this: shrink labor intensity where the work is bad, expand the denominator where care is scarce, move humans toward the bottlenecks that remain genuinely human, become AI-naturalized rather than AI-curious, build a serious task map of the enterprise, treat agents as a new labor class, use the current capital mobilization while it lasts, and do the whole thing with enough candor and generosity that the public believes the winnings are being shared rather than strip-mined. That's the doctrine. The rest is execution.

The Care Denominator Has to Expand Too

One more clarification before the playbook, because otherwise the labor thesis can sound too much like austerity in better clothes. The goal isn't simply to shrink the numerator. The goal is to *expand the denominator of care*. If AI makes one clinician, one care manager, one scheduler, one coder, or one analyst more productive, the institution has two choices. It can do the same work with fewer people, or it can do more work with the same people. In a normal industry, that's mostly a margin and growth question. In healthcare, it's a moral question because latent demand is everywhere.

Primary care is under-touched. Behavioral health is under-supplied. Chronic disease is under-managed. Medication adherence is under-supported. Frail elders are under-watched until they fall, decompensate, or show up in the emergency department. Caregivers are under-supported. Rural

patients are under-reached. Patients with social needs are under-navigated. None of that disappears because we automated a documentation workflow. If anything, lower marginal cost reveals how much care the current system has been rationing through friction, price, staffing scarcity, and administrative exhaustion.

So the serious version of AI in healthcare isn't just labor compression. *It's labor reallocation plus care expansion.* Compress the administrative numerator. Expand the care denominator. Move human attention away from clerical metabolism and toward trust, persuasion, longitudinal support, serious illness, behavioral change, and the messy social work of helping people actually live differently. If the denominator doesn't expand, the public will experience AI as a layoff machine. If the denominator expands and the transition is humane, the public may experience AI as the first credible path toward affordable abundance.

The Numerator-Denominator Choice

The central managerial question is *numerator* or *denominator*. Do you do the same amount of care with fewer inputs, or do you do dramatically more care with roughly the same inputs? The answer, annoying because it's true, is *both*—by domain, in sequence, and with a very clear moral vocabulary. Pretending the numerator is sacred is avoidance. Pretending the denominator will magically expand while every incumbent job category remains untouched is arithmetic denial.

The numerator is the input base: staffing, contracted services, administrative overhead, management layers, real estate, agency spend, SWB, and the whole cost structure of the enterprise. *The denominator is the output base:* lives covered, geographies served, patients touched, visits delivered, panels managed, episodes handled, procedures performed, members enrolled, risk lives managed, contracts served, and access expanded. **In a post-AI world, the 150 have two basic strategic choices.** They can keep the numerator roughly flat and massively expand the denominator, doing much more with the same people. Or they can keep the denominator roughly flat and shrink the numerator, doing the same work with fewer people. The winners will probably do both, but the order and morality matter enormously.

Now for the unpleasant sentence: the 150 need to contract labor intensity. Not because labor is bad. Not because employees are disposable. The reason is more basic and less theatrical. *Healthcare is unaffordable, labor is the dominant input, and GenAI finally makes substitution possible in a sector that has evaded almost every prior productivity wave.* A system spending something like \$2.9 trillion on labor inside a \$5.3 trillion sector can't treat labor as an untouchable object of reverence while patients carry medical debt, employers suffocate under premiums, and the public sector bends around Medicaid, Medicare, subsidies, provider taxes, and every other financing contraption we've built to disguise the system's cost.

Jensen's formulation—three times the revenue with only a quarter more staff—is the right formula for the age. Big Tech is the precursor. The most admired companies are no longer boasting about headcount as empire. They're boasting, sometimes insufferably and arrogantly, about operating leverage, revenue per employee, tiny teams with enormous output, and growth without commensurate staffing. Healthcare is behind. It still too often confuses staff growth with mission growth, as though adding people were intrinsically a sign of civic virtue. Sometimes it is. Often, it's just the old production function failing to imagine another way to create capacity. That habit needs to end.

This is how healthcare becomes affordable: not by pretending care is cheap, and not by starving the bedside, but by lowering the numerator required to produce a unit of care. The denominator can and

should expand—more lives managed, more outreach, more prevention, more longitudinal contact. But if the labor numerator remains sacred, the system will simply convert AI into another expensive overlay. The 150 have to make personal productivity become enterprise productivity, which is where healthcare has historically failed.

There should be a moratorium, or near-moratorium, on nonessential new hires. Close open requisitions. Treat attrition as an asset rather than a crisis. Require every headcount request to answer a simple question: why is this not an AI-enabled workflow redesign problem? The burden of proof should shift from “why eliminate?” to “why hire?” That's a cultural inversion, and many leaders will hate it because it feels ungenerous, frightening, and alien to healthcare's shortage psychology. They should do it anyway, and they should do it before the workforce fully understands that the old shortage narrative isn't a permanent shield against substitution.

The One-Year Window

The window for attrition-first rightsizing is short. Maybe a year, maybe two, but not a decade. Once workers fully internalize that technological unemployment isn't science fiction but operating strategy, voluntary attrition will *slow*, unionization pressure will intensify, political resistance will harden, and every unfilled requisition will become a battleground. Use the inhalation before the exhalation. Close the open recs now. Let attrition do the humane part of the work before layoffs have to do the brutal part.

This also means out-counseling mediocrity, a phrase I know sounds unpleasant because it *is* unpleasant. But cultures of false kindness become cruel to the high performers who carry the institution. Healthcare has a habit of treating rigor as meanness and mediocrity as an unfortunate but permanent feature of the terrain. That posture is going to become unaffordable. A rigorous, meritocratic organization that uses AI to remove low-value work and redeploy talent into higher-leverage roles will be better for the best humans, not worse. Superstar employees resent mediocrity more than leaders realize. They resent the meeting custodian, the status translator, the person whose contribution is an atmospheric mist of process, the manager who adds friction and calls it alignment. AI will make some of that visible, and the serious organizations shouldn't flinch when it does.

Management layers need particular scrutiny. Agentic systems reduce coordination costs, or at least they should. If middle management exists primarily to move information through meetings, translate metrics, chase status updates, reconcile work queues, and narrate progress upward in a tone of strategic calm, then the layer will compress. That's not an insult to middle managers. Some are indispensable operators, translators, coaches, and culture carriers. But many are human routers in an information-scarce organization. Once the information layer becomes continuous, machine-readable, and agentially monitored, the human routing layer has to justify itself again.

The inherited org chart is a 19th-century artifact wearing a Workday interface. The classical *pyramid*—executives, senior management, middle management, supervisors, administrators—was built for information scarcity and human coordination. AI first turns it into a *diamond*: the administrative base compresses, middle coordination changes, and senior leaders become more leveraged by agents. Then, in some functions, the diamond starts to invert. Fewer coordinators. Fewer status translators. Fewer meeting custodians. More accountable operators supervising fleets of agents. The 1855 org chart doesn't survive contact with synthetic labor unchanged.

My original notes for this chapter were even blunter: *this is the new HR*. Managing humans alone was

the industrial-firm problem. Managing humans plus agentic workers is the post-AI problem. HR, IT, operations, compliance, and finance begin to converge because the workforce itself has changed. Moderna's move to combine IT and HR wasn't a curiosity; it was a harbinger.¹³⁵ A serious CHRO will need to understand model governance; a serious CIO will need to understand workforce substitution; and a serious CEO will need to understand both well enough not to be managed by either vendor theater or internal obstruction.

Horizontal consolidation in healthcare should change too. For decades, nonprofit mergers often preserved two of everything: two CEOs, two headquarters, two leadership teams, two cultures, two sacred local duplications, all maintained in the name of community sensitivity, political expedience, and the great *Noah's Ark principle* of healthcare consolidation. In most industries, one justification for mergers is eliminating redundancy. Healthcare often neutralized that logic and then wondered why consolidation didn't produce meaningful cost reduction. Under GenAI, redundancy will become harder to defend. Mergers may reaccelerate, and the operating thesis will have to include labor and management-layer compression. If two organizations combine and preserve every duplicative administrative layer, they aren't integrating. They're performing a civic pageant of integration while protecting the old numerator.

Invest in Compute, Not Beds

Reader Note: "Invest in compute, not beds" returns in the hospital chapter as a capital-allocation doctrine. It appears here first as a labor-substitution consequence: if cognition becomes an infrastructure layer, compute becomes a care asset.

The 150 should also stop building so much real estate (the sector *already* has some \$1.5 *trillion* of buildings and land) as though beds are the only serious capital asset. *Invest in compute, not beds*. The commandment is simple enough to sound unserious, which is usually when a commandment is useful.

Not literally never beds. Trauma, ICU care, surgery, obstetrics, complex inpatient medicine, regional access, and real clinical acuity still require physical capacity. The hospital isn't going away, and anyone who says otherwise should be gently escorted away from the strategic-planning session and given a tour of an ICU. But the next strategic dollar has to be interrogated with a new suspicion: is this a real-estate dollar, a labor dollar, or a compute dollar? The old healthcare answer was almost always the first two. Build the tower. Hire the staff. Expand the service line. Add capacity through buildings and people. The new answer increasingly needs to be the third.

If marginal care becomes more virtual, proactive, home-based, AI-mediated, and continuous, then the next trillion dollars of healthcare capital can't simply be another real-estate binge. Hospitals will still matter, but the strategic center of gravity *shifts toward the intelligence layer* that lets care move out of the hospital when appropriate, stay connected between visits, coordinate around the patient rather than the facility, and reduce the expensive human labor required to move information through a fragmented system. Compute isn't a gadget category here. It's the new productive asset. It's the means by which capital becomes cognition and cognition becomes capacity.

This is where numerator and denominator reconcile. The denominator expansion path is the more morally attractive one if the system can execute it: more touchpoints, more prevention, more

¹³⁵ Forbes: "Moderna's Game-Changing Reorg Mergers HR and IT" published on August 28, 2025 ([Link](#))

behavioral-health support, more adherence management, more navigation, more care in the home, more earlier diagnosis, more rural access, more primary care, more human presence where presence matters. But denominator expansion must be real. If leaders keep the labor base intact and merely call vague growth aspirations a strategy, the math will punish them. The numerator path is harsher but sometimes necessary: do the same work with fewer people. The adult strategy is probably both. Shrink the labor trapped in administrative loops, expand the human labor devoted to scarce relational bottlenecks, and let the care model grow where lower marginal cost reveals latent demand. That's not austerity. That's substitution in the service of abundance.

None of this can be democratized down to the department level. Department heads, consciously or unconsciously, see their own disintermediation coming and will resist. They will ask for more study, more pilots, more governance, more committee review, more validation, more stakeholder engagement, more time. Some of that will be sincere. Some of it will be the organism protecting itself. The CEO has to be assertive. Maybe autocratic, in the narrowly virtuous sense I use elsewhere: not cruel, not impulsive, not contemptuous of expertise, but clear that this isn't a moment for consensus-driven diffusion through committees whose members are structurally incentivized to protect the old workflow. The CEO can't delegate this to innovation theater.

Become AI-Naturalized, Not Merely AI-Curious

The 150 don't need to become AI-native. They can't. They weren't born inside the new paradigm, and pretending otherwise will only produce bad startup theater inside large institutions. But they must become *AI-naturalized*. There's a difference. *Naturalization* means the institution learns the language, customs, tools, risks, economics, and operating tempo of the new world without pretending it has the metabolism of a startup.

The incumbents have real advantages: data, patients, contracts, regulatory standing, trust, brand, distribution, capital access, clinical expertise, and the authority to put new tools into real workflows. The insurgents have the opposite advantages: speed, youth, technical talent, risk tolerance, iconoclasm, venture capital, and a willingness to ask disrespectful questions. The winning model is pirate plus navy. The pirate alone lacks legitimacy and scale. The navy alone is too slow and ceremonially overdressed. Together, perhaps, they can move.

This means partnerships and acquisitions need to become more aggressive and more discerning. Partner with the frontier labs. Acquire insurgents where it makes sense. Build with firms that have real model access, real engineering talent, and no squeamishness about labor substitution. But don't buy wrappers and call it strategy. The wrapper economy will be brutally commoditized. A pleasant interface sitting on top of someone else's model, with no proprietary workflow, no distribution, no data advantage, no regulatory depth, no liability posture, and no measurable enterprise productivity, isn't a strategy.

The durable advantage is workflow ownership, data access, distribution, governance, and the capacity to turn model capability into enterprise productivity. That last clause matters most. The 150 don't have to out-compute Google, OpenAI, Anthropic, Microsoft, Amazon, Meta, or xAI. They won't. What they can own is the deployment substrate: the workflows, the patients, the clinicians, the claims, the call centers, the care managers, the employed physician base, the trust envelope, the local legitimacy, and the ability to convert intelligence into care. That's their strategic jurisdiction, and they need to stop treating it like a procurement function.

AI diffusion must become a core competency. Not an innovation-office hobby. Not a monthly demo

day. Not a thousand disconnected pilots with no measurement architecture. Diffusion. Installation. Workflow redesign. Backpropagation of local learning into enterprise learning. The institution has to become an AI-learning organism. The operating model should be *tight-loose-tight*: tight on goals, loose on experimentation, tight on measurement. Tight on goals means the leadership team knows what matters—documentation time, denial cost, call volume, prior-auth latency, care-gap closure, access, agency spend, management-layer compression, denominator expansion. Loose on experimentation means teams can test tools, recompose workflows, and find use cases in the fog of war. Tight on measurement means every experiment answers the questions that matter: what changed, what was saved, what broke, what scaled, what got eliminated, what did we learn, and how does that learning propagate?

And yes, democratize experimentation where appropriate. Let a thousand flowers bloom, but don't let a thousand flowers rot into pilot purgatory. Lionize the people who produce measurable productivity gains. Make them internal celebrities! Reward the nurse manager who redesigns discharge follow-up with AI and actually reduces readmissions or call volume. Reward the scheduler who collapses phone volume. Reward the analyst who eliminates a monthly manual report rather than lovingly automating its generation forever. Make productivity improvement prestigious. Healthcare is culturally excellent at venerating compassion, clinical heroism, and mission. It has to learn to venerate labor productivity too, without sounding like it has joined a cult of spreadsheet cruelty.

POCs must scale to production. This is where healthcare historically fails. It admires novelty, convenes panels, issues cautious statements, pilots something promising, and then returns to the workflow exactly as before. That won't do. The 150 need installation capacity: product managers, workflow engineers, model-risk leaders, clinical translators, data stewards, evaluation teams, operational owners, and executives with authority to change staffing. A pilot that can't alter headcount, contractor spend, throughput, quality, access, or patient experience isn't a pilot. It's a demonstration of institutional indecision.

This is why CEO-as-CIO isn't a joke. The chief executive doesn't need to code (although with Claude Code, I think they ought to be coding!). But the chief executive must understand the substitution logic, the agentic logic, the economics of inference, the bottlenecks, the governance risks, and the labor implications deeply enough not to be managed by consultants, vendors, or internal obstructionists. This can't be another technology wave that healthcare abdicates until someone else defines the paradigm. The cultural test is whether the institution can move from admiration to metabolism. Healthcare is very good at admiring technology from a safe distance. We invite the founder, ask the cautious question, praise the pilot, and then watch the immune system kill the thing. AI-naturalization means the technology becomes part of how the enterprise thinks, measures, staffs, procures, trains, governs, and rewards. Anything less is theater.

Build the Healthcare GDPval

Reader Note: GDPval reappears in the hospital CEO playbook. I introduce it here as a labor-mapping method; later it becomes an enterprise operating discipline for seeing the work rather than merely the org chart.

The most important practical artifact the 150 can build is a serious map of the work. Not the org chart. The work. The org chart is an inherited mythology of roles, titles, reporting lines, budgetary

territories, and political accommodations. AI doesn't automate titles. It automates tasks. So the task map comes first.

The practical starting point is a healthcare GDPval. I don't mean a glossy consulting deck with a heat map and a 2x2. I mean a rigorous, occupation-by-occupation, task-by-task, workflow-by-workflow evaluation of what current and imminent AI systems can do relative to expert humans. Use O*NET-style task decomposition, but healthcare-specific. Use real work products. Use blinded comparisons. Use multimodal inputs: notes, claims, images, audio, spreadsheets, call transcripts, denial letters, prior-auth packets, care plans, quality reports, pathology slides, radiology studies, patient messages, scheduling queues. Stop guessing where the work is exposed and measure it.

The healthcare GDPval should be deliberately uncomfortable. Use real work products, blinded human comparisons, expert graders, and multimodal task bundles. The goal isn't to produce a pretty heat map. The goal is to identify which work can be automated now, which work becomes augmentable with oversight, which work should disappear because the workflow is upstream stupidity, and which work becomes newly possible because the marginal cost of expertise falls. If the output doesn't change hiring, outsourcing, capital allocation, compensation, org design, and leadership incentives, it's decorative.

Every task should be classified across four dimensions: automatable, augmentable, eliminable, and newly generative. Automatable means the machine can do it with acceptable oversight. Augmentable means the human remains primary but becomes materially more productive. Elimenable means the workflow itself should disappear because AI changes the upstream process. Newly generative means the technology creates a care or administrative capability that didn't previously exist at viable cost. That fourth category may be the most important, because too much of the current discourse is trapped in the substitution frame: human did X, machine now does X. But the most powerful technological transitions create new Xs. They reveal latent demand, new workflows, new products, new forms of care, new modes of surveillance and support, new kinds of longitudinal engagement that were impossible when every incremental touch required another expensive human minute.

Then map the bottlenecks. Amdahl's Law should discipline the whole exercise. If you automate one step but the system remains constrained by a human decision elsewhere, you haven't transformed the system; you have moved the queue. If documentation gets faster but prior authorization remains slow, the bottleneck moves. If scheduling gets faster but clinician supply remains fixed, the bottleneck moves. If diagnosis gets faster but follow-up capacity remains weak, the bottleneck moves. Transformation is bottleneck hunting. The institutions that win won't be the ones that sprinkle AI across the enterprise like digital parsley. They will be the ones that find the constraints and remove them.

Jevons matters here too. If AI lowers the cost of a service and demand is elastic, utilization rises. In healthcare, latent demand is enormous. If behavioral health becomes cheaper, demand explodes. If primary care access becomes monthly rather than yearly, people will use it. If diagnostics become frictionless, the system will find more risk, more disease, more borderline findings, more follow-up needs. If AI makes outreach cheap, we'll contact patients more. If care management becomes scalable, we'll manage more care. The labor question is therefore not simply "how many jobs disappear?" It's "where does demand reappear once scarcity is relieved?" Human labor should move toward the non-automated *bottlenecks*: trust, adherence, navigation, longitudinal support, complex care, social needs, behavioral health, home-based care, serious illness, caregiver support, patient education, and the messy human interfaces where healthcare actually succeeds or fails.

The GDPval shouldn't become another consultant-produced artifact destined for a shared drive, slowly fossilizing under version-control names like "AI Workforce Exposure Final v7_clean_REVISED.pptx." It should be operational: a living map of tasks, exposure, verifiability, error tolerance, labor cost, technology maturity, bottleneck status, and redeployment path. If it doesn't change hiring, vendor selection, workforce planning, capital allocation, outsourcing, org design, and leadership incentives, it's decorative. The output should be uncomfortable. It should name roles that will shrink, roles that will change, roles that should be consolidated, roles that should be redeployed, and roles that should become more valuable because they're bottlenecks to abundant care. The point isn't prediction as prophecy. The point is to stop flying blind.

The New HR: Humans, Agents, and the Bottleneck Workforce

Reskilling is necessary, but not sufficient. Historically, technological transitions moved through *deskilling* and *reskilling*. The loom deskilled the artisan; the new industrial system then created roles operating, repairing, managing, and improving the machines. This time we need to add a third phase: *non-skilling*. Some work will simply be automated away without a commensurate new job in the same vicinity. We shouldn't lie about that.

That triad—deskilling, reskilling, non-skilling—should be part of the healthcare vocabulary. Deskilling lets lower-cost labor do work previously reserved for higher-cost experts. Reskilling moves people toward the machines, the exceptions, the bottlenecks, and the newly valuable human tasks. Non-skilling is the brutal coda: the work simply disappears without a nearby replacement. Healthcare leaders will be tempted to soothe everyone with reskilling language. They should fund reskilling aggressively. They should also refuse to lie that reskilling absorbs everything.

Reskilling is the comforting word. Non-skilling is the frightening one. Healthcare should fund the first while planning honestly for the second. Some roles will move to better bottlenecks. Some roles will be redesigned. Some roles will simply disappear because the workflow was never care in the first place; it was administrative accretion around a broken financing architecture. Still, the reskilling project matters. Move people from documentation to outreach. From claims chasing to care-gap closure. From call queues to relationship management. From manual reporting to model evaluation. From transaction processing to exception resolution. From prior-auth paperwork to patient navigation. From administrative duplication to human support in the newly abundant care system. Some people will make the transition. Some won't. The plan should be compassionate enough to handle both.

The skills that rise in value will share a few economic properties. They will be hard to substitute, complementary with AI, tied to outputs whose demand expands as price falls, and scarce enough that labor supply can't immediately flood the market. In plainer English: do things AI can't easily do, do things needed to deploy or govern AI, do things the world wants much more of when price falls, and cultivate rare skills that others can't rapidly copy. Healthcare has many such areas if we redesign for them.

For individuals, the advice is uncomfortable but clear. Use the tools aggressively. Become cybernetic. Build social skill, leadership, taste, communication, strategy, and judgment. Avoid long, expensive training pathways whose terminal work is routine application of established knowledge. Save money. Live lightly. Learn fast. Favor smaller, growing organizations where agentic leverage lets one person do the work of a team. Be one step ahead of automation, bank the winnings where you can, and preserve your humanity while becoming far more capable. I know that's a strange injunction. It's also

the world we're entering.

The deeper organizational point is that HR and IT begin to converge. If agentic systems are workers, then workforce planning includes digital labor. If human workers are supervising agents, then training includes evaluation, prompt-craft, escalation, adversarial testing, judgment preservation, and model-risk fluency. The CHRO and CIO should suddenly have much more to talk about. Moderna combining IT and HR wasn't a curiosity; it was a harbinger. The new HR isn't only about hiring, retention, compensation, performance management, and culture. It's about governing a mixed workforce of humans and agents, deciding where synthetic labor substitutes, where it complements, where it creates risk, and where it threatens the developmental ladder by which humans become competent.

This is the biggest organizational metamorphosis since the industrial firm learned to manage the modern workforce. The 150 should treat it accordingly.

The End of Outsourcing, the Work-from-Home Tell, and the Bot War

Outsourcing should be treated as one of the first procurement categories to attack. This isn't because outsourcing firms are bad. Some are excellent, and many have performed necessary work the healthcare enterprise was too inefficient, distracted, or expensive to perform itself. And the people on the other side of this in Manila and Hyderabad deserve the same human compassion and dignity as those lucky enough to be born within our borders. But my point is structural: outsourcing reveals decomposability. If a workflow can be parceled out, contractually specified, monitored through SLAs, performed behind a computer, and managed at distance, it's probably agentifiable.

So every contracted service should be reviewed through the AI lens. What are we paying for that's really document review, data entry, claims follow-up, call handling, scheduling, coding, reporting, analytics, credentialing, compliance compilation, denial management, or knowledge retrieval? What's the tolerance for error? What's the objective function? What data would be required? What human oversight is necessary? What would it cost to build or buy an agentic alternative? The old outsourcing wave moved work to cheaper humans. The agentic wave moves work to cheaper cognition. If you can outsource it, do it behind a computer, and tolerate bounded error with escalation, the burden of proof has shifted. Why isn't this an agentic workflow?

Work from home should be evaluated with similar seriousness. Again, not as a moral complaint about remote workers. The issue is strategic. Remote work revealed which tasks are digitized, decontextualized, and separable from embodied institutional presence. The human work that remains valuable should become more in-person, more social, more relational, more tied to trust and tacit context. The screen-only work will be competed against agents. That doesn't mean punitive return-to-office theater, which is often just managerial nostalgia with badge-swipe data. It means redesigning human roles around the things humans still do better. If the job's value proposition is that it can be done anywhere behind a screen, then its automation exposure is high.

The payer-provider bot war is another transitional hazard. The first chapter of AI in healthcare may be inflationary because existing adversarial systems will weaponize AI before they're redesigned. Providers will use AI to optimize coding and documentation. Payers will use AI to deny, delay, audit, adjudicate, and counter-optimize. My bots counteract your bots. Administrative musculature expands before it collapses. That's a transitional pathology, not the end state.

The second chapter is equilibrium: bots fighting bots, margins shifting, transaction costs still too high but increasingly automated. The third chapter should be deflationary: human labor exfiltrates from the adversarial transaction layer, and some of the artificial friction begins to disappear because the

cost of sustaining it falls toward compute. The danger is that we simply automate the insanity rather than eliminate it. Faster denials, faster appeals, faster documentation, faster coding games, faster friction. That would be a grotesque misuse of a civilization-shaping technology. The opportunity is to use automation to reveal how much of the insanity was never necessary.

That's why governance matters. If AI only makes the payer-provider arms race faster, we'll increase productivity inside a morally defective architecture. If AI lets us redesign the architecture, then labor substitution becomes healthcare reform by other means. The goal should be to remove transaction costs, not merely accelerate the transaction-cost war. If the bots are only fighting over who gets to keep the surplus, the public will be right to revolt.

Use the Bubble

A brief capital-markets aside belongs here because labor substitution won't happen without the current capital mobilization. *We're in a bubble*. The word shouldn't frighten us. Not all bubbles are the same. Some are mean-reversion bubbles, detached from intrinsic value and destined mostly to incinerate capital. Others are infrastructure or inflection bubbles (read the excellent book *Boom* for a real treatise on this). They mobilize capital, talent, risk tolerance, storytelling, FOMO, youth, and definite optimism around a civilizational buildout. Railroads, electrification, telecom, the internet—each had excess, fraud, hype, wreckage, and the usual collection of prophets, charlatans, geniuses, and lucky idiots. Each also built substrate.

The AI bubble is an inflection bubble. It will produce malinvestment, overcapacity in some places, air pockets, absurd valuations, companies with beautiful demos and no business model, and a cathartic scapegoat or two when confidence and cash burn diverge. Fine. Bubbles are races between narrative and solvency. Founders become propagandists because capital markets require belief before the infrastructure exists. Edison understood this. Gates understood this. Elon and Sam understand this. The story pulls capital forward; capital builds the future; some investors get incinerated; society keeps the tracks.

The scale is historically weird. The large hyperscalers are spending at a level that starts to rhyme with national mobilization, and JPMorgan and others have talked about multi-trillion-dollar data-center and AI-infrastructure buildouts by the end of the decade.¹³⁶ Some of this will be malinvestment. Some will be incinerated. Some of it will be absurd. But the substrate will remain: chips, power, cooling, networking, model labs, tooling, robotics, and the industrial base of synthetic labor. Railroads bankrupted investors and reorganized the continent. Fiber incinerated balance sheets and gave us the internet. This bubble can be financially stupid and strategically indispensable at the same time

Healthcare should use the bubble. That sounds a little mercenary because it is. The current flood of capital into models, chips, data centers, energy, tooling, robotics, inference optimization, and agentic infrastructure is creating a substrate healthcare could never build on its own. The 150 shouldn't sit out the mobilization because some valuations are silly. They should exploit it: partner, buy, learn, shape, demand healthcare-specific tooling, and use their data, distribution, and legitimacy while the insurgents still need real-world domains.

When the air pocket comes—and it will—the winners will be the organizations with large balance sheets, clear strategy, and enough technical fluency to go shopping. Many AI companies will fail. Some will have real technology trapped inside broken business models. Some will have talent. Some will have workflows that need distribution. The prepared incumbent can acquire at the trough. The

¹³⁶ Morgan Stanley Institute: "AI Is Now a Macro Variable. Are You Positioned" published on March 9, 2026 ([Link](#))

unprepared incumbent will issue a white paper on responsible AI while someone else buys the future.

Most of private equity should be nervous here, by the way. Capital will matter more, but not all capital is equal. Fifteen years of ZIRP produced a generation of financial engineers very good at leverage gerrymandering and sometimes mediocre at technological diffusion. In the AI era, capital is de-commodified by domain expertise and AI fluency. Healthcare capital allocators who understand labor substitution, functional verifiability, workflow ownership, and AI partnerships will have an advantage. Those holding legacy labor-intensive assets with no credible automation path may discover that “platform company” was just a spreadsheet euphemism. The winners won't be the people with capital alone. They will be the people with capital, domain understanding, AI fluency, workflow access, and enough imagination to see which labor-intensive assets become dead weight and which become more valuable once synthetic cognition is installed. But I digress.

She Who Redeploys the Surplus Humanely, Wins

There's a labor analogue to the liability thesis from the clinical AI chapter: she who redeploys the surplus humanely, wins. The institution that captures labor savings while expanding access, lowering premiums or prices, and treating displaced workers with generosity will have strategic and moral legitimacy. The institution that captures the savings as margin alone may win a quarter and lose the country.

That's not sentimentality at all. I'd call it political economy. A sector that employs one in seven workers—and supports far more through local multipliers—can't behave like a thin software company and expect social permission. Healthcare's oligopolists have a civic burden commensurate with their civic footprint. If the 150 use GenAI to make healthcare more affordable, more abundant, more longitudinal, more humane, and less administratively insane, the transition can be defended. If they use it to enrich incumbents while communities hollow out and patients see no material relief, the backlash will be ferocious and deserved.

So the doctrine has to be explicit. Attack the labor denominator, yes, but do it in the service of abundance. Freeze nonessential hiring, but explain why. Use attrition, but prepare transition paths. Compress management layers, but protect real operators. Invest in compute, not beds, but preserve physical capacity where bodies actually need institutions. Partner with pirates, but keep the moral seriousness of the navy. Build agents, but govern them as a labor class. Automate bad work, but redesign apprenticeship before judgment cavitates. Capture savings, but share enough of the winnings that the public understands this wasn't another elite extraction event with better software.

That's the playbook. It's not gentle, and it's not optional. The 150 aren't being asked to become startups. They're being asked to become serious installation institutions in the most important technological transition of their professional lives. The work isn't to admire AI. The work is to install it, govern it, humanize it, and use it to reconstitute healthcare before someone else does it to healthcare.

And that, finally, is the difference between leadership and panelism.

CODA—BUILD THE MACHINE, CARE FOR THE PEOPLE

The Healthcare Bargain Has to Be Designed

So let me end where this chapter began, with Nightingale rather than Aristotle. *Hospitals are an intermediate stage of civilization.* Healthcare institutions aren't sacred because their current workflows are sacred. They are sacred only insofar as they serve care, healing, trust, and the reduction of suffering. That's the moral premise of this whole chapter. AI gives healthcare a chance to change the form without betraying the purpose.

The opportunity is enormous. AI may make care more affordable, more continuous, more proactive, more longitudinal, more synoptic, and less administratively deranged. It may reduce documentation burden, compress revenue-cycle sludge, make behavioral health more accessible, detect deterioration earlier, close gaps before they become hospitalizations, and make expertise less dependent on ZIP code, income, institutional proximity, and social capital. That's the salvation side of the ledger, and it's very real.

The side effects are real too. AI may dislocate healthcare workers, weaken local economies, shift power from labor to capital, degrade judgment, hollow out apprenticeship, turn humans into validators of machine output, and provoke a backlash that smothers the technology in procedural cement. It may liberate patients from an unaffordable system while injuring workers who helped build that same system. The aggregate line may go up while the particular life goes sideways. That's the cruelty of the transition.

Healthcare doesn't have the luxury of pretending one side of that ledger is fake. The techno-optimists are right that abundance matters. The labor skeptics are right that people get hurt. The guilds are right that safety matters. The insurgents are right that safety can become a pretext for incumbency protection. The economists are right that the lump of labor is usually a fallacy. The workers are right that their own jobs aren't abstractions. All of these things can be true at once, which is why the problem is hard.

This isn't an IT project, a chatbot strategy, or another vendor-friendly occasion for everyone to say “transformation” and then return to the old workflow. It's the *industrialization of intelligence* inside the most labor-addicted sector of the American economy. The old denominator was IT spend. The new denominator is labor. The old question was how to buy software. The new question is how to redesign the production function of care.

The implied covenant—come here, work hard, endure the bureaucracy, acquire the credential, climb the ladder, and healthcare will provide durable middle-class stability—is going to fracture in some roles, some functions, some regions, and some institutions. Pretending otherwise isn't kindness. It's cowardice with a better communications plan. The new covenant should be explicit: fewer humans doing bad work, more humans doing human work, more care at lower cost, and real generosity toward those displaced by the transition.

The Chapter, Compressed

Here is the chapter, compressed. The argument isn't “AI will automate healthcare.” The argument is cleaner and more severe: AI changes the denominator, exposes the labor stack, travels through identifiable mechanisms, and forces the 150 to design a new bargain before the old one breaks in public.

First, the denominator is labor. AI in healthcare isn't primarily about the three-to-five-percent IT

budget. It's about the roughly \$2.9 trillion labor substrate inside a \$5.3 trillion sector. The proper question isn't how much to spend on AI, but how much of the current labor architecture survives when cognition becomes industrially scalable.

Second, shortage is a present reality but an untenable strategy. The physician, nursing, and home-care shortage forecasts are real, but they are also artifacts of a production model that scaled care by adding humans. A shortage is a price signal. Once inference becomes cheap enough and agents become reliable enough, the shortage forecast becomes a map of where capital will attack first.

Third, the exposure is stacked. Healthcare is labor-addicted, labor-expensive, productivity-deprived, administratively overgrown, data-superabundant, biologically complex, and regionally embedded. Its weaknesses are the reason the opportunity is so large: labor addiction creates substitutability; technology debt creates leapfrog potential; data sludge creates the model substrate; biology creates returns to intelligence; unaffordability creates the moral permission structure for substitution.

Fourth, the first wave is administrative cognition. The earliest compression will come where work is verifiable, decontextualized, teachable, and error-bounded: RCM, payer operations, prior authorization, documentation, call centers, scheduling, coding, quality reporting, and the machinery that keeps the payer-provider war moving. Don't start with the surgeon. Start with the claim.

Fifth, mechanism beats occupation labels. A job title doesn't tell you enough. The work has to be decomposed into tasks and tested against the rails of substitution: functional verifiability, decontextualization, teachability, tolerance for error, agentic execution, and bottleneck migration. The map must be built from work products, not vibes.

Sixth, agents are a workforce, not a feature. The managerial imagination has to move from seats, licenses, and dashboards to digital labor: agents that prosecute workflows, learn from outcomes, escalate exceptions, and gradually turn SWB into kWh. HR and IT begin to converge because the workforce now includes humans and non-human workers.

Seventh, the long arc may be expansionary, but the middle is where the pain lives. Jevons matters. Lower-cost expertise can expand demand for monitoring, prevention, behavioral health, chronic-care engagement, and longitudinal support. But new demand doesn't guarantee the old roles survive, nor does it guarantee that displaced workers land in the new bottlenecks at the same wage, in the same place, or on the same timeline.

Eighth, healthcare labor is civic infrastructure. Hospitals and health systems aren't ordinary employers. In many regions they are payroll, procurement, philanthropy, political ballast, and middle-class scaffolding. Labor compression will spill into restaurants, housing, tax bases, unions, community colleges, and local politics. The multiplier isn't academic. It's lived.

Ninth, the human risks aren't ornamental. Cognitive offloading can weaken judgment. Verification can become ceremonial. Apprenticeship can cavitate. Work can become narrower, more surveilled, and more machine-paced. If humans become passive approvers of machine output, healthcare will have automated labor without preserving agency.

Tenth, the 150 need a doctrine. Map the work. Stop reflexive hiring into exposed roles. Build or buy agentic capacity. Redeploy humans toward bottlenecks and genuinely human work. Share enough of the surplus with workers, patients, employers, and communities that the public experiences AI as abundance rather than extraction. Govern the whole thing with enough seriousness that safety is real and not merely a protectionist incantation.

Just for fun, I asked Claude to read this entire, unwieldy chapter, and distill it down to less than 10

words. This is what Claude came up with, and I think it's a clever distillation: *tell, map, stop, build, redeploy, share, and govern*. Tell the truth about labor dislocation. Map the work, not the titles. Stop hiring reflexively into roles whose task content is already exposed. Build the agentic capacity. Redeploy humans toward the bottlenecks. Share the winnings. Govern the machine like a civic institution rather than a procurement department with a press release.

That's the adult thesis. *Build the machine. Care for the people*. Use intelligence to make care abundant, and use the surplus to make the transition humane. Healthcare can't preserve every incumbent role and call that compassion while patients drown in cost. It can't automate aggressively and call that courage while workers and communities are discarded as externalities. The task is synthesis: reconstitute healthcare around abundance without forgetting that abundance is supposed to serve human beings.

Move like builders. Govern like adults. Tell the truth like people who respect their workers. Share the winnings like institutions that understand their civic footprint. And never forget that the point of all this intelligence, if the word still means anything, is more care, better care, cheaper care, less suffering, and a more humane system than the one we inherited.

Build the machine. Care for the people.

That's the only version of this revolution worth defending.

BEFORE WE TURN THE PAGE

Healthcare's exposure is now visible. But exposure is not destiny; leadership is. The next chapter therefore moves from the sectoral diagnosis to the institutional actor who must decide whether the transition becomes purposeful statecraft or merely something done to the enterprise from outside.

Chapter 6 · Future of U.S. Hospitals and Health Systems

Nation-States, CEO Statecraft, and the Diffusion Mandate

"A nation which doesn't shape events through its own sense of purpose eventually will be engulfed in events shaped by others."

— ROCKEFELLER SPECIAL STUDIES PROJECT, DIRECTED BY HENRY KISSINGER (1956–60)

"In the fields of observation, chance favors only the prepared mind."

— LOUIS PASTEUR, 1854

A WORD ON NAVIGATING THIS CHAPTER

This chapter is the memo to the health-system CEO. It treats hospitals and health systems as quasi-sovereign institutions and asks how an organization built for deliberation, credentialing, local politics, physical plant, and human labor can diffuse AI fast enough to avoid being shaped by others.

The chapter moves from frame to mandate to operating model to market structure. First comes the frame: Coase, coordination tax, health systems as nation-states, and diffusion as statecraft. Then comes the CEO mandate: ownership, speed, one enterprise model, workforce upskilling, and the conversion of private productivity into enterprise productivity. Then comes the operating model: labor decomposition, GDPval-like task maps, ontology, org-chart redesign, capex shifting from atoms to bits, and the workforce reset. Finally, it turns outward to care delivery, payer abrasion, M&A, stratification, balance-sheet statecraft, and the first 180-day CEO plan. In other words, this isn't an 'AI strategy' sidecar. It's statecraft for the quasi-sovereign institution.

At first blush, Kissinger and Pasteur may seem like improbable travel companions for a chapter on hospitals and health systems. One belongs to statecraft: sovereignty, purpose, deterrence, legitimacy, self-determination, the choreography of power. The other belongs to scientific preparedness: the disciplined mind, the trained eye, the intellectual posture that turns apparent accident into discovery. But that's exactly why they belong here. The health-system CEO now needs both. Kissinger gives us the strategic obligation of large institutions: shape events through purpose or be engulfed by events shaped elsewhere. Pasteur gives us the epistemic posture for an exponential age: chance, discontinuity, and technological surprise favor the prepared mind.

Honestly, I might have started this chapter with the line often attributed, likely apocryphally, to Einstein: "If you can't explain something simply, you don't understand it well enough." That may be the more embarrassing and more appropriate epigraph, because it has taken me a mortifying number of pages to explain what I think health-system strategy should be in the AI era, and the length itself is a confession. I'm not writing from the complacent altitude of settled doctrine. I don't yet understand the right strategy cleanly enough to render it in one page; maybe no one does. We're early. The model capabilities are jagged. The workflows are messy. The liability architecture is immature. The labor implications are morally uncomfortable. The regulatory moat buys time but not immunity. And whenever we think we've arrived at some stable interpretive lens, the next model release, policy shift, market convulsion, or Chinese open-weight announcement

detonates the framework from underneath.

So take what follows not as doctrine, not as immutable law, not as some Decalogue descending from the cloud. Take it as a set of provisional postulates, operating hypotheses, pattern recognitions, and, yes, intuitions from someone trying to see around the corner with you. The point isn't to pretend to omniscience. The point is to cultivate the prepared mind, move under uncertainty, and iterate faster than institutions like ours have historically been willing or able to do.

This chapter is an operating memo to the Health System CEO. I won't subject you again to the broad historical prologue that sits in the opening essay: the Republic of Letters, the lessons of technological diffusion, the half-life of strategic frameworks, the inventor-versus-installer distinction, Hayek's dispersed-knowledge problem, the individual-versus-enterprise productivity paradox, and the insistence that incumbents have less time than they think. I still love those ideas, and they have relevancy for us here. But enough philosophizing. It's time to get tactical. This chapter begins where the hospital CEO actually has to begin: with the institution in hand, the board in front of her, the operating model beneath her, the workforce around her, the community watching her, and the models improving faster than any normal health-system planning cycle can tolerate.

That's why this chapter has to be addressed to the Health System CEO. Not to the innovation committee, not to the digital steering group, not to the AI governance council, and not even to the CIO, though the CIO will matter enormously and may become either one of the protagonists of this transition or one of its subtlest, if well-intentioned, saboteurs. This is written for the person who can still command the whole organism while there's still time: the person who can decide that AI is no longer a side project, a pilot, a sandbox, or an anodyne innovation-center artifact, but a mandate.

The AI-alert CEOs have already autocorrected. They are, to slip in a Matrix meme, red-pilled. They don't need to be proselytized that generative AI is real, that frontier models are sprinting forward at a civilization-altering pace, that agents are here, that administrative work is exposed, or that the old proof-of-concept metabolism is too slow. Good. Necessary. But also irrelevant now. The question is what you do next, while the models are still jagged, the use cases still embryonic, the enterprise still confused, the regulatory moat still buys a little time, and the asymmetric first-mover advantage hasn't yet evaporated into the banal air of everybody-has-it-now.

This chapter ends with an unambiguous CEO plan: what to do in the first 180 days, what to stop doing, what to build, what to buy, what to partner on, what to measure, and how to convert private productivity into enterprise productivity. I will get there plainly, if eventually. For the impatient among you, go ahead and skip to the end and read the 15-point plan. But I selfishly encourage you to read this all in sequence, because the plan only makes sense if the architecture underneath it is clear.

The CEO's Map of the Chapter

The argument moves in six parts. **First**, the frame: Coase, coordination costs, health systems as quasi-sovereign institutions, and the historical lesson, already elaborated in the opening chapter, that power belongs less to *inventors* than to *diffusers*. **Second**, the CEO mandate: why AI can't be delegated downward, why the CEO has to govern with speed and decisiveness, why one frontier model and agentic harness should be diffused broadly, why the organization needs a *ministry of education* rather than a training module, and why the CEO now has to behave, at least in this domain, as something uncomfortably close to a CTO. **Third**, the operating model: labor substitution, workflow decomposition, GDPval-like task maps, data ontology, org-chart redesign, capex shifting from *atoms to bits*, and the workforce reset. **Fourth**, the care-delivery payoff: not a second, inferior clinical-AI chapter, but the CEO implication of the clinical-AI argument elsewhere in this essay: de-bureaucratize first, build diffusion muscle, then move to clinical AI quickly where the business model and liability architecture make its value legible. **Fifth**, the market consequences: payer-provider abrasion, health-system stratification, non-contiguous M&A, AMCs and RCM incumbents under pressure, and a much more serious version of what it means to weaponize the balance sheet. **Sixth**, the plan: a hard, prescriptive, measurable 180-day punch list.

The architecture is deliberately causal: frame before mandate, mandate before operating model, operating model before care redesign, care redesign before market structure, and market structure before the 180-day plan. If one of these sections feels like a detour, that's the reason it's here: each piece answers one layer of the same statecraft question—how does a large, trusted, slow-moving, labor-heavy civic institution become an *AI-diffusing operating company* before someone else turns it into a tenant of their model?

The route, naturally, given your humble author's unconquerable proclivities, will still meander through Coase, statecraft, Silicon Valley, the Gulf, the industrial revolutions, the 1855 New York and Erie Railroad org chart, honeybees, *Tanzsprache*, and a few other historical and metaphysical tributaries. But the structure isn't ornamental. Each detour is meant to answer the same practical question: what should the CEO of a large health system do *now*, while asymmetric advantage is still available?

The reason to care about that tactical plan is that this isn't merely a technology strategy. It's an institutional survival strategy. An *anti-disintermediation* strategy. The coming contest among health systems won't be decided by who *invents* AI. That race is elsewhere. It will be decided by who *diffuses* it fastest, most horizontally, most safely, and most economically through the enterprise. Hegemony in technological revolutions rarely belongs to the first person who glimpses the machine. It belongs to the sovereign, company, institution, or CEO that installs the machine into the metabolism of ordinary life.

One last note before the frame. Hospitals and health systems, whether they like the comparison or not, are *now large enough to think of themselves as something closer to nation-states than ordinary firms*. They govern populations. They allocate scarce resources. They employ millions. They manage legitimacy. They negotiate with other sovereign power centers. They sit on real estate, data, capital, labor, philanthropy, clinical knowledge, public trust, and political influence. In many regions, they are the dominant institutional fact of civic life after government itself. Which is why this is no longer just a management problem.

It's a statecraft problem.

Part I: The Frame—Coase, Statecraft, and Diffusion

This first part gives the chapter its operating frame. If the health system is now a quasi-sovereign institution, then its CEO has to think less like a departmental manager and more like a head of state: What territory do I govern? What population do I serve? What legitimacy do I need to preserve? What foreign powers—payers, regulators, hyperscalers, frontier labs, unions, capital markets, insurgents—am I negotiating with? What infrastructure must I build? What bureaucracy must I reform? What technology must I diffuse before someone else diffuses it into me?

The frame starts with Coase because the firm exists to solve coordination problems, and American healthcare is now mostly *coordination tax*. It then moves to the nation-state analogy because health systems are no longer ordinary operating companies; they are civic institutions with geopolitical-like constraints. And finally it moves to diffusion because the winner in this technological revolution won't be the institution that admires AI most eloquently. It'll be the institution that installs it. This part is therefore not throat-clearing. It is the conceptual permission structure for the rest of the chapter: once coordination costs fall, sovereignty, scale, labor, capital, and clinical design all have to be rethought.

What Coase Saw, and What AI May Finally Change

Let me begin by disinterring a dusty essay from almost a century ago that suddenly speaks to us with surprising relevance. In 1937, Nobel laureate Ronald Coase wrote *The Nature of the Firm* and asked a question that looks almost too obvious to be profound:¹³⁷ if markets are so powerful, why do firms or organizations exist at all? Why isn't economic life simply an endless chain of spot contracts? Coase's answer was that the *market is costly to use*. There are search costs, contracting costs, negotiation costs, monitoring costs, enforcement costs, dispute costs, and the thousand little frictions that arise when humans and institutions try to coordinate across boundaries. The firm—the organization—exists because, beyond some point, it's cheaper to coordinate activity administratively than to transact for it endlessly in the open market. The firm is a machine for reducing the *coordination tax*.

That frame is almost embarrassingly clarifying for American healthcare. We're a \$5 trillion-plus healthcare economy, roughly 18% of GDP, with hospital care alone weighing in at \$1.6 trillion, and yet the signature emotional experience of the system is *friction*. The patient can't get through. The physician can't find the information. The payer can't understand the claim, or pretends not to. The nurse can't escape the documentation. The revenue-cycle team can't get paid without ritual combat. The whole system is saturated with handoffs, portals, authorizations, denials, appeals, eligibility checks, coding rules, compliance rituals, work queues, resubmissions, reconciliations, and administrative ceremonies that exist because the coordination fabric is broken—broken to the tune of roughly \$1 trillion in *coordination-tax expenditure* per year.

This is why generative AI matters in healthcare in a deeper way than the usual product language suggests. The first wave isn't really about pithier emails, faster chart summaries, or one more documentation assistant, useful though those things are. Those are surface manifestations. The real significance is that models and agents may reduce the cost of searching for information, interpreting unstructured records, translating across incompatible systems, navigating rules, drafting appeals, preserving context across handoffs, routing work, supervising queues, and making one institutional actor legible to another. In plain English: **AI may begin to dissolve the coordination tax that justified the administrative superstructure of modern healthcare in the first place.**

¹³⁷ Ronald H. Coase: "The Nature of the Firm" published in November 1937 ([Link](#))

That's the Coasean claim. Hospitals became enormous in part because the coordination costs of American healthcare became enormous. AI now begins to compress those costs. And once that happens, the question isn't whether a model can automate a single task inside the current enterprise. The question is whether the very reasons we organized healthcare firms the way we did are beginning to erode. This theme—the mitigation or elimination of the collective coordination tax—will permeate this entire chapter.

Let me make the implication sharper. If the firm exists because coordination is expensive, then a technology that makes coordination radically cheaper changes the boundary of the firm. Some things that were centralized may become distributed. Some things that were local may become centralized. Some things that required managerial hierarchy may require agentic orchestration. Some things outsourced to BPOs may be repatriated into agentic workflows. Some functions that were defended as “strategic” may reveal themselves as clerical artifacts of pre-AI transaction costs. Coase doesn't give us a tidy answer, but he gives us the right question: where should coordination live once machine intelligence makes previously expensive coordination cheap?

That's not a CIO question. That's a CEO question.

Why the Nation-State Analogy Matters

One last metaphor, or perhaps one last frame, before we get tactical. Yes, I'm going there: I'm analogizing hospitals and health systems to nation-states. I know this sounds a little grandiose, maybe even self-indulgent, but the analogy isn't ornamental. *It forces the right unit of analysis.* The claim isn't that health systems have armies or flags or foreign ministries, though some have enough internal politics to make the Holy Roman Empire look administratively crisp. The claim is that, at sufficient scale, the largest systems behave less like ordinary firms and more like sovereign institutional actors. They govern populations, command workforces, allocate scarce capital and labor, absorb shocks, negotiate with other power centers, manage legitimacy, and decide how fast a region experiences technological change.

Hospital care remains the single largest spending category in American healthcare—roughly \$1.6 trillion, or about 5.5% of U.S. GDP. **The top 100 health systems alone represent nearly \$1.2 trillion in annual revenue, more than 70% of the hospital sector; the top 10 control approximately \$445 billion.** These aren't quaint charitable institutions with a mission statement and a few ambulatory annexes. By sheer scale, more than *fifty* would qualify for inclusion in the Fortune 500. They are vast operating systems: balance sheets measured in the tens of billions, more than \$1 trillion in real estate, sophisticated philanthropic enterprises, employed physician bases that encompass over half of all U.S. doctors, and deep reservoirs of political influence, labor-market power, data exhaust, and institutional memory.

The nation-state analogy gives the CEO a better set of questions. What is your population strategy? What is your civil service? What is your intelligence function? What is your treasury strategy? What is your labor compact? What is your foreign policy toward payers, regulators, unions, hyperscalers, frontier labs, and insurgent companies? What is your infrastructure plan? What is your legitimacy strategy? What do you build yourself, what do you import, what do you ally around, and what do you deter?

Those questions sound dramatic until you remember the scale of these institutions. In many regions, the health system is the largest employer, the major real-estate actor, the philanthropic center, the physician platform, the public-health backstop, the graduate-medical-education engine, the political

bargaining unit, and the site where the public discovers, in the most intimate way, whether the institution can be trusted. That's not an ordinary firm.

The analogy also clarifies the historical lesson. Who wins after industrial revolutions? Not usually the first person to invent the new technology, nor the first company to demo it, nor the first minister to admire it at a polite distance. The winners are *diffusers*. The first three industrial revolutions—*mechanization*, *electrification*, and *computerization*—changed the world only when societies installed those technologies broadly enough to reorganize production, labor, capital, education, management, and daily habit. I belabored this notion enough in my opening chapter, so I'll simply say the same is true here. A health system that doesn't shape this transition through institutional will eventually will be engulfed by a transition shaped by frontier labs, payers, regulators, insurgents, capital markets, autocratic diffusion regimes abroad, and more aggressive peer systems.

That's why "nation-state" isn't meant as a flattering analogy. It's a burden. A nation-state has to decide. It has to mobilize. It has to make tradeoffs. It has to govern. It can't endlessly admire the external environment while the external environment rearranges its sovereignty. Health systems now face a similar problem. The question is whether they remain sovereign over the production function of care or become dependent provinces inside someone else's AI operating model.

Diffusion, Not Invention

One last elaboration on this installation theme. Most people in healthcare still think the central question is *invention*. Which model is best? Which startup has the sharpest demo? Which ambient vendor, coding vendor, triage workflow, agentic tool, or frontier lab? These questions matter, but they aren't the main event. We've invention. We're *drowning* in invention. What healthcare has always been bad at—spectacularly, almost comically bad at—is diffusion, assimilation, installation, and the conversion of scattered technical possibility into institutional habit.

That's the historical personality of the sector. We were slow on the *internet*, slow on *mobile*, slow on *cloud*, slow on *analytics*, slow on *enterprise SaaS*, slow on *remote monitoring*, and slow on almost every technology wave that reorganized other industries. Healthcare could get away with it because it was insulated by regulation, reimbursement, local market power, credentialism, physical care delivery, and the stubborn fact that people still get sick in specific geographies and need beds, nurses, physicians, devices, blood, drugs, and doors that open at 1 a.m.

Not this time. Or, more precisely, not *indefinitely* this time. Healthcare still has a short, time-delimited reprieve: liability, verification, trust, HIPAA, licensure, reimbursement, procurement, cyber, and the physicality of care all slow deployment. *Capability and deployment aren't synonymous*. But a reprieve isn't immunity. The correct inference isn't "we can wait." The correct inference is "we've been granted a narrow window to prepare while diffusion is still hard."

This is the prepared-mind point. The technology is moving so quickly that even its prime progenitors struggle to keep up—listen to the podcast Dario and I did, where he rather endearingly admits it's hard even for *him* to keep pace. The answer isn't panic. It's facility. Learn the tools. Learn the nomenclature. Study the exponentials. Stop treating ignorance of AI as a harmless executive eccentricity. We don't get to say, "I'm not really a technology person," anymore. Not with a straight face. Not if you're the CEO.

And this is where the argument turns from diffusion as a management problem to diffusion as a structural problem. Healthcare is slow not merely because its leaders are cautious, though many are. It's slow because the sector is wrapped in an almost geological accumulation of rules, permissions,

liabilities, local customs, accreditation rituals, reimbursement traps, and institutional veto points. That thicket has historically made the sector maddeningly resistant to change. In this one peculiar moment, it also buys time. But time, badly interpreted, becomes anesthesia.

The correct posture isn't "we're protected." It's "we've a temporary moat, and therefore a temporary obligation." Use the moat to prepare the institution. Use the moat to build governance. Use the moat to teach the workforce. Use the moat to map the tasks. Use the moat to create safe enterprise access. Use the moat to partner with the labs before the labs decide they don't need you. *Use the moat to become a diffuser before diffusion becomes easy enough that your advantage is gone.*

The Regulatory Moat, and the View from Beijing and Riyadh

Reader Note: The regulatory-moat / Beijing / Riyadh comparison repeats material from the Clinical AI and China chapters. I repeat it here because the hospital CEO needs to see regulation not only as protection, but as a timing risk in a global installation race.

Healthcare's regulatory moat *is* real. This is one of the most regulated sectors in America—with nuclear power raising its hand from the back of the room—and the numbers are almost comically overwhelming. Mercatus and QuantGov identified 805,817 state healthcare regulatory restrictions across 44 states and the District of Columbia as of July 2020¹³⁸—restrictions meaning textual injunctions like "shall," "must," "may not," "prohibited," and "required." At the federal level, their Healthcare RegData work identified another 49,312 healthcare regulatory restrictions in the 2018 Code of Federal Regulations.¹³⁹ So before we even get to municipal rules, sub-regulatory guidance, accreditation standards, medical-board requirements, payer policies, HIPAA interpretations, Stark, Anti-Kickback, EMTALA, state licensure rules, certificate-of-need regimes, privacy law, facility codes, cybersecurity standards, and the whole cathedral of CMS guidance, we're already staring at roughly 855,000 state and federal healthcare regulatory injunctions. Round it carefully, and we're in the neighborhood of 900,000 healthcare "shall/must" obligations before the softer but still coercive layers of the system are even counted. We don't function in anything approximating a normal operating environment. This is regulatory kudzu.

This matters because complexity slows diffusion. The hospital CEO doesn't operate in the same world as the software founder or the consumer-platform executive. Every meaningful clinical or operational change has to pass through licensure, privacy, reimbursement, malpractice, accreditation, medical-staff politics, EHR constraints, payer contracts, state law, federal law, and local institutional custom. That makes healthcare maddeningly slow. But in this one peculiar moment, slowness also buys a little time. *The regulatory thicket is a moat.* It gives hospitals and health systems a reprieve that less protected sectors—software, media, advertising, professional services—don't have. And all one has to do is glance at the WSJ front page to see the convulsions rocking each of these less-regulated sectors weekly.

But a reprieve isn't immunity. The danger is that extra time gets misread as safety. It isn't safety. It's only a delay in the arrival of the reckoning. And the delay isn't global. China and the GCC will almost certainly move faster in clinical AI because they have fewer veto points, more concentrated *state*

¹³⁸ Mercatus Center – George Mason University: "State Healthcare RegData 1.0: A Quantification of State Healthcare Regulations" published on October 14, 2020 ([Link](#))

¹³⁹ Mercatus Center – George Mason University: "Healthcare RegData: Quantifying the Volume of Healthcare Regulations" published on April 15, 2020 ([Link](#))

capacity, more permissive data regimes, and a different appetite for risk. Their advantage isn't that they have better values. It's that they have fewer procedural brakes.

That's why the **Cleveland Clinic–G42 collaboration** matters—and as an aside, have a listen to my podcast with Tom Mihaljevic for more on how he thinks about this.¹⁴⁰ This wasn't just another decorative international partnership announcement, or another ribbon-cutting exercise with better lighting. It was a signal from one of the world's great clinical institutions that some of the future may be built in jurisdictions where the data, regulatory, capital, and state-capacity environment permits faster movement. In 2025, Cleveland Clinic and Abu Dhabi-based G42 announced a joint task force to evaluate, prioritize, and accelerate AI projects in healthcare;¹⁴¹ G42's healthcare portfolio company, M42, spans 480 facilities in 27 countries and leads initiatives such as the Emirati Genome Program and Malaffi, Abu Dhabi's health information exchange. That's precisely the kind of environment where clinical AI may diffuse faster than it can inside the American regulatory thicket.

So yes, *you can see pieces of the future of clinical AI in Beijing and Riyadh*. That doesn't mean we should import their politics. It means we should *study their diffusion*. Clinical innovation may happen faster in the Gulf or China than in the United States, and some of it may be *reverse-imported* back into American medicine. The American regulatory moat should therefore produce urgency, not complacency. It gives our incumbents time to prepare their organizations, build safe governance, partner with frontier labs, ontologize and vector-enable their data, and develop the muscle memory of *iterative deployment*. It doesn't guarantee they'll use that time well.

And that brings us back to the institution itself. If diffusion is the contest, and if the moat merely gives us time, then the decisive variable is leadership. Not governance in the ornamental sense. Not innovation in the theatrical sense. *Leadership*. Someone has to make this consequential enough for the whole organism to move.

Part II: The CEO Mandate—Ownership, Speed, and Diffusion

This part turns from *frame* to *command*. The institution won't be transformed by ambient enthusiasm, innovation-center theatre, or a thousand local experiments with no scaling path. It will be transformed only if the CEO makes AI one of the two or three central operating imperatives of the enterprise. The CEO doesn't need to become a machine-learning engineer, though with Claude Code perhaps we should all lower our excuses. But the CEO does need enough fluency to allocate capital, compress bureaucracy, adjudicate internal resistance, govern risk, select partners, and hold the organization accountable for measurable deployment.

The thesis of this part is simple: AI won't transform health systems. CEOs will. This part exists because diffusion is a leadership problem before it is a technical problem. The technology can be purchased. The operating model has to be commanded.

AI Won't Transform Health Systems. CEOs Will.

AI won't transform health systems. *CEOs will*. Tools don't diffuse themselves into large, risk-averse, professionally guarded institutions. Whatever the CEO *personally* and *visibly* prioritizes is what the organization will deliver. If AI isn't one of the two or three things that matter at the board and ELT level, the system will produce pilots, slide decks, governance charters, and polite theater while better-mobilized peers pull ahead.

¹⁴⁰ Incumbents and Insurgents: "Tom Mihaljevic and Elad Walach" published on April 17, 2026 (Link)

¹⁴¹ Cleveland Clinic: "Cleveland Clinic and G42 to Advance Healthcare through Artificial Intelligence" published on March 25, 2025 (Link)

This can't be delegated downward. The CEO has to own the transformation personally: enough technical fluency to make astute capital-allocation decisions, enough operational authority to adjudicate resource conflicts, enough courage to stop legacy work, and enough accountability to tie AI outcomes to the board, the ELT, compensation, promotion, and resource allocation. This isn't a CIO project, a consultant project, or an innovation-center project. *The CIO is indispensable; the CEO is accountable.*

The governance model has to be fast, not senatorial. A small cross-functional group—clinician, finance, operations, legal, technologist—*chaired by the CEO* or by a true designee with direct board visibility, should set policy and remove blockers. If the approval process takes longer than building the pilot, the structure has failed. Then build a transformation unit, not an innovation center: an operating group that maps cost structure, identifies functionally verifiable work—yes, that recurrent leitmotif—evaluates the build-buy-partner landscape, and drives enterprise deployment. Start non-clinical, build trust through outcomes, and then extend (quickly) into clinical domains with evidence gates, validation, regulatory pathway mapping, workforce planning, and radical transparency with clinicians. Measure everything: cost-to-serve trajectory, quality, speed, workforce impact, adoption, deployment, and financial outcomes. Tie rewards to results. Every successful pilot needs a predetermined path to scale.

This is why the CEO mandate has to come before the tooling discussion. The tools will change. The vendors will change. The models will improve, cheapen, and proliferate. But the institutional bottleneck will remain the same: will the CEO make AI a governing priority, or will the organization metabolize it into the old machinery of committees, pilots, and gently exhausted consensus?

There's a board implication here too. Boards need to stop treating AI literacy as optional decoration. If AI is a labor, capital, clinical, cyber, liability, and strategy issue, then a board that can't ask intelligent questions about model access, data sovereignty, workflow exposure, labor substitution, clinical validation, and capex reallocation isn't governing the enterprise. *It's observing it.* Every board should have an AI operating dashboard, an AI risk dashboard, and an AI labor-exposure dashboard. Not because dashboards save us, God help us, but because boards govern what they can see. Right now too many boards can see philanthropy, debt, days cash on hand, nursing turnover, quality events, payer mix, and strategic-plan adjectives. They can't yet see the emerging production function. That has to change.

Before the CEO posture comes the informational problem, because this is where the hospital chapter would otherwise feel too managerial and not structural enough. AI first superpowers the individual. The CEO's task is to convert that private productivity into enterprise productivity. That is why this detour belongs here, and why Friedrich Hayek—of all people—turns out to be useful inside a chapter about hospitals and health systems.

From Private Productivity to Enterprise Productivity

There is, however, a strange and slightly paradoxical feature of this first chapter of diffusion. I think there's something intrinsic to this technology that makes *individuals* super-powered, *but not yet enterprises*. Think of the strange alchemy that happens with GenAI. The alchemists of yore—including some of the most generational, even civilizational, minds like Sir Isaac Newton—were alchemists in their spare time (literally). Alongside the development of Newtonian physics and calculus, they experimented with the ancient dream of transmuting base metals into gold. Their mythological quest was the *Philosopher's Stone*, the substance that could convert base metals into gold, something prosaic into something magnificent. They never found it. But in a strange way we may have stumbled onto something even more magical. As Marc Andreessen recently analogized, in the early twenty-first century we learned how to transmute the most superabundant commodity on the planet—*sand*—into the rarest and most powerful thing we know: *thought*.¹⁴²

And with this new Philosopher's Stone, the LLM, something extraordinary happens. *One person becomes many*. One person becomes polymathic. One person suddenly has access to the ingested corpus of human civilization—5,000 years of recorded knowledge and the intellectual residue of roughly 117 billion humans—compressed into a form factor that fits comfortably in one's pocket. *One to many*. One individual, fortified with this synthetic intelligence partner, begins to approach something like quasi-omniscience and can plausibly multiply their productivity by an order of magnitude. The lone analyst becomes a research team. The solo programmer becomes an engineering department. The physician becomes a walking medical library.

And yet here we encounter a paradox. If individuals are suddenly becoming so dramatically more capable, why aren't enterprises experiencing the same commensurate leap in productivity? Why does the individual suddenly feel expanded, polymathic, omni-disciplinary—while the organization around them remains stubbornly unchanged?

Part of the answer is practical. Enterprises still haven't figured out how to fully ontologize their data, ensure data sovereignty, and prevent the exfiltration of sensitive corporate information. Naturally organizations proceed with caution. But I suspect something more foundational is at work here—something structural about the nature of organizations themselves.

What's the purpose of a corporation, or an enterprise, or an organization, after all? The aggregation of many human minds into one of these shared-purpose-defined groupings exists to do the *opposite* of what an LLM empowers the individual to do. The individual augmented by AI operates in a *one-to-many* configuration: a single mind amplified into many capabilities. The corporation, by contrast, is fundamentally a *many-to-one* system. It aggregates many individual talents and organizes them toward a single objective—or at least a manageable set of objectives. Legal, marketing, finance, engineering, product, operations: these functions exist precisely to coordinate and channel many disparate capabilities into a unified strategic direction.

But the technology itself behaves in the opposite direction. It's intrinsically one-to-many, distributing cognitive power outward rather than concentrating it inward. There is therefore something structurally oppositional between the way corporations are organized and the way this technology initially expresses its power.

Which leads to an intriguing observation about this first chapter of AI diffusion: the technology initially acts as a *centrifugal* force, pushing power outward to the edges of organizations rather than

¹⁴² Marc Andreessen: "The Techno-Optimist Manifesto" published on October 16, 2023 (Link)

reinforcing it at the center. Individuals become dramatically more capable before institutions do. Knowledge workers equipped with these systems suddenly operate with far greater autonomy and effectiveness, even while the enterprises employing them struggle to translate those gains into system-wide productivity improvements.

That dynamic won't persist forever. Eventually organizations will learn how to capture and coordinate those gains, largely through the automation and restructuring of white-collar workflows. When that happens the vector of power will reverse. The same technology that initially empowered individuals will become a centripetal force, drawing capability back toward the institutional center as organizations reorganize themselves around machine intelligence.

But for the moment we are living in the centrifugal phase of the revolution. The individual has discovered the Philosopher's Stone before the institution has figured out how to refine it. And that's precisely the tension that makes the next section, oddly enough, a Friedrich Hayek story.

Friedrich Hayek in the Hospital: When Dispersed Knowledge Becomes Machine-Legible

The private-productivity problem leads, oddly enough, to Friedrich Hayek writing in 1945. If the first chapter of diffusion makes individuals super-powered while enterprises remain sluggish, the second chapter may begin when the *institution learns how to capture the knowledge at its own edges*. One person, fortified with a language model, suddenly becomes polymathic, capable of synthesizing information and executing tasks that once required entire teams. The individual becomes dramatically more capable while the enterprise itself struggles to translate those gains into system-wide productivity improvements. But that centrifugal phase may only be the opening act of a longer story, because as the technology matures the direction of force may begin to reverse.¹⁴³

In his seminal essay "The Use of Knowledge in Society," Hayek articulated one of the most elegant explanations ever written for why decentralized markets outperform centralized planning. The central economic problem, he argued, isn't simply allocating known resources efficiently but capturing and *coordinating dispersed knowledge*—the tiny fragments of information about preferences, scarcity, opportunity, and local conditions that exist only in the minds of millions of individuals operating at the frontier of the economy. No central planner, however intelligent or well intentioned, can possibly aggregate all of that tacit knowledge in real time. The genius of the price system is that it functions as a communication network, condensing enormous quantities of distributed information into simple signals such as rising or falling prices. Markets coordinate human behavior spontaneously, allowing complex systems to adapt without requiring anyone to understand the entirety of the organism.

That insight became one of the intellectual foundations explaining why centrally planned economies were brittle and sclerotic. The politburo or the central committee might possess authority, but it lacked the informational bandwidth necessary to allocate resources intelligently. Leaders at the apex of those systems could only *perceive a narrow slice of the underlying reality*, while the granular knowledge embedded in millions of local decisions remained inaccessible. Central planning failed not merely because leaders were incompetent or malicious or stupid—though history offers plenty of examples of all three—but because the information required to run a complex system could not be centralized. The frontier knowledge remained dispersed, and the center never truly saw what the periphery knew.

The arrival of machine intelligence forces us to revisit Hayek's insight from a new angle. What

¹⁴³ Friedrich A. Hayek: "The Use of Knowledge in Society" published in September 1945 ([Link](#))

happens when the technology itself begins capturing and synthesizing exactly the kind of dispersed knowledge that Hayek believed could never be centralized? Imagine an enterprise in which intelligent agents are embedded across the operational surface area of the organization—product development, engineering, finance, HR, legal, sales, operations—each continuously observing activity, extracting signals, and summarizing the informational frontier of the enterprise. Now imagine an Ultron agent sitting at the center of the organization, pulling together those signals from every corner of activity into a continuously updated synthesis of what is actually happening inside the institution.

In that environment the informational position of leadership changes dramatically. The CEO—or the leadership team—no longer operates entirely through filtered reports and bureaucratic intermediaries. Instead they gain access to something approaching a continuously updated map of the organization's operational reality. *The leader at the apex becomes far closer to omniscient than any leader in history has ever been.* Perfect information remains a theoretical fiction, of course, but radically improved information isn't. Once intelligent agents diffuse throughout the enterprise and continuously aggregate frontier knowledge, leadership no longer operates in informational darkness. The signals that once lived only at the edges of the organization—inside product teams, sales conversations, engineering threads, operational bottlenecks—begin to flow inward in ways that were simply impossible in the managerial systems of the twentieth century.

Seen through this lens, the technology that initially behaved as a centrifugal force, empowering individuals, may ultimately become a centripetal force that strengthens the informational capacity of institutions themselves. The individual worker may discover the Philosopher's Stone first, but once the organization learns how to orchestrate that intelligence infrastructure across its workflows the center of gravity begins to shift inward again. Leadership historically required delegation because the cognitive and informational limits of human beings made centralized oversight impossible. A CEO couldn't attend every meeting, read every engineering thread, monitor every operational anomaly, or understand every regulatory nuance across a large enterprise. Those constraints were artifacts of biological cognition, and once machine intelligence begins aggregating the signals from the frontier of the organization those limits begin to soften.

Hayek's insight therefore remains correct but incomplete for this new environment. The dispersed knowledge problem doesn't disappear, yet the mechanisms available for aggregating that knowledge change dramatically when intelligence itself becomes computational. If information from the edges of an organization can be continuously captured, synthesized, and transmitted to the center, then the leadership problem begins to look very different from the one Hayek was describing in 1945. The CEO of the future may operate less like a distant coordinator of semi-autonomous silos and more like the conductor of a highly instrumented system whose internal signals are visible in real time.

And once leaders can see more clearly, they are expected to decide more decisively. The informational excuse for institutional inertia begins to weaken, and the leadership model required to navigate technological upheaval begins to change accordingly. That realization is exactly why Hayek belongs in a hospital-and-health-system chapter: if machine intelligence reshapes the informational architecture of the enterprise, it inevitably reshapes the kind of CEO statecraft the enterprise requires.

CEO as Autocrat

So let's take this to the logical conclusion, and here I'm deliberately going to use an uncomfortable word: *the CEO has to become autocratic*. Not because autocracy or authoritarianism is admirable, and not because healthcare should import the worst instincts of Silicon Valley *founder-mode* machismo. Hospitals are civic institutions, trust-dependent institutions, moral communities, and regional labor anchors. But diffusion against bureaucratic inertia requires muscular executive will. Committees don't create technological revolutions. Consensus doesn't compress a decade of adoption into eighteen months.

This will feel alien because it contradicts the nonprofit healthcare ethos: *first do no harm*, sacred covenant with workers, local autonomy, subsidiarity, physician fiefdoms, committee governance, pilots, proof points, and *everyone gets a say*. There's wisdom in that ethos. There's also paralysis in it. The same habits that once looked like prudence in peacetime can become, in moments of discontinuity, a narcotic.

The mandate has to be clear: *AI competency isn't optional for the senior team*. Everyone has to use the tools. Everyone has to learn enough to understand how their domain changes. *The senior team becomes the first missionary corps*, then their lieutenants, then their lieutenants' lieutenants, until the process daisy-chains through the enterprise. You'll need authenticity and vulnerability, because leadership pretending to be omniscient will kill experimentation. But after the vulnerability comes the mandate: *we're learning, and therefore we're moving*.

And that distinction matters. Humility can't become drift. Saying "we're early" can't become a euphemism for institutional immobility. The CEO's job is to create a strange and difficult combination: *humility about the map, but decisiveness about the march*. We don't know everything yet. Fine. No one does. But we know enough to begin, and the organizations that begin seriously will compound learning while the organizations still "evaluating the landscape" will compound excuses.

The autocracy I'm describing is temporary, bounded, and strategic. It's autocracy about priority, not autocracy about truth. The CEO shouldn't pretend to know exactly which model wins, which vendor survives, which workflow scales fastest, or which clinical application clears the liability threshold first. But the CEO should be autocratic about the operating premise: AI is now a central production technology, and the enterprise will learn, measure, diffuse, and redesign around it. That premise isn't up for monthly re-litigation by every committee with a clever objection.

Healthcare has confused participation with governance for too long. Participation matters. Clinician input matters. Worker input matters. Patient trust matters. But an institution can't ask every incumbent workflow for permission to transform itself. The old workflow isn't a neutral witness. It's an interested party, about to be radically disintermediated.

CEO as CTO: *Founder Mode* with Guardrails

The uncomfortable word *autocratic* needs one companion phrase, or else it sounds like I'm smuggling Silicon Valley founder-mode bravado into a sector that should know better. The companion phrase is *with guardrails*. Healthcare isn't Twitter. A hospital CEO can't wake up at 2 a.m., fire half the organization, rewrite the medical staff bylaws by tweet, and call it operational courage. Hospitals are civic institutions, trust-dependent institutions, moral communities, and regional labor anchors. The CEO can't behave like a caricature of Elon sitting menacingly over a manufacturing line. But she also can't behave like a ceremonial chairperson presiding over consensus theater while the production function of the enterprise is being rewritten outside the walls.

So yes: some version of *founder mode* is now required. Not the cultic version. Not the Adam Neumann version. Not the tone-deaf WSJ editorial version of the Cloudflare CEO boasting of his AI firings. Not the self-intoxicated version in which charisma metastasizes into megalomania. I mean founder mode in the more sober operating sense: the CEO is close to the work, close to the technology, close to the talent, close to the capital-allocation choices, and close enough to the workflows to know when the professional-managerial layer is translating urgency into delay. *The CEO has to become the CTO in the only sense that matters here*: she has to understand the technology well enough to decide, to prioritize, to say no, to kill work, to choose partners, to ask better questions, and to prevent AI from being metabolized into the old machinery of committees, pilots, procurement, and gently exhausted consensus.

This is where the Hayek point from the opening chapter returns in operational form. The enterprise used to suffer from *dispersed knowledge*. The frontier of the organization knew things the center didn't. Nurses, coders, revenue-cycle workers, clinic managers, pharmacists, analysts, schedulers, and physicians lived with the actual work, while the CEO received filtered reports from a managerial ladder descended spiritually from the railroad age (not kidding on this, actually). *AI changes the informational geometry*. Agents embedded across the enterprise can observe, summarize, route, compare, and synthesize the frontier of the organization in ways no biological leadership team could ever do manually. The center doesn't become omniscient, because nothing human does. But it becomes less blind.

And once the center can see more clearly, it's expected to decide more decisively. The old excuse that the organization is too large, too federated, too locally idiosyncratic, too politically delicate, or too difficult to understand starts to weaken. A CEO with a governed intelligence layer can see variation, duplication, bottlenecks, labor exposure, workflow friction, model adoption, and private productivity (not yet translating to enterprise productivity) in something closer to real time. That doesn't abolish *subsidiarity*. It does, however, discipline it. Local knowledge still matters. Local vetoes don't automatically deserve deference just because they are local.

That's why the CEO's own behavior matters so much. Use the models. Show the senior team the bad prompt and the better fifth prompt. Vibe-code badly. Ask the model to interrogate a board memo. Use it to pressure-test a payer-contracting argument, summarize a medical-staff dispute, draft a community-college retraining partnership, or build a first-pass task map of the revenue-cycle function. *Let the organization see that you're learning in public*. Vulnerability lowers the psychological barrier. The mandate raises the operating floor. The message isn't: I know everything. The message is: I'm learning, and therefore we're moving.

That's the version of founder mode healthcare can morally tolerate: humble about knowledge,

autocratic about priority, transparent about uncertainty, serious about guardrails, and ruthless about converting private productivity into enterprise productivity. The CEO doesn't need to become a machine-learning engineer. But the CEO does need to become the institutional diffuser-in-chief. If the CEO can't do that, the health system won't become AI-native. It will become AI-observed.

One God Model, One Agentic Harness, Diffused Everywhere

This next exhortation may sound a bit theological, and maybe that's unavoidable. The language around frontier models keeps drifting toward gods, oracles, priests, idolatry, and *dark enlightenment*—more on that in the deification chapter, where I let myself go even further into the metaphysics of this. But the point here isn't metaphysical. It's practical. *Pick one god model, one agentic harness, and diffuse it everywhere.* So not just the model, e.g. Claude, but the agentic harness that's the indispensable scaffolding *around* the model that brings it to life, e.g. Claude Cowork and Claude Code.

Don't create a balkanized island chain of disconnected AI consumer tools, departmental experiments, boutique vendor wrappers, and local AI hobbies that no one can govern, measure, secure, or learn from at enterprise scale. Choose one frontier partner. I admire Anthropic, for reasons of safety culture, enterprise seriousness, constitutional AI, and my own experience with the product, though the principle matters more than the logo. The strategic move isn't to "buy a chatbot." It's to build the health-system diffusion platform: HIPAA-safe, no-data-exfiltration, data-sovereign, audit-ready, integrated into the enterprise, and designed not as a clever toy for the already converted, but as a workforce-transformation layer.

The first requirement is universal, safe access. I don't mean everyone wandering off to use whatever consumer chatbot they like while promising, Scouts-honor style, not to paste PHI into it. That kind of laissez-faire indiscipline can become malpractice with a friendly interface. I mean broad enterprise access with contractual protection, logging, privacy preservation, data boundaries, role-based controls, and security architecture fit for healthcare. And then, just as importantly, an enormous amount of *education and training* about how to use it. Don't confuse indefinite comparative shopping with strategy. Intelligence at the frontier will proliferate, cheapen, and improve. The scarce thing won't be eventual access to a model and harness. *The scarce thing will be institutional absorption:* the capacity to become an AI-literate, AI-proficient, AI-experimenting, AI-diffusing organization, all within sanctioned bounds.

That's why every serious system needs what Dario somewhat mischievously calls a **ministry of education**. Not a training department with better branding. Not a compliance module. Not a cheerful intranet page with five approved prompts and a laminated "responsible AI" pledge. I mean an institutional apparatus for *upleveling the workforce* until frontier intelligence becomes as ordinary as email, Epic, Excel, or the phone. The ministry has two arms. The first is a **forward-deployed teacher** model: people who go into revenue-cycle pods, nursing units, finance teams, HR teams, legal teams, compliance teams, clinics, call centers, and physician groups and teach the model against real work. Not abstractions. Real workflows. Real friction. Real deliverables. The second is a **forward-deployed engineer** model for the harder tasks requiring integration, orchestration, agents, data access, workflow redesign, and governance. The more enlightened and fast-moving frontier labs are beginning to provide some version of this themselves, and health systems should take advantage of it.

Then create an **R37-style lab** inside the system (I've mentioned R37 in the context of our RCM company R1 repeatedly in this essay): a small, agile, young, technical, irreverent, close to the workflows, *protected by the CEO*, and measured on deployment, not demos. R1's R37 framing is useful because it isn't innovation theater. It's applied AI aimed at identifying, building, and scaling solutions

with measurable revenue-cycle impact. And the name itself is deliciously apt: *Move 37* in AlphaGo, the eerie, beautiful, non-human move that looked wrong until it was revealed to be brilliant. That's exactly the posture needed here. Not another innovation center. Not another pilot museum. A *protected deployment engine*. Launch, diffuse, measure. Outcomes, cost to serve, denials, days in AR, documentation time, call-center load, manager spans, headcount avoided. *Not vibes. Measures.*

The operating doctrine should be **tight, loose, tight**. *Tight* on data exfiltration restrictions, governance, sovereignty, PHI, approved models, logging, accountability, and what can touch regulated data. *Loose* on *experimentation inside those boundaries*. Let staff use it. Let them teach one another. Let the use cases surface from the people actually living inside the work. Then *tight* again on what gets standardized, integrated, embedded, and scaled. That's the only sane response to shadow AI: surface it, govern it, learn from it, and convert private productivity into enterprise productivity.

That last phrase is the hinge. *Private productivity* is already happening. People are already using these tools in the shadows to write faster, synthesize faster, code faster, appeal faster, summarize faster, and think faster. But unless the organization governs, teaches, measures, and scales those gains, the productivity accrues to the individual while the enterprise remains unchanged. *The CEO's job is to make the private breakthrough institutional.* That's diffusion. That's the whole game.

One more point on the god model and harness. Picking one doesn't mean theological permanence. It means operational coherence. The model layer may change. The partner may change. The architecture should be portable enough not to become hostage to one vendor forever. But during the first phase of institutional naturalization, coherence beats abstract optionality. If every department is using a different model, different wrapper, different prompt library, different security posture, and different definition of success, the enterprise will learn nothing as an enterprise. It will generate a thousand anecdotes and zero operating system.

The Upskilling Guide: Where Humans Still Matter

There's a noble version of this transition, and I don't want to lose it inside all the rampant talk of labor reduction, margin expansion, and *buyer-seller-distressed-asset* Darwinism. Our job is to *upskill everyone fast*. Everyone. Not because every worker can be rescued into some pristine, permanent, AI-proof role. That would be another bedtime story, and we have enough of those circulating already. Some of this upskilling will be transitional. Some of it may be ephemeral. Some of it may even be a little quixotic—almost romantic—as we race against the machines. But it's still noble. It's also smart. *Institutions that abandon their people to fear will get concealment, resistance, shadow AI, quiet quitting, and a thousand small acts of defensive sabotage.* Institutions that teach aggressively, honestly, and practically will get *discovery*.

The upskilling agenda has two registers. The first is personal effectiveness. Every executive, physician leader, nurse manager, analyst, coder, pharmacist, scheduler, and operator has to learn how to *use the models* to get smarter, faster, more *cybernetic*—more prepared. Ignorance in adjacent domains is no longer charming. The old “there be dragons” portions of the map are disappearing because everyone now carries, or soon will carry, an omniscient polymath in their pocket. If you don't understand a term, ask. If you need a first draft, ask. If you need a counterargument, ask. If you need to understand a vendor, a regulation, a payer policy, a disease state, a staffing model, a capital structure, or a workflow, ask. Skillsmaxxing isn't (just) a Silicon Valley affectation. It's the new literacy.

The second register is *enterprise role design*. Upskilling can't simply mean teaching everyone to write better prompts, though that matters too. It has to teach the organization where humans still

contribute most. I would use four filters.

First, look for low-substitution work: data-poor, messy, long-horizon, context-rich tasks where the model can't yet see enough of the world, or hold enough tacit reality, to act alone. Second, look for complementary work: hybrid human-agent workflow management where the person orchestrates, verifies, supervises, escalates, and translates between machine output and institutional consequence. Third, look for high-elasticity-of-demand work: areas where making the work cheaper creates the need for more of it, not less of it. Healthcare has many of these: more patient interactions, more care coordination, more outreach, more synchronization, more longitudinal check-ins, more medication-adherence support, more behavioral-health integration, more home-based nudges, more social-needs intercession. Fourth, look for inelastic-labor-supply work: rare expertise, embodied skill, trust relationships, taste, clinical judgment, leadership, persuasion, presence, and those complex human capabilities that don't scale simply because demand rises. For a fuller explanation of this heuristic, have a look at the excellent 80,000 Hours blog post: *How Not to Lose Your Job to AI*.¹⁴⁴ No

That's the guide. **Low substitution. Complementarity. High elasticity of demand. Inelastic labor supply.** Those are the domains where human contribution remains most defensible and valuable. Train people toward those. Move them away from work whose only defense is that the legacy workflow hasn't yet been decomposed. The goal isn't to pretend that everyone becomes an AI engineer. The goal is to help the whole workforce understand the shape of the new labor market: what the model can do, what the model can't yet do, what the model makes more valuable, and where the human has to move up the stack.

This is also why upskilling has to be fast, practical, and embedded in work. A classroom module won't do it. A certificate won't do it. Send the *forward-deployed teachers*—supplied by your frontier-model partner or a plucky startup (I have a few to recommend) or trained internally—into departments. Let them sit beside the scheduler, the nurse manager, the analyst, the pharmacist, the HR business partner, the supply-chain operator. Teach against the task. Show the before and after. Let people experience the tool not as an abstraction or a threat, but as *leverage*. Then capture the use case, measure it, and scale it. Upskilling isn't a benevolent HR program. It has to become the operating system of diffusion.

And that last point matters because the upskilling project isn't merely humanistic. It's organizationally self-serving and strategic. You can't diffuse AI through an organization whose people are terrified of it, hiding their use of it, or waiting for permission from a committee that meets monthly. *You need the workforce to become a discovery engine.* You need them finding friction, surfacing use cases, testing prompts, teaching one another, and telling the enterprise where the real work is broken. The CEO can mandate diffusion. The *ministry of education* can accelerate it. But the use cases live in the work. And the work lives with the people.

¹⁴⁴ 80,000 Hours: "How not to lose your job to AI" published in June 2025 (Link)

Ontologize the Enterprise

A word on data, because the health system that wants models to reason inside its enterprise has to *make the enterprise intelligible to machines*. That means partnering with a competent firm (again, I have a few to suggest) to ontologize the data, vector-enable the knowledge base, and give agents governed access to the workflows where work actually happens. Hospitals sit on mountains of data—notes, labs, images, claims, schedules, contracts, supply-chain orders, call logs, policies, HR tickets, physician rosters—but much of it's trapped in local semantics, half-structured fields, PDFs, scanned documents, departmental dialects, and tacit human memory. A model can't reason over what the enterprise can't represent.

Start with the ontology: a canonical map of the objects and relationships inside the enterprise. Patient, encounter, diagnosis, payer, plan, authorization, denial, appeal, physician, service line, order, result, medication, supply, contract, vendor, claim, location, role, workflow, task, error, metric. Map local codes to SNOMED, LOINC, ICD, CPT, RxNorm, FHIR resources, payer policies, and internal definitions. Make the nouns and verbs of the enterprise legible. Otherwise, you're asking the model to reason inside a fog bank.

Then vector-enable the unstructured layer: notes, policies, denials, contracts, transcripts, call-center records, operating procedures, clinical guidelines, and board materials. Put retrieval around it. Put access control around it. Put provenance around it. Put auditability around it. This is where the generic model becomes enterprise intelligence. Not because it has magically absorbed your institution, but because you've made your institution retrievable, searchable, governed, and *semantically coherent*.

This isn't a data-lake vanity project. We've had quite enough of those. This is the *precondition* for agentic work. Revenue-cycle agents need payer policy, contract language, claim history, chart documentation, denial reason, and appeal precedent. Clinical agents need notes, meds, labs, imaging, guidelines, care plans, and longitudinal context. Supply-chain agents need vendor contracts, utilization, inventory, price, and substitution rules. HR agents need policies, roles, benefits, employee inquiries, training records, and escalation pathways. Finance agents need general ledger detail, budget history, variance drivers, volume trends, labor assumptions, and capital plans. If these objects remain balkanized, AI becomes a clever interface pasted onto institutional incoherence. If they're ontologized and vector-enabled, *AI becomes a reasoning layer across the enterprise*.

The question, of course, is who to do this with. This isn't work I would hand to a generic systems integrator with a 300-page deck, a rate card, and a burial instinct. You need a partner that understands ontology, security, workflow, healthcare data, and enterprise deployment. Palantir is an obvious exemplar here, because ontology isn't decoration in their model; it's the substrate. Palantir isn't the only one—this is an increasingly well-populated competitive space (let's chat). This is also where your frontier-model partner matters because the model layer and the data layer have to become mutually intelligible. And it's where verticalized healthcare incumbents that have real workflow knowledge—especially in revenue cycle, claims, payer policy, and operational data—may matter more than pure technologists. The partner archetype isn't "vendor." It's co-builder.

The larger point is simple: *before the health system can become AI-native, it has to become machine-legible*. Before agents can coordinate work, the work has to be represented. Before models can reason across the enterprise, the enterprise has to stop being a thousand local dialects held together by heroic employees and bad interfaces. Ontology isn't glamorous. Vectorization isn't

exactly a CEO cocktail-party word. But this is the plumbing of institutional intelligence. And without it, the god model remains outside the walls, looking at the hospital through a keyhole.

Part III: The Operating Model—Labor, Workflow, Data, and Organization

The argument now turns from *mandate* to *operating model*. If Part II was about *who* has to lead this transition, Part III is about *what* actually changes inside the health system once the CEO leads. And the answer isn't "AI strategy" in the abstract. It's labor, workflow, data, organization design, capital allocation, and, ultimately, the operating logic of the enterprise itself.

I'm going to move more briskly here than I did in the Oppenheimer labor chapter, because the healthcare labor chapter already made the fuller denominator argument: healthcare as landfall, the roughly \$2.9 trillion labor denominator, the exposure stack, the regional labor architecture, the substitution rails, and the civic burden of labor reconstitution. I won't re-litigate that whole argument here. Consider this section the CEO operating translation. The healthcare labor chapter tells us why the labor denominator is exposed. This chapter tells the CEO what to do with that knowledge.

The conclusions are straightforward, even if the work isn't. **First**, decompose work to the task level, not the job-title level. **Second**, distinguish what is *augmentable* from what is *substitutable*. **Third**, build a GDPval-like internal map so models are tested against real deliverables, not demos. **Fourth**, move from holding company to operating company, because AI can't scale through local mythology and *Noah's Ark* duplication. **Fifth**, rethink the org chart itself, because agents change what labor is. **Sixth**, shift capital from *atoms* to *bits*. **Seventh**, accept that the workforce reset is coming, and speak about it honestly before the workforce concludes, correctly, that leadership is euphemizing its way toward the same destination.

In other words, Part III is where the abstract denominator argument becomes managerial anatomy. The CEO needs to know which organs of the enterprise are clinical, which are administrative exoskeleton, which are coordination tax, which are sacred, which are obsolete, and which are newly valuable because agents make them scalable.

That's the operating model. Now to the brutal arithmetic. The causal order is important: first make the work visible, then decompose it, then decide what deserves a human, then redesign the organization around that answer.

The \$890 Billion Question, Recapitulated for the CEO

Reader Note: This is the hospital-CEO recapitulation of the healthcare-labor denominator from Chapter 5. The same labor math returns here because it's now a management question: how much work does the enterprise actually need humans to do?

How much labor do hospitals actually need? Call it the **\$890 billion question**, or the **\$1 trillion question**, depending on which AHA labor estimate, hospital boundary, and year one is using. The precise number matters for budgeting, but the strategic reality is the same: hospitals and health systems are gigantic labor machines. Something on the order of \$1 trillion a year goes to labor. That number is so large it almost ceases to communicate. It becomes one of those abstract healthcare mega-statistics, like \$5.3 trillion of national health expenditure or 18% of GDP. True, but inert.

Disaggregate it, and the picture becomes more useful. Our working decomposition remains roughly **73% clinical labor and 27% administrative, managerial, infrastructural, or otherwise non-direct-care labor.**¹⁴⁵ In dollar terms, using the \$1 trillion frame, that's about **\$730 billion clinical and \$270 billion administrative and infrastructure.** Nursing is the gravitational center—perhaps \$300 billion to \$310 billion, roughly 31% of the total labor pool. Physicians, including employed physicians, hospital-based physicians, contracted physician labor, residents, advanced clinical supervision, and professional fees embedded in the system, are perhaps \$140 billion, or 14%. Allied health—therapists, lab personnel, radiology techs, pharmacy, respiratory therapy, imaging, diagnostic and therapeutic roles—adds roughly \$170 billion, or 17%. Clinical support—nursing assistants, medical assistants, patient-care techs, phlebotomists, transport-adjacent roles, orderlies, psychiatric aides, and the under-described connective tissue of hospital operations—adds another \$110 billion, or 11%.

Then comes the administrative and infrastructural layer. Revenue cycle—billing, coding, clinical documentation integrity, prior authorization, denial management, claims follow-up, eligibility, registration, payer correspondence, cash posting, the whole priesthood of American medical reimbursement—is probably on the order of \$55 billion of labor. IT is perhaps \$30 billion. HR \$15 billion. Finance \$18 billion. Management and supervisory labor \$45 billion to \$50 billion. Supply chain around \$20 billion. Facilities, environmental services, food service, plant operations, maintenance, security, laundry, and related supervision perhaps \$50 billion. Legal and compliance around \$12 billion. Marketing, communications, and other administrative functions fill in the rest. The spreadsheet isn't revealed scripture, but it's clarifying enough to make the basic point: the hospital is a gigantic clinical labor organism surrounded by a gigantic administrative exoskeleton.

That distinction prevents two opposing mistakes. The first is the lazy critique that all hospital expense is administrative bloat, which is false. The second is the sentimental defense that because hospitals are fundamentally clinical, their labor architecture is morally immune from redesign, which is also false. The hospital is clinical at its core, but administratively encrusted around that core. The job isn't to pretend nurses are replaceable widgets. The job is to collapse the coordination burden that makes the clinical organism so expensive, distracted, and slow.

The key distinction is *augmentable* versus *substitutable*. Most public conversation leaps immediately to replacement: will AI replace doctors, nurses, pharmacists, radiologists, coders, managers? That's the wrong first question. The better question is *how much of this labor machine is spent on structured cognition that can be externalized into software?* Once you ask it that way, the exposure map clarifies. Revenue cycle lights up like a Christmas tree because it's language-heavy, rules-heavy, document-based, queue-based, adversarial, repetitive, and at least partly verifiable—hence my repeated evangelism about R1 and its best-practice approach here. A denial appeal is a text artifact. A prior-authorization packet is a text-and-data artifact. A coding query is structured interpretation plus communication. Eligibility verification is rules and data. Claims follow-up is queue work. Clinical documentation integrity is pattern recognition plus language. None of this is easy, because American reimbursement is a surrealist opera of exceptions, local rules, payer-specific behavior, and procedural arcana. But it's the right kind of hard for models and agents: not physical, not usually real-time life-or-death, already supervised, already measured.

That's why I would estimate that perhaps 75% of revenue-cycle labor is *augmentable* and perhaps 30% to 35% is ultimately *substitutable* or compressible through agentic workflow. HR operations, finance reporting, compliance drafting, legal review, marketing, supply-chain exception handling, IT

¹⁴⁵ Bureau of Labor Statistics: "May 2023 National Industry-Specific Occupational Employment and Wage Estimates" published on April 3, 2024 (Link)

service desk, internal communications, policy summarization, procurement analytics, and management reporting follow the same logic at different intensities. Draft the memo. Summarize the policy. Reconcile the variance. Extract the contract clause. Prepare the audit response. Classify the ticket. Route the exception. Build the board deck. These aren't fantasy use cases. They are the daily metabolism of the modern enterprise, and hospitals are giant enterprises, whether the mission language likes that fact or not.

Clinical labor is different. Nurses are the largest bucket, but nursing isn't primarily a document problem. Nursing includes documentation, yes, and the documentation burden is absurd, corrosive, and ripe for AI support. But *nursing is embodied cognition*: noticing the patient, touching the patient, turning the patient, medicating the patient, escalating the patient, educating the family, coordinating the floor, sensing when something is wrong before the record fully knows it. Some of that can be supported by AI. Very little can be substituted by AI in the near term. Physician labor is exposed differently: less physical than nursing but more liability-bound and epistemically accountable. AI can help with chart review, inbox triage, note drafting, differential diagnosis scaffolding, literature synthesis, prior-auth letters, discharge summaries, patient messages, and second-opinion reasoning. But the physician remains, for now, the legal and moral bearer of the decision. The model doesn't own the order, the procedure, the consent conversation, or the adverse event. For the fuller, bolder, more dangerous argument about physicians and *medical superintelligence*, see the clinical-AI chapter. I won't smuggle all of that back in here.

Roll it up and the planning frame remains something like this: roughly **\$400 billion of hospital and health-system labor is meaningfully augmentable** by generative AI and agentic systems, and perhaps **\$100 billion is plausibly substitutable or structurally compressible** over a medium-term horizon. To be clear, this isn't rigorous math, defensible by a forensic, line-by-line spreadsheet accounting. This is a Larsen, (hopefully) educated extrapolation and intuition around matching emerging AI system capabilities with the surface area exposure of the health system vertical. This analysis will be an evolving, kaleidoscopically changing one. So in other words, I'm not saying that \$100 billion of jobs vanish tomorrow. I'm saying the labor architecture changes. Some work gets eliminated. Some gets automated. Some gets supervised rather than performed. Some gets redeployed. Some becomes dramatically more productive. The point isn't a crude headcount purge. The point is to redesign the labor-to-output ratio of the enterprise.

This is also why Jevons' paradox and the lump-of-labor fallacy have to be handled carefully. Yes, increased efficiency can create more demand. Yes, the amount of human work in an economy isn't fixed. Yes, healthcare has enormous unmet need: more coordination, more visits, more intercessions, more longitudinal support, more behavioral-health integration, more medication optimization, more social-needs synchronization, more closing of gaps in care, more patient outreach, more care at home. But that doesn't mean every administrative workflow deserves immortality. Filing taxes once a year doesn't become more socially valuable because software makes tax filing easier. Much of hospital administrative labor exists because of compliance, regulation, payer friction, documentation ritual, and inherited *proceduralism*. We don't need more of that simply because we can do it faster. We need less of it.

There's a second way to see the labor arithmetic, and it's even more uncomfortable because it makes the margin question explicit. Compare the average nonprofit hospital labor structure with a more disciplined operator like HCA. The numbers shift by year, definition, payer mix, accounting treatment, and what exactly gets captured in salary, wages, benefits, contract labor, and professional fees, so I want to be careful here: this is a stylized model, not revealed scripture. But the directional implication

is arresting. If, through agentification of administrative functions, workflow decomposition, management-layer compression, RCM automation, and broad de-bureaucratization, the national not-for-profit hospital sector moved from roughly 57% of net patient service revenue spent on salaries, wages, and benefits toward something closer to HCA's approximately 43.5%, the margin effect would be enormous. In my model, that kind of migration takes **average operating margins from roughly 1.3% to something like 14.5%**. Put differently, it implies about \$212 billion of reduced hospital-sector expenditure—roughly a 73-basis-point reduction in the share of U.S. GDP allocated to healthcare, moving healthcare from about 18.0% of GDP to 17.3%. Again, stylized. Directional. But directionally explosive. This isn't merely a better EBITDA bridge. It's a macroeconomic event.

The right CEO doesn't ask, "How do I replace people?" The right CEO asks, "What work inside this institution still deserves a human being?" That's a different question, and a better one. It honors human labor by refusing to waste it. It reserves humanity for judgment, trust, touch, accountability, persuasion, supervision, empathy, and embodied care. It also forces the organization to admit that a great deal of current human activity exists because we had no better coordination technology. Now we do.

Build Your Own GDPval

That brings us to method. It's not enough to say labor is exposed, or revenue cycle lights up, or structured cognition can be externalized into software. A health system needs a way to see its own work. Not as departments. Not as job titles. Not as headcount lines in a budget. As tasks, deliverables, inputs, outputs, error costs, review structures, and measurable outcomes.

This is where GDPval is useful. I explored this at length in the *Oppenheimer* labor chapter, so I'll just recapitulate the CEO implication here: GDPval is OpenAI's attempt to measure model performance not on academic puzzles, not on bar-exam trivia, not on a sterile benchmark that rewards cleverness in the abstract, but on economically valuable work products drawn from real professional life. The benchmark spans 1,320 specialized tasks across 44 occupations and nine major GDP sectors, built from representative work products created and vetted by experienced industry professionals. That's the important move. The model isn't being asked to hallucinate brilliance from a naked prompt. It's being asked to *do work*.

And the evaluative structure matters too. Outputs are judged against expert human work, not against a classroom answer key. That shifts the conversation from "is the model smart?" to the more operationally relevant question: is the work product good enough, reviewable enough, cheap enough, and fast enough to change the labor model? That's precisely the question every health system should be asking. Not whether AI is impressive. Not whether the demo is beautiful. Whether it can produce economically legible work the institution currently pays human beings to produce.

So build your own GDPval. Or call it *Hospitalval* if the branding morality police insist. Break work into tasks, not job titles. "Nurse," "coder," "case manager," "financial analyst," "scheduler," "denials specialist," "pharmacist," "medical assistant," and "manager" are too coarse. The unit of analysis has to be the deliverable: draft the denial appeal, reconcile the medication list, prepare the discharge summary, answer the employee benefits question, route the referral, summarize the chart, extract the payer policy, generate the board memo, close the coding query, schedule the follow-up visit, prepare the staffing communication, classify the service-desk ticket, write the prior-auth packet.

Then score each task with a simple but rigorous taxonomy. Is it language-heavy? Rules-based? Repeatable? Reference-dependent? Reviewable? Verifiable? What is the cost of error? Does it

require physical presence, licensure, empathy, legal accountability, final judgment, or embodied skill? Is it short-cycle or longitudinal? Is the output a text artifact, a structured field, a decision, an intervention, a handoff, a physical act, or a trust relationship? From that decomposition, the answer begins to reveal itself. Some work should be augmented because the human remains essential but the cognitive scaffolding can be externalized. Some work should be automated because it's repetitive, bounded, and auditable. Some work should be eliminated because AI exposes it as a compensatory ritual created by bad systems, bad interfaces, payer friction, regulatory accretion, or pure inertia. Some work becomes newly generative because cheaper intelligence enables something the old labor model never made economically plausible.

Outsourcing offers a clue. If you can outsource it, you can probably automate meaningful parts of it. If you can BPO it, decontextualize it, script it, queue it, measure it, and manage it by SLA, it's a candidate for agentic compression. That doesn't mean every outsourced function disappears. In fact, the opposite may be true in some domains: you may choose an external partner precisely because the partner has the data scale, engineering velocity, and workflow density to deploy AI across a high-dimensional function better than you can. But it does mean the structure of the work is already legible enough to be engineered. Revenue cycle, coding QA, documentation integrity, IT support, HR operations, finance reporting, procurement, credentialing, compliance, and call centers all have this property. They aren't simple. But they are decomposable.

Finally, *measure*. No more innovation theater. Every scaled use case needs a financial and operational hypothesis attached to it: reduce cost to collect by X, reduce denial cycle time by Y, lower documentation time by Z, improve schedule fill rate, reduce call abandonment, increase first-contact resolution, reduce management-reporting labor, widen span of control, close open requisitions, avoid incremental hires, reduce outsourced spend, improve operating margin. Some benefits will be hard to isolate. Fine. Make estimates. Run counterfactuals. Build the measurement muscle anyway. The systems that can quantify savings will move faster because they will know where to reinvest, where to cut, where to scale, and where the emperor has no clothes.

From Holding Company to Operating Company

Once the work is visible, the next conclusion follows: *the health system must become an operating company*. This is the structural prerequisite. Healthcare built a lot of *holding companies* masquerading as systems. Multi-hospital names, common branding, vague strategic coherence, and then the operational reality of *Noah's Ark*: two of everything, or five of everything, or seventeen of everything. Two supply chains. Two treasuries. Two revenue-cycle empires. Two HR stacks. Two cyber teams. *Sometimes even two CEOs*. Two clerical bureaucracies gazing at each other across the merger agreement. We called them systems because the letterhead said so. Too often they were **SINOs—systems in name only**.

AI makes that less tolerable because agents require legible work. You can't automate what you can't see. You can't scale what remains local mythology. You can't export a superior operating model if every geography runs on tacit memory, heroic workaround, and politically protected variation. To diffuse AI, you have to make workflows explicit, decontextualized, measurable, and redesignable. That's what an operating company does. It knows the work. It sees the work. It can compare the work across sites. It can decide which local variations are clinically necessary and which are inherited barnacles.

This is why HCA remains such a formidable exemplar. Whatever else one wants to say about HCA, it has the unmistakable instincts of an operating company par excellence (and the \$100b+ market

capitalization to show for it). Staffing ratios. Spans of control. Throughput. Standardization. Salary, wages, and benefits as a percentage of NPSR. An unsentimental relationship to operational engineering. That doesn't mean every nonprofit system should become HCA aesthetically or spiritually. It does mean the field should study what it means to run healthcare as visible engineered throughput rather than a federation of sacred local improvisations.

The *numerator-denominator* question belongs here. You can contract the numerator—fewer humans doing the same work—or expand the denominator—the same human labor serving more lives, more geographies, more visits, more panels, more interventions. The best systems will do both. They will take the administrative numerator down and expand the clinical denominator outward. Not by squeezing bedside care until it squeals, but by dismantling the human middleware that made scale so strangely unrewarding.

This also changes M&A, but we will come to that in Part V. For now, the CEO point is simpler: you can't acquire your way into an AI operating company if your own system remains a branded confederation of local exceptions. You have to build the operating model before you export it.

AI Is the New HR

Once AI becomes part of the labor model, org design itself starts to wobble. **AI is the new HR.** That sounds like a slogan, but I mean it quite literally. For more than a century, the org chart has been built around the assumption that intelligence is scarce, human, hierarchical, and slow to transmit. Authority sits above. *Information flows upward. Instructions flow downward.* Managers supervise humans because humans are the only cognitive units in the system. The org chart descends, spiritually at least, from the railroad diagram: layers, spans, reporting lines, and control points built for an age in which coordination required vertical stacking.

LLMs and agents attack that assumption. They don't merely make individuals faster. *They change what a unit of labor is.* A nurse manager plus a staffing agent, a policy summarizer, a communications drafter, and a variance analyst isn't the same managerial unit as a nurse manager alone. A revenue-cycle specialist supervising agents that prepare appeals, check policies, classify denials, and assemble documentation isn't the same job as a person manually working the queue. A physician with chart synthesis, inbox triage, literature review, patient-message drafting, and prior-auth support isn't the same physician as before. The job doesn't disappear in each case. But the boundary of the job changes. The worker becomes, in part, an orchestrator of non-human labor.

That means HR and IT converge. Workforce planning must include agent planning. Training must include model fluency. Performance management must ask not merely what the employee did, but how effectively the employee used intelligence as leverage. Job descriptions need to change. Promotion criteria need to change. Span-of-control assumptions need to change. The whole apparatus by which organizations define work, allocate labor, and supervise productivity has to be rebuilt around blended human-plus-agent teams. This is why delegating AI to IT is structurally wrong. IT can provision tools. It can secure platforms. It can integrate systems. But it can't, by itself, redesign work.

The better analogy for the future organization is less the old hierarchy than a swarm with governance. I'm taking the imagery here, in part, from Alex Karp's *Technological Republic*: starlings, honeybees, scouts at the periphery, shared signals, rapid local discovery, and central doctrine. The CEO sets the direction. The governance group sets the boundary. The diffusion lab builds and scales. The workforce discovers. The agents multiply capacity. The operating company captures and standardizes what works. That's not anarchy. It's a new kind of *managed decentralization*: centralized

strategy, decentralized discovery, centralized scaling.

Karp's honeybee analogy is persuasive because bees have a form of distributed intelligence that's both local and collective. The German word *Tanzsprache*—literally “dance language”—refers to the waggle dance by which worker bees communicate the direction, distance, and quality of a food source to the hive. It's a beautiful biological metaphor for what AI-native organizations have to become: scouts at the edge discovering opportunity, translating local insight into a shared signal, and then mobilizing the collective without waiting for every instruction to travel up and down a rigid hierarchy. Health systems need their own *Tanzsprache*: a way for frontline discoveries, workflow hacks, prompt innovations, agentic use cases, and operational breakthroughs to move quickly from the periphery to the center and then back out across the enterprise.

The irony is that the 1855 New York and Erie Railroad org chart, so often treated as the ancestor of modern managerial hierarchy, was itself created to solve a coordination problem in a technology system that had *suddenly become too large and too fast for old habits* (sound familiar?). It wasn't, at its origin, a monument to bureaucracy. It was an attempt to make a sprawling railroad legible. Over time, however, the org chart hardened into a kind of managerial theology: boxes, reporting lines, span, level, title, and the comforting pretense that authority and intelligence move neatly up and down a tree. That's the anachronism now. Clean up your 1855 org chart. Rethink it. The chart is no longer the work, and the boxes are no longer the only cognitive units.

From Atoms to Bits: Capex After AI

There's a capital-expenditure implication hiding inside all this. Health systems have spent generations thinking in *atoms*: beds, buildings, towers, clinics, parking structures, cath labs, ORs, MOBs, and the real-estate empire that proves, to boards and communities alike, that the system is permanent.¹⁴⁶ Some of that plant is essential. Some of it's sacred in the literal sense: the place where care happens, where babies are born, where the dying are comforted, where trauma is treated at 2 a.m. But a lot of it's too much, too expensive, too fixed, and too dependent on volumes that AI-enabled care may begin to redirect, virtualize, or prevent.

I'll be bold: **no more reflexive building**. Maintain the plant. Modernize what must be modernized. Build where access, safety, or strategic necessity require it. But absent exigent circumstances, slow your roll on new construction. The old reflex was more demand, more beds, more FTEs, more square footage. The new reflex should be more intelligence, more throughput, more inference, more orchestration, more home-based capacity, more virtual capacity, more edge monitoring, and more software-defined care. The capex frontier shifts from *atoms* to *bits*: GPUs, god model partnerships, enterprise licenses, vector databases, ontology layers, inference orchestration, edge deployment, cybersecurity, and the data infrastructure that lets the hospital become more than a physical plant with a billing department.

Token utilization becomes a capital-allocation signal. In the old world, a new building could be photographed, blessed by donors, named by philanthropy, and shown on the cover of the annual report. In the new world, the most important capital asset may be invisible: a governed model environment diffused through tens of thousands of workers, running billions—eventually trillions—of tokens against denials, care gaps, patient messages, staffing, scheduling, coding, and clinical synthesis. Boards will need to learn to love the invisible asset. The question won't be how many beds did we add, but how much human friction did we remove per dollar of inference? How many administrative hours did we retire? How many patients did we synchronize without another building? How much

¹⁴⁶ RevistaMed: “RevistaMed Updates the Size and Scope of the Medical Real Estate Sector 2024” published on May 21, 2024 ([Link](#))

demand did we prevent, redirect, or manage at home? Healthcare has lived too long in the world of atoms. The next operating leverage comes from bits.

This will be difficult culturally because buildings are legible to donors, politicians, trustees, and communities. Compute isn't. A tower gives everyone something to point at. A model-governance environment doesn't. A parking garage looks like capacity. An ontology layer looks like a procurement memo written by someone who should probably get outside more. But the invisible assets may matter more. The hospital of the future will still have walls, beds, ORs, ICUs, and human beings doing sacred embodied work. But the operating leverage will increasingly come from the intelligence layer wrapped around that physical plant.

The CEO's question should therefore change. Not "What do we need to build?" but "What care requires atoms, and what care can be moved into bits?" Every capital request should now compete against a compute request. Every building should have to answer why the same strategic objective can't be achieved through throughput, home care, virtual care, AI-enabled coordination, panel expansion, or administrative simplification. Not because buildings are bad. Because capital is scarce and the production function is changing.

The Workforce Reset

Let's not tell bedtime stories. This will become, in significant part, about labor reduction. Hard to avoid that conclusion when roughly 24 million Americans work in healthcare and social assistance, when hospitals alone employ millions, and when healthcare remains one of the few major industrial sectors where growth has been driven more by workforce expansion than by productivity.¹⁴⁷ So yes: 2026 has to be the year of efficiency for healthcare staffing.

That sentence will make many people in healthcare queasy, and I understand why. Hospitals aren't merely employers. They are civic anchors, training grounds, community institutions, and, in many geographies, the most important source of stable middle-class employment. The *sacred covenant with workers* isn't a fiction. It has moral content. But the premise underneath that covenant is changing. Healthcare can't remain the labor absorber of last resort for the entire American economy if intelligence itself is becoming cheaper, more abundant, and increasingly capable of doing the administrative work into which we've poured human beings for decades.

The first move should be *attrition*. Close nonessential open requisitions. Put a moratorium on new administrative hires. Shift the burden of proof onto managers asking for headcount: demonstrate why AI *won't* be able to do this set of tasks in the next twelve months. Not why AI can't do the entire job title. Why AI can't do the *tasks* that justify the incremental hire. *This distinction matters because managers defend roles, while AI attacks tasks*. Use vacancy closure, role redesign, retraining, redeployment, and span-of-control expansion before crude layoffs. Be compassionate and generous on the exit when exits are necessary. But don't confuse compassion with denial.

The Overton window on hospital layoffs will shift—and fast. It's already shifting in the broader economy, but healthcare lags, as usual, wrapped in mission language, chronic-shortage narratives, and the understandable emotional difficulty of reducing staff in organizations built around care. For now, ride the cover of attrition. Don't backfill reflexively. Treat every departure in a language-heavy, rules-heavy, screen-mediated function as an invitation to redesign the work, not as an automatic requisition.

Then flatten layers of management. Silicon Valley has already widened spans of control and

¹⁴⁷ US Bureau of Labor Statistics: "Health Care and Social Assistance: NAICS 62" published on May 15, 2026 ([Link](#))

compressed managerial ranks; some companies have effectively *tripled the number of direct reports* expected of managers. I'm not claiming the healthcare data are clean enough to prove that exact multiple inside hospitals. This is a directional theme, an operating postulate, not a finished empirical law. But observationally, healthcare hasn't absorbed the same span-of-control productivity. It should. Apply the question to every management layer: what reporting, synthesis, coordination, escalation, scheduling, performance monitoring, and communication work can now be done by agents, or by managers amplified by agents? The answer will make many middle layers look less inevitable than they used to.

Remote work is another automation map. The pandemic taught us which tasks could move behind a screen. Billing, coding, scheduling, claims follow-up, utilization management, record review, chart abstraction, IT support, finance, HR, tele-triage, remote monitoring, portions of radiology and pathology, medication review, counseling, care coordination, quality work, and much of the managerial apparatus all became more screen-mediated than healthcare wanted to admit. If a task can be done remotely, it's not *automatically* automatable, but it's much more likely to be decomposable into digital segments. Covid ran the experiment for us. We should read the results.

This will be labor-dislocating. Don't sugarcoat it. Staff are smart. They can see what is happening in other industries. They can feel the tools improving. They know that if a model can draft the memo, summarize the chart, write the appeal, build the presentation, reconcile the policy, classify the ticket, and prepare the staffing communication, some work will change and some work will vanish. The honest version is better: we will reduce the administrative burden first; we will use attrition before layoffs where possible; we will retrain and redeploy where sensible; we will be generous and humane when roles disappear; and we won't lie to you that every legacy job has a permanent claim on the future.

The reward is **deflation**. Take labor out and lower prices. This is the moral counterweight to the brutality of the labor argument. If health systems use AI merely to claw back margin and preserve high prices, they will deserve the political hostility that follows. The promise is different: *reduce administrative labor, lower the cost to serve, simplify the patient experience, and then lower prices.* Hospitals can become the good guys again. Not through another ad campaign about community benefit, but through actual deflation. Take out cost. Lower prices. Given the industrial scale of the sector, this is a service to the country with national economic and societal implications. That's the covenant that can replace the old one.

Part IV: Care Delivery After De-Bureaucratization

The argument now turns from the operating model to the care model. This is where the glamour lives, which is precisely why we've got to be disciplined about sequencing. It's tempting, always, to leap straight from generative AI to autonomous diagnostics, AI doctors, N-of-1 therapeutics, predictive deterioration, continuous monitoring, computational biology, synthetic biology, and *medical superintelligence*. All of that matters. Some of it will be civilization-shaping. But for the health-system CEO, the first battle is more prosaic, more immediate, and more operationally tractable: *de-bureaucratization*.

A note of discipline before we begin: I take up the full clinical-AI and liability architecture in the clinical-AI chapter, and I don't want to reproduce that argument here under another heading. This chapter has a different job. It asks what the hospital CEO should do so the enterprise becomes capable of absorbing clinical AI when the moment arrives, and in some domains that moment is already arriving. The sequence isn't "administration forever, clinical later." The sequence is more

precise: use administrative AI to build diffusion muscle, then move clinical AI through domains where validation, business model, workflow, and liability can be made legible.

The organizing principle is therefore *sequence without lethargy*. Move first where the institution can learn safely. Move clinical where the evidence, incentives, and liability architecture are ready. And don't confuse prudence with delay, because delay is still a decision—usually a decision made on behalf of the incumbent workflow.

The care-delivery claim is therefore simple: *rationalization precedes reconceptualization*, or in the most aggressive systems, rationalization and reconceptualization proceed in parallel under different risk envelopes. You first have to make the enterprise legible to itself, build safe model access, decompose work, remove sludge, prove that private productivity can accrue to the organization, and develop the muscle memory of iterative deployment. Then use that same muscle to redesign care.

Rationalization Before Reconceptualization

Let's prudently start where the work is less regulated, more tractable, more rules-based, more reviewable, and less likely to kill someone if the first draft is wrong. Start in revenue cycle, scheduling, coding assist, denials, prior authorization, HR service, IT service desk, finance reporting, compliance, procurement, policy operations, call centers, and documentation support. Start where the enterprise can actually learn.

This isn't because clinical AI is unimportant. It's because the enterprise has to build *diffusion muscle* before it tries to diffuse into the most safety-critical, liability-bound, professionally guarded domains. Ambient documentation is the bridge because it reduces physician and nurse burden while leaving the human clearly in the loop. It's *augmentation*, not *substitution*, and early evidence is suggestive enough to matter: clinicians experience relief when the machine takes some of the clerical weight vest off their backs.¹⁴⁸ But even here the strategic point isn't the note. The note is only the visible artifact. Once the model is allowed to listen, summarize, draft, classify, and route, the care process around the visit becomes newly redesignable. The administrative metabolism around the note is the prize.

After that, clinical diffusion should move through the *employed physician base* and through settings where the business model lines up. Provider-sponsored health plans, Medicare Advantage, Medicaid risk, case rates, advanced primary care, safety-net no-payment environments, and underinsured populations become strategic laboratories because the economic value of avoided utilization is visible. Fee-for-service remains hydraulically resistant to the most beautiful forms of clinical AI because prevention and synchronization often destroy billable activity. Where the system bears risk, or where the patient is unfunded and therefore the old reimbursement machinery offers no rescue, clinical AI becomes more obviously valuable.

148 Aaron A Tierney et al: "Ambient Artificial Intelligence Scribes to Alleviate the Burden of Clinical Documentation" published on February 21, 2024 ([Link](#))

From Fragmentation to Synchronization

Reader Note: Synchronization was introduced clinically in Chapter 3 and returns in the payer and behavioral-health chapters. Here it becomes a hospital operating model: AI matters because it can keep medical, behavioral, pharmaceutical, social, and longitudinal context alive across the enterprise.

The clinical opportunity is *synchronization*. Medical, behavioral, pharmaceutical, social, and longitudinal care remain grotesquely unsynchronized. Patients abandon drugs. Referrals leak. Gaps in care remain open. Behavioral health sits off to the side like an afterthought. Specialty care often proceeds without a whole-person context. The patient journey isn't a journey so much as a sequence of disconnected institutional encounters, and *the patient, shamefully, becomes the integration layer*.

AI can eventually help synchronize this: closing gaps, coordinating outreach, identifying abandonment risk, integrating medication management, matching patient-specific interventions, surfacing social determinants, and maintaining longitudinal context across sites and teams. The patient with diabetes, heart failure, depression, financial insecurity, medication complexity, and transportation barriers isn't a set of billing encounters. He's a person with intrinsic dignity who needs all of these domains to converge: medical, behavioral, pharmaceutical, and social. Our current institutions don't manage that domain integration well because the information is scattered, the incentives are misaligned, the humans are overburdened, and the workflows are episodic. An AI-enabled care model can keep context alive between visits, detect abandonment risk, prompt outreach, reconcile medications, coordinate behavioral health, surface social needs, and ensure that the next best action doesn't get lost in the gap between departments.

This is where employed physicians matter. Provider organizations already control, employ, or affiliate with a large share of the national physician base—more than half, depending on how one draws the perimeter.¹⁴⁹ That physician base is the diffusion channel for clinical AI, but only if the models are deployed in a way that makes the physician more powerful rather than more surveilled. The doctor should experience AI as relief from clerical suffocation and as force multiplication: larger panels supported safely, better pre-visit synthesis, smarter outreach, automated inbox triage, prior-auth support, clinical-gap closure, patient education, medication optimization, access to an increasingly omniscient medical superintelligence, and the ability to coordinate with a care team that actually sees what the doctor sees.

That's how clinical AI moves from toy to operating model.

Capitation, APCs, and Provider-Sponsored Plans

The business-model problem is that clinical AI doesn't fit naturally into fee-for-service. If a model prevents an admission, avoids a procedure, keeps a patient adherent, coordinates behavioral health, or manages chronic disease longitudinally, where exactly is the fee-for-service code? If it reduces demand, who gets paid? If it substitutes continuous intelligence for episodic intervention, how does the old reimbursement hydraulic recognize the value? It mostly doesn't. Fee-for-service pays for activity, and some of clinical AI's most beautiful use cases remove activity.

¹⁴⁹ American Medical Association: "Physician Practice Characteristics in 2024: Private Practices Account for Less Than Half of Physicians in Most Specialties" published on in May 2025 ([Link](#))

That's why clinical AI works best under capitation, delegated risk, case rates, Medicare Advantage, Medicaid risk, provider-sponsored health plans, advanced primary care, safety-net no-payment environments, and any other arrangement where preventing utilization is economically legible. Advanced primary care, especially when risk-bearing, is the natural diffusion engine for clinical AI. A team-based primary care model that includes physicians, nurses, pharmacists, behavioralists, social workers, medical assistants, and care navigators can manage far larger panels with more precision if it has AI support for outreach, documentation, triage, gap closure, medication reconciliation, patient education, and longitudinal monitoring, to say nothing of the more captivating medical-superintelligence possibilities contemplated elsewhere in this essay. The PCP becomes less a harried fifteen-minute tollbooth and more the clinical conductor of a synchronized system.

Provider-sponsored health plans matter for the same reason. **Health systems, don't sell your health plan prematurely!** I understand the temptation: risk-based capital, network adequacy, SG&A, payer economics, regulatory burden, and the accumulated disappointments of provider-sponsored plans over the last thirty years. But if clinical AI's value accrues most naturally to the *owner of the premium dollar combined with delivery capabilities*, then giving up that asset may be a categorical strategy error at precisely the wrong time. *The winning combination is premium dollar, ambulatory assets, employed or aligned physicians, advanced primary care, home-based care, pharmacy integration, behavioral health, and clinical AI.* That combination lets a system use AI to lower total cost of care, not merely increase coding intensity.

Liability, in One CEO Paragraph

The gating issue for clinical AI isn't capability alone. *It's blame allocation.* I spill a great deal more ink on this in the clinical-AI chapter, so I'll state only the hospital-CEO implication here: *don't wait passively for liability to be solved by someone else.* Hospitals should help architect the liability stack. Define validated clinical envelopes. Demand indemnification and liability-bearing partnerships where the model is doing real clinical work. Use malpractice captives strategically. Instrument model use, model override, model under-ride, uncertainty flags, equity effects, adverse events, latency, outcomes, and cost. Create clinician safe harbors for good-faith use of approved tools. Refuse vendor disclaimers that push all meaningful clinical risk back onto the physician while the vendor claims credit for the benchmark.

The standard cannot be infallibility. That's the frankly *hypocritical standard* we impose on new technology while tolerating massive human variation, diagnostic error, delay, fragmentation, and preventable harm from the current system. The standard should be human equivalence or human superiority inside a defined use case, with a transparent validation envelope and post-market monitoring. Today, malpractice anxiety pushes clinicians to treat AI as advisory, especially when it contradicts conventional care. Tomorrow, in some domains, failing to use an approved AI tool may be the negligent act. Hospitals should prepare for that standard-of-care inversion now.

The Sacred Trust

A word of caution, because this can't become merely an austerity manual with better metaphors. Healthcare is consecrated work. I don't mean that sentimentally. I mean that there's an intrinsic ethical dimension to the care of the sick that distinguishes it from ordinary commerce. Nurses, physicians, therapists, pharmacists, aides, and social workers occupy a sacred relationship with patients at moments of fear, pain, dependency, shame, hope, grief, and vulnerability. AI can augment this work, but if we use it to replace the human relationship with a simulacrum, if we reduce the encounter to an uncanny-valley chatbot with a cheerful tone and no soul, we will have committed a category mistake of grave proportions.

The point of de-bureaucratization is to *restore* humanity, not strip it away. Get the nurse out of the chart and back to the bedside. Get the physician out of the inbox and back into the conversation. Get the pharmacist out of fax warfare and into medication optimization. Get the care manager out of manual list-building and into actual intercession. If the administrative labor savings simply become margin while the patient gets a machine instead of a human at the moment of need, the politics and morality of this transition will curdle quickly.

So the covenant that replaces the old labor covenant has to be this: fewer humans wasted on administrative sludge, more human attention available for the parts of care that still require presence, trust, judgment, and love. *First rationalization, then reconceptualization.* First remove the sludge. Then redesign the mechanism.

Part V: Market Structure—Stratification, M&A, and Partnership

Let's take a moment and zoom out. Up to this point, the argument has mostly lived *inside* the walls of the health system: labor, workflow, data, organization design, clinical AI, liability, capitation, and the strange new task of teaching a giant institution how to absorb machine intelligence. But none of this stays internal for very long. Once some systems learn how to diffuse AI and others don't, the market structure changes. Labor intensity becomes destiny. Administrative complexity becomes either an exploitable vulnerability or a source of newly removable cost. Payer-provider abrasion changes character. RCM incumbents get compressed and one or two emerge as dominant—again, I'm betting on R1. Academic medical centers lose some monopoly on biomedical legitimacy. Health-system M&A accelerates. Balance sheets become instruments of strategic self-determination. Partnerships become existential. *And leadership, once again, becomes the scarce variable.*

So this part of the chapter is about the *external* consequences of the internal transformation. If the previous sections asked, "What should a CEO do inside the enterprise?" this one asks, "What happens to the field when a few CEOs actually do it?" Again, that's the tricky part of oligopolies—you really only need one or two first movers to catalyze the chain. The answer, I think, comes in five movements: **first**, the payer super-cycle breaks and providers briefly rise; **second**, systems stratify by diffusion capability; **third**, the old priesthoods and clerical layers lose protection; **fourth**, non-contiguous M&A becomes more plausible because the operating model can travel; and **fifth**, the balance sheet becomes a weapon rather than a passive reserve.

The Old Anti-Disintermediation Playbook Won't Work

Before the field stratifies, one more piece of statecraft needs to be named. The conventional incumbent *anti-disintermediation* playbook won't be enough. Healthcare incumbents have historically defended themselves through *regulatory enclosure*, *narrative warfare*, *interoperability chokepoints*, local market power, credentialing, procurement friction, and the tendency of any complex institution to convert delay into virtue. These tactics may still buy time. They won't create advantage. Time isn't strategy.

The reason is that AI works on the very substrate healthcare has used to protect itself: cognition, documentation, expertise, bureaucracy, and professional interpretation. *The thing behind the moat has learned to read the moat.* A payer policy manual, a denial rationale, a specialty guideline, a credentialing rule, a contract clause, a medical-staff bylaw, a compliance memo, a board packet—all of them become machine-readable, machine-queryable, machine-contestable, and eventually machine-executable. Incumbency still matters, but only if it's converted into speed. Trust, distribution, data, clinical legitimacy, litigation memory, and installed workflows are powerful advantages. Left inert, they become sentimental assets on a depreciating ledger.

That's the combinatorial math. Incumbents have things insurgents can't conjure quickly: patients, trust, distribution, workflows, regulatory muscle memory, and institutional legitimacy. Insurgents have speed, engineering culture, iconoclasm, capital intensity, and willingness to rebuild from first principles. The winning strategy isn't barricade or surrender. It's *partnerability*: pair insurgent velocity with incumbent legitimacy before the insurgents route around the incumbents and before the platforms swallow the application layer. Pirates and the navy, to channel Steve Jobs. The pirates explore. The navy holds territory. Healthcare needs both.

The Payer Super-Cycle Breaks, and Providers Briefly Rise

Permit a quick macro digression, because it helps contextualize what happens to the health-system hierarchy after AI diffusion. **The 2010s were the payer super-cycle:** the ascendancy of managed care, the subordination of providers—what I termed *disintermediation* and *commodification* in my 2023 paper, *Reform or Reformation*—and the quiet but decisive transfer of leverage from hospitals and health systems to payers. The payers had the premium dollar, the balance sheet, the government-program tailwinds, the Medicare Advantage and complex Medicaid machinery, the ambulatory strategy, the physician-control strategy, the *hospital-less integrated delivery and financing network*—more on this in the payer chapter—and the ability to make the old acute-care operating model look increasingly overbuilt. Hospitals faced oversupply, sluggish demand, site-of-care shifts, payer-mix degradation, physician scarcity, private-equity ambulatory insurgents, and a handful of enormous public MCOs with near-trillion-dollar aggregate market values at peak fighting fragmented, subscale provider systems. It wasn't a fair fight.

Now the center of gravity has shifted back toward providers, at least temporarily. Not permanently, and not nostalgically. This isn't the restoration of the old hospital-centered cosmos. It's a phase reversal driven by the fact that providers had the first obvious GenAI use cases: ambient documentation, acuity capture, coding optimization, revenue-cycle acceleration, denials management, prior authorization, documentation integrity, and administrative counter-weaponry against payer friction. These are immediate, remedial, cash-yielding uses of intelligence. *They are also inflationary*, at least in the first phase. Better documentation leads to better coding. Better coding leads to higher reimbursement. More aggressive appeal machinery leads to more recovered dollars. Providers that have

been under-armed technologically for a decade suddenly found a weapon.

The payer response has been revealing. Downcoding, auto-reductions, unilateral clawbacks, procedural obstruction, and the same old request to fax the chart. *Fax it.* In the year of our Lord 2026. The fax machine isn't just a comic anachronism. It's a tell. It reveals what the old payer moat, for some but not all, really was: delay, opacity, policy-manual asymmetry, procedural fatigue, and the ability to make the other side surrender. The payers built an empire not only on capital and underwriting, but on asymmetries of time and intelligibility. The less noble of them could surround the provider in fog: denials, carve-outs, obscure policies, site-of-care fights, coding asymmetries, appeal rituals, resubmissions, and waiting. That wasn't incidental. It was part of the business model.

Generative AI attacks that moat directly. Your policy manual can be 200 pages or 2 million pages; my model can parse it in *femtoseconds*—look it up, as we're going to have to get comfortable with these timescales: *one quadrillionth of a second*. Your denial rationale can be opaque; my agent can compare it to the record, the contract, the guideline, and the literature. Your appeal process can be exhausting; my system can generate the appeal, track the deadline, attach the support, and escalate the exception. Insurers have long counted on very few patients and too few providers appealing denials through to the end. That assumption is now vulnerable. Algorithm meets algorithm. *The old administrative intransigence becomes computationally tractable.*

This is why the current provider moment is real and strategically important, but it's not the end state. It's a waystation. In the short term, AI weaponizes the current fee-for-service battle lines and can be inflationary. Providers use AI to code more completely; payers use AI to deny more aggressively; providers use AI to appeal; payers use AI to downcode; the whole abrasion layer becomes a computational arms race. AI first *escalates* the war because both sides use the new weapon inside the old model. Then, over time, the more profound effect should be *deflationary* because the administrative logic on which both sides depended begins to decay.

There's a narrower market implication hiding inside the broader payer-provider fight: the revenue-cycle management market will be incinerated—or at least radically compressed and reconstituted. That may sound strange because RCM has been one of the great growth markets produced by American healthcare dysfunction: large, fragmented, labor-heavy, adversarial, and built on payer complexity, provider underinvestment, patient confusion, coding ambiguity, denials, appeals, and the chronic inability of institutions to understand one another. But that's precisely why it's exposed. A business that exists to intermediate complexity becomes vulnerable the moment intelligence makes that complexity more tractable.

That's why I'm such a partisan for R1—yes, again, with the obvious disclosure that TowerBrook is an investor in it. R1 is an *incumbent moving at the speed of an insurgent*, using its data, installed base, workflow intimacy, payer-provider knowledge, and scale not to defend the old labor model, but to transform it. That's the right incumbent move: don't wait to be disintermediated by creative destruction; use creative destruction as your own weapon.

Stratification: Prepared Minds, Purgatory, and Distressed Sovereigns

Health systems aren't monolithic, and the lazy habit of talking about "providers" as if they are one uniform category obscures the coming stratification. *AI diffusion will become the major determinant of competitive position.* Let that sentence reverberate in your mind for a second. If it's true—and I've high conviction it is—then we should phenotype the field into three categories, one of which of course betrays my deep and ineradicable Catholic schooling: **prepared minds, purgatory, and distressed sovereigns.** Though perhaps you should use something less medieval for your board deck.

The first category, perhaps 10% of the field, is the *prepared minds*. These are the emerging AI-operating systems. Their CEOs are learning, traveling to Silicon Valley, using the tools, vibe coding with Claude, meeting the AI-native founders, listening to the podcasts, studying the nomenclature, partnering with frontier labs, building diffusion capacity, tightening hiring discipline, closing open requisitions, mapping workflows, and treating AI not as an innovation program but as a treasury issue, labor issue, operating-model issue, and eventually solvency issue. *They understand that the advantage is asymmetric only while the work is still hard.* These are the likely acquirers.

A separate, somewhat insulated category may have structural assets that buy time, but not permanent immunity. They may have a profitable health plan, a standout reputation, a fortress balance sheet, philanthropic depth, geographic monopoly power, or some special clinical franchise. They may not yet be AI leaders, but they have enough insulation to observe, partner, and move deliberately. Time isn't victory, but it's time, and in a discontinuity even a small reprieve can matter.

The danger zone is *purgatory*: middling-to-weak operational performance, predominantly fee-for-service reimbursement, high labor intensity, limited balance-sheet flexibility, no real AI diffusion muscle, and a culture still waiting for the turnkey future to arrive. Some purgatory systems will ascend if they master the disciplines enumerated here. Others will descend into something colder and more financial than theological: *distressed sovereignty*. Again, sorry—sixteen years of Catholic indoctrination here. But this is the category in which labor intensity becomes destiny. If salary, wages, and benefits remain north of 55% of NPSR and the organization lacks the will or skill to compress administrative labor, widen spans of control, standardize work, and export a better operating model, then the system isn't merely inefficient. It's exposed.

That taxonomy matters because it predicts the next market structure. *Prepared minds acquire other systems. The structurally advantaged middle buys time. Purgatory becomes inventory. Distressed sovereigns become rescue cases, merger candidates, or stranded civic assets.* That sounds cold, and it is, but the coldness isn't gratuitous. It's the industrial logic of a sector whose cost structure is colliding with a technology that substitutes capital for labor and turns diffusion capability into operating advantage. Once intelligence becomes cheaper, the systems that remain most dependent on expensive human coordination become the most fragile. And once a few systems learn how to remove that coordination burden, the strategic sorting accelerates.

Who is Disadvantaged? AMCs, the Gerontocracy, and the Clerical Priesthood

The coming shift won't be gentle to every incumbent. Academic medical centers are complicated in this story because they're both essential *and* exposed. They remain among the great, proud institutions of American medicine: sites of tertiary and quaternary care, research, training, prestige, and complicated social missions. But their center of gravity isn't obviously moving in their favor. Please have another glance at my *Generative Epistemology* chapter for the more fulsome rendering: the old research establishment—NIH mechanisms, departmental chairs, editorial boards, peer-review ordination, grant cycles, academic promotion committees, the whole 'gerontocratic' architecture of legitimacy—isn't where the most dynamic research, synthetic and computational biological work is now being born. The frontier is moving toward hyperscalers, VC-backed startups, foundation-model labs, biological AI companies, and young computationally native teams with capital, compute, speed, and very little reverence for the old priesthood. This isn't the end of academic medicine. It may instead be the end of its monopoly on biomedical legitimacy.

The clerical priesthood is even more exposed. Coders, billers, schedulers, claims processors, prior-authorization teams, denial managers, HR service workers, IT service desks, policy analysts, report builders, call-center staff, and middle-management reporting layers are all doing work that models and agents can increasingly absorb, draft, classify, route, summarize, and supervise. That doesn't mean every person in those roles disappears. It means the identity of the role changes. The old clerical worker becomes an exception manager, model supervisor, workflow engineer, data steward, or patient-facing navigator—or the role disappears through attrition. Healthcare should be honest about this. The human middleware that flourished in the old coordination-tax regime won't be protected forever by the complexity that created it. Complexity was their employer. AI is coming for complexity.

The Non-Contiguous M&A Wave

Let me state the prediction directly, because it's too important to leave as an implication: *AI diffusion will accelerate health-system M&A dramatically*, and the most interesting deals won't be the familiar same-market combinations that antitrust lawyers, local newspapers, and state attorneys general know how to hate. The more important wave will be inter-market, non-contiguous horizontal consolidation. We should expect more multi-state, non-contiguous, \$50 billion-plus and eventually \$100 billion-plus health-system conglomerates. In fact, I predict we will see the creation of *multiple \$50 billion to \$75 billion health systems* across 2027. National powerhouses. Conglomerates in the old sense, perhaps, but finally with the possibility of becoming real operating companies rather than branded federations of local workarounds.

This is one of the largest consequences of the whole argument. The AI-proficient system that cracks de-bureaucratization doesn't merely become a better local operator. *It acquires an exportable operating model.* That's the thing hospital consolidation has almost always lacked. For decades, health-system mergers promised scale economies and then too often delivered *SINOs*: systems in name only. *Noah's Ark*. Two of everything. Two CEOs, two revenue-cycle empires, two supply-chain teams, two HR stacks, two treasury functions, two cyber groups, two coding bureaucracies, two call-center logics, two clerical priesthoods, two local kingdoms with one logo and very little true operating-company discipline. We got bigger, not better. We aggregated fragility. We didn't really engineer scale.

AI changes that possibility space. Not automatically, not magically, and not for everyone, but for the

systems that actually know how to diffuse it. Revenue cycle can be centralized. Coding support can be centralized. Documentation infrastructure can be centralized. HR service can be centralized. Cyber can be centralized. Treasury can be centralized. Supply chain and procurement can be centralized. Call centers, prior authorization, denials management, policy analysis, management reporting, contract review, compliance drafting, and service-desk operations can be run through common agentic infrastructure. Managerial span of control can widen because headquarters no longer has to rely entirely on layered human reporting structures to know what is happening across the enterprise. AI gives the system something like a nervous system.

That's why non-contiguous consolidation becomes more attractive. Local adjacency still matters for care delivery, physician networks, referral patterns, access points, brand, and politics. Bodies remain stubbornly local, and medicine isn't a SaaS product no matter how many venture capitalists try to make it sound like one. But adjacency matters much less in the functions AI is best at collapsing: revenue cycle, procurement, cyber, treasury, HR operations, finance reporting, coding support, documentation infrastructure, scheduling logic, management analytics, contact centers, and the whole administrative membrane between the patient, the payer, the physician, and the institution. Once those functions become software-mediated and agentically supervised, the operating model travels. And if the operating model travels, the old geographic constraints on scale begin to loosen.

The regulatory weather matters too. *A more permissive antitrust environment*, a less hostile FTC and DOJ posture—good riddance, Lina Khan and Jonathan Kanter—and the broad reopening of capital markets all create permission for dealmaking. But permission isn't the driver. AI proficiency is the driver. Regulatory relaxation supplies the oxygen; labor intensity supplies the desperation; AI supplies the operating mechanism. That combination is why this wave won't be just another cyclical return of the bankers. It will be a strategic sorting mechanism. The AI-proficient systems become buyers. The systems with structural insulation buy time. The labor-heavy, middling, fee-for-service-dependent systems become inventory.

And here is the crucial social condition: **this wave has to be deflationary** or it will deserve the political backlash it gets. Hospital consolidation has spent too long looking like price power dressed up as strategy. If the AI-enabled merger merely aggregates revenue, preserves duplicated labor, increases local leverage, and then asks commercial payers and employers to pay more for the privilege, then spare us the press release about transformation. But if the acquirer actually removes *Noah's Ark* duplication, standardizes the operating model, reduces administrative labor, lowers cost to serve, and then lowers prices, the moral and political proposition changes. For the first time, hospital mergers could plausibly produce real returns to scale and real deflation. That's the prize. Not bigger for its own sake. Bigger because the operating model is better, portable, cheaper, and less human-labor-dependent.

Weaponize the Balance Sheet

This is where I start sermonizing the nonprofit strategics, because the balance sheet is no longer a passive repository of reserves, bonds, real estate, and dutifully diversified investment-policy piety. If AI is an operating-model transition, a labor-substitution transition, a clinical-reinvention transition, and a consolidation transition, then capital allocation isn't adjacent to strategy. It's the circulatory system of strategy. Treasury and strategy can no longer behave like separate religions that meet occasionally at board dinners and politely admire each other's rituals.

The first rule is defensive: **build a fortress balance sheet**. That sounds boring, but boring matters when the field is about to stratify. The AI-proficient systems will need capital for model partnerships,

data architecture, agentic workflow, cyber, strategic equity, worker transition, and, eventually, M&A. The systems with weak liquidity, high labor intensity, deferred capex, and no AI diffusion muscle will discover that the future is cheaper to admire than to install. In this transition, financial flexibility becomes strategic oxygen.

The second rule is offensive: stop being merely a customer of the companies that will automate you. Nonprofit health systems are major customers, validators, distribution channels, data-context providers, workflow laboratories, and legitimacy engines for the companies that will define the next healthcare operating model. When you buy, integrate, validate, and scale an insurgent product, you are not just procuring software. You are helping create enterprise value. If you are helping create the value, participate in the value creation. This is a framework TowerBrook has been championing for years.

And to be clearer, this isn't venture tourism. It's not a request that every health system suddenly cosplay Sand Hill Road, hire a few ex-bankers, and spray mission-adjacent capital into fashionable AI startups because everyone else at the conference sounded excited. Quite the opposite. The investment discipline has to be tighter than that. Build only what is truly differentiating. Buy what is commoditized. Partner where the external market has engineering velocity, data scale, workflow specialization, or operating discipline you cannot replicate. And where your workflows, brand, distribution, clinical substrate, or data exhaust materially de-risk the company, take economics. That's not speculation. That's *anti-disintermediation*.

There are really three capital lanes here. **First**, infrastructure: the god-model partner, agentic harness, ontology layer, vectorized knowledge base, cyber posture, and governance architecture that the whole enterprise will run on. **Second**, workflow platforms: revenue cycle, prior auth, documentation, scheduling, call centers, supply chain, HR service, finance reporting, and the other domains where an outside partner may already possess the data density and engineering velocity to move faster than you can. **Third**, strategic options: minority equity, co-development rights, preferred deployment rights, advisory-board access, and other structures that let the institution benefit when its workflow intimacy helps create the category winner.

Epic, to some, is the *Empire Strikes Back* archetype. Some assert that a monopolist does what monopolists do: partner with insurgents while they are small, learn from them, tolerate them while they are non-threatening, and then bundle a native feature that seeks to eviscerate them. *I don't necessarily subscribe to this view*, but it's out there. And if there is any truth to it, then this is the *anti-disintermediation* strategy in its pure form. It's rational. It's powerful. It's also a reminder that the platform layer won't politely leave all value to the application layer. If you're a health system, you need to know which platform dependencies are useful, which are dangerous, and where you need optionality before the platform's gravitational field becomes destiny.

R1 is the more interesting counterexample, and yes, again, I'm biased here because TowerBrook invested in it. But the reason I keep coming back to it is that it is an *incumbent moving at the speed of an insurgent*. It has a large installed base, deep payer-provider domain expertise, proprietary structured and unstructured data, workflow intimacy, integrations across the revenue cycle, and a CEO posture that is missionary rather than defensive. That's the right incumbent move: don't wait to be disintermediated by creative destruction; use creative destruction as your own weapon.

Revenue cycle is the cleanest example of the partnership principle. If you're a hospital CEO, don't persuade yourself that your internal RCM department, already fighting payer policy drift, denials, coding complexity, labor shortages, brittle workflows, and a labor market that no longer loves you

back, is somehow going to outbuild the category pioneer. Partner with the pioneer. Hold the partner accountable. Take economics where you can. But don't confuse control with doing everything internally. That's how incumbents turn structural advantage into paralysis.

This is where my ironic Karl Marx invocation becomes too useful to resist: *own the means of production*. The old 60/40 investment-policy piety, outsourced to managers with no domain expertise and no AI insight, starts to look almost negligent when the institution has direct influence over which healthcare technology platforms win. *Diversification is sensible when you can't influence outcomes*. But health systems *can* influence outcomes in healthcare technology because, in many categories, they *are* the market. They can fund their mission not merely through philanthropy and bond markets, but by matching treasury to strategy and owning equity in the infrastructure that will define the next operating model.

The anti-disintermediation strategy isn't to wall yourself off from insurgents. *It is to bring them inside the tent before they burn it down from outside*. Put insurgents on advisory boards. Give them access to workflows. Co-develop products. Invest in them. Structure incentives so that your institution benefits when their tools scale. Reduce *mutual incomprehension* between the oligopoly of the 150 and the AI 10. Healthcare executives are too used to supplicants coming to them; in this moment *the mountain has to go to Mohammed*. Go west to San Francisco. Go to the frontier labs. Go to the twenty-seven-year-old founder who doesn't know the Stark Law but knows how to build the agent that will make half your administrative stack obsolete. Then teach her healthcare before she teaches herself the wrong version of it.

There's a moral version of balance sheet weaponization too. If a health system invests in the tools that compress administrative labor, then the winnings can't disappear into institutional self-congratulation. Some of the surplus has to return to the mission: lower prices, better access, worker transition, clinical reinvestment, behavioral health, primary care, and the human work that remains sacred. Capital strategy can't be separated from legitimacy strategy. If the hospital profits from the automation of work and the community sees no relief, the balance sheet will have been weaponized against trust.

Learn from the Gulf too, not because we should emulate monarchies per se, but because *state capacity* matters. G42, Mubadala, and similar sovereign-capital ecosystems will move quickly. They can reinvent abroad, reimport models, partner with frontier labs, and deploy with fewer veto points. China will move quickly too. The counterfactual isn't some tidy American healthcare future designed only by the AHA, Epic, CMS, and the top 150. The counterfactual is other regimes, with other politics and other values, diffusing faster. Complexity gives American health systems more time. It doesn't guarantee they'll use it.

The Leadership Turn

Ok, here is a melancholy reflection I've been having: with all this complexity and fog-of-war confusion, how many of the Healthcare 150 can actually make this turn? My guess is certainly not all. Maybe one-fifth? The exact number is intuition more than statistical rigor, but the direction feels right. The barrier isn't merely technical. *It's temperamental and institutional*. The current CEO class was selected for stewardship, diplomacy, consensus-building, regulatory navigation, philanthropy, physician politics, community legitimacy, and the careful management of giant organizations under constraint. Those are real, time-honored and long-cultivated skills. They aren't at all contemptible. But the skills that get you to the top of an *ancien régime* aren't always the skills that let you break it.

There's also a generational awkwardness here. The average health-system CEO isn't in her 60s, and the leadership demographics are still shaped by a slower institutional world. Douglas Adams—favorite author of Elon, Demis et al.—had the best line: anything that exists when you're born seems normal; anything invented between roughly fifteen and thirty-five feels revolutionary and exciting; anything invented after thirty-five feels unnatural and offensive to the proper order of the universe. That captures the cultural friction well. GenAI arrived late enough in the careers of many senior healthcare leaders that it feels like an intrusion rather than a native extension. Millennials and Gen Z are showing more receptivity to experimentation with AI, but so far this is proving to be a younger person's game. Poor Gen X—my own much-persecuted cohort—may once again be skipped over, waiting patiently for the baton only to discover that the baton is now an API.

But averages don't matter much in punctuated-equilibrium moments. All you need is one. One CEO. Or a handful. This isn't a democratized call to action. It's a siren call to the rare *insurgent-incumbent CEO*: the atypical leader from the *ancien régime* who sees the turn early, seizes self-determination while it's still available, and demonstrates that a health system can move faster than the category believed possible. Once one or two systems show the power, others will follow Darwin-style. The rare individual matters because technological change is often authored by the unrepresentative person acting at the right moment under the right conditions.

The old experts will resist because experts rarely self-disrupt. Vinod Khosla's line is directionally correct: experts know all the reasons something won't work. Healthcare is especially vulnerable to this because it's parochial, credentialist, self-referential, and suspicious of outsiders. The revolution is happening *outside* healthcare, which is precisely why healthcare has to humble itself. Again: go to Silicon Valley. Listen to the podcasts. Meet the founders. Learn the language. Spend time with the AI labs. Spend time with the twenty-six-year-old building a thing your senior team can't yet name. Not because Silicon Valley is morally superior. It's not. But because the dynamism is there.

This is the *prepared mind* again. The first task is psychological. Don't shield the organization from the exponentials. Expose them. Don't feed the staff bedtime stories about permanent stability. Tell them the truth with humanity. Don't hide behind the CIO. Don't let governance become anesthesia. Don't wait for the perfect product. Don't wait for consensus. Don't wait for another system to take all the arrows. This is a foreign motion for healthcare, but the foreignness is the point.

There's a further ambiguity in the CEO role that's worth making explicit. The *CEO has to be humble and autocratic at the same time*. Humble because no one really knows exactly how this will unfold. The technology is incipient, exponential, opaque, jagged, and changing faster than any normal enterprise planning cycle can metabolize. A serious CEO should say, out loud, "We're learning." That sentence matters. It gives the organization permission to learn too. It prevents the theater in which executives pretend to have mastered a domain that did not exist in its current form eighteen months earlier. But the CEO also has to be autocratic about priority. Humility about knowledge can't become timidity about action. We're learning, and therefore we're moving. We're uncertain, and therefore we're experimenting. We don't know the whole map, and therefore we're building the capacity to navigate it faster than our peers.

This is the difference between vulnerability and drift. Vulnerability says, "I'm learning this with you, and I expect you to learn it too." Drift says, "Because we don't know everything, we will wait." The former builds a prepared mind. The latter forfeits the future.

Part VI: The CEO Plan—From Doctrine to Deployment

The final movement returns to tactics. You've suffered through the EJM recitation of history, economics, labor arithmetic, clinical caution, liability architecture, market-structure implications, M&A prediction, and perhaps more bee metaphors than any one person has a right to make. Now we can stop circling the thing and state the plan plainly.

This is the part where all the metaphysics have to cash out. Coase becomes workflow. Statecraft becomes governance. Pasteur becomes readiness. The industrial revolutions become diffusion. The god model becomes an enterprise platform. The sacred covenant becomes a new labor architecture. The nation-state analogy becomes a CEO calendar.

So here is the punchline: move. Not recklessly. Not hastily. Not with the self-intoxicated bravado of a founder tweeting from a private jet. But move with speed and decision. The old healthcare habit is to admire the future until it becomes someone else's operating model. That won't work here. The future isn't asking to be admired. It's asking to be installed.

The First 180 Days: A 15-Point CEO Plan

Let me close where I began, with the tactical memo to the CEO. If I were running a large health system and believed even half of what I've written here, I wouldn't spend the next 180 days commissioning another white paper, scheduling another strategy offsite, or inviting six consultants to tell me the future is uncertain over stale biscotti. I would *move*.

The list is intentionally operational. It's not a strategy retreat masquerading as a plan. It's meant to force the first six months into decisions: access, governance, education, task mapping, labor discipline, measurement, clinical sequencing, partnership, and capital allocation. The point isn't to finish the transformation in 180 days. The point is to make institutional motion irreversible.

1. Make yourself proficient. Use the models every day. Not ceremonially. Not once before a board meeting so you can say you've "played with it." Use them. Learn the nomenclature. Learn prompting. Learn agents. Learn what the models are good at and where they still fail stupidly. Show your senior team your own learning curve. Let them see the first bad prompt and the better fifth prompt. AI literacy is now executive competence. Ignorance is no longer charming. "I'm not really a technology person" is no longer a defensible sentence for a CEO.

2. Own the transformation personally. Don't delegate this to the CIO, a consultant, a strategy office, or an innovation committee. The CIO is indispensable. The CEO is accountable. This isn't a tooling project. It's labor architecture, operating model, clinical strategy, capital allocation, market positioning, and eventually institutional survival. The CEO sets priority, resolves conflicts, allocates capital, breaks ties, kills legacy work, and reports measurable results to the board. If AI isn't in the CEO's scorecard, it will become theater.

3. Govern with speed. Create a small, cross-functional AI governance group with direct CEO and board visibility: clinical, finance, operations, legal, compliance, data, cybersecurity, and technology. Small. Strong. Fast. If the group becomes senatorial, disband it and start again. If approval takes longer than building the pilot, the structure has failed. Governance should make experimentation safe. It shouldn't become anesthesia.

4. Pick one frontier god-model partner and one agentic harness. Stop drifting through endless comparative diligence. Pick one serious frontier partner and build a governed enterprise

environment: HIPAA-safe, no-data-exfiltration, data-sovereign, audit-ready, role-controlled, and integrated into the workflows where work actually happens. I admire Anthropic, but the principle matters more than the logo. End the shadow-AI hypocrisy. Give your people a safe place to work. You can't govern what you force underground.

5. Build the *ministry of education*. This isn't a training module. This isn't an intranet page with five approved prompts. This is a workforce transformation apparatus. Deploy forward-deployed teachers into revenue cycle, nursing, finance, HR, legal, compliance, call centers, clinics, pharmacy, and physician groups. Teach against real work. Then deploy forward-deployed engineers against the harder problems: integration, orchestration, agents, data access, workflow redesign. The organization doesn't become AI-native through a webinar. It becomes AI-naturalized when people learn on the task.

6. Create an R37-style transformation lab. Not an innovation center. Not a pilot museum. A deployment engine. Staff it with young, technical, operational, irreverent people who are close to the workflows and protected by the CEO. Measure them on deployed workflows, not demos. The posture is launch, diffuse, measure. Outcomes, cost to serve, denials, days in AR, documentation time, call-center load, manager spans, headcount avoided. *Not vibes. Measures.*

7. Apply Tight-Loose-Tight. *Tight* on data, PHI, security, model access, governance, auditability, procurement standards, and scaling criteria. *Loose* on experimentation inside those boundaries. Let staff use the tools. Let them teach one another. Let the use cases surface from the people actually living inside the work. Then *tight* again on what becomes standardized, integrated, embedded, and scaled. This is the only sane answer to shadow AI: surface it, govern it, learn from it, and convert private productivity into enterprise productivity.

8. Build your own GDPval. Don't manage AI exposure at the level of job titles. Job titles lie. Tasks tell the truth. Build a GDPval-like map of the enterprise: tasks, deliverables, inputs, outputs, error costs, reviewability, physicality, licensure, empathy, legal accountability, and cycle time. "Coder" is too coarse. "Draft the denial appeal" is the unit of analysis. "Nurse" is too coarse. "Prepare the discharge summary" is the unit of analysis. "Manager" is too coarse. "Create the staffing variance narrative" is the unit of analysis. *AI attacks tasks before it attacks jobs.*

9. Classify the work. Every meaningful task should be labeled: augment, automate, eliminate, newly generate, or preserve as irreducibly human. *Augment* where the human remains essential but the cognitive scaffolding can be externalized. *Automate* where the task is bounded, repetitive, reviewable, and auditable. *Eliminate* where AI reveals the work as a compensatory ritual created by bad systems, bad interfaces, payer friction, regulatory accretion, or pure inertia. *Newly generate* where cheaper intelligence creates a care or operating capability that did not previously exist at viable cost. Preserve what still deserves a human being: trust, touch, judgment, persuasion, accountability, presence, ethical interpretation, embodied care.

10. Freeze nonessential administrative hiring. Close open requisitions unless managers can prove why AI can't do the work within twelve months. Not the whole job. The work. The task. The deliverable. Attrition is your friend. Use it. A resignation in a language-heavy, rules-heavy, screen-mediated function isn't automatically a vacancy. It's an invitation to redesign the workflow. The default should no longer be *backfill*. The default should be *interrogate*.

11. Start non-clinical, then move to clinical with discipline. Attack revenue cycle, denials, prior authorization, coding, scheduling, HR service, IT service desk, finance reporting, compliance, procurement, call centers, contract review, policy operations, and documentation support before pretending you're

ready for autonomous clinical genius. Start where the work is tractable, measurable, reviewable, and less likely to harm a patient if the first draft is wrong. Administrative simplification isn't the consolation prize. It's the training ground.

12. Measure everything. Track adoption, token use, cycle time, cost to collect, denial overturns, clean-claim rates, documentation time, call resolution, manager spans, requisitions closed, hires avoided, labor dollars saved, outsourced spend reduced, quality, safety, equity, satisfaction, and operating margin. Every scaled use case needs a financial and operational hypothesis. If you can't measure it perfectly, measure it directionally. If you can't isolate it completely, build the counterfactual. No more innovation theater. No more demo applause. No more "promising early results" with no P&L consequence.

13. Prepare honestly for labor contraction. This will reduce staffing needs in administrative and managerial layers. Don't lie about that. Use attrition before layoffs. Retrain where possible. Redeploy toward human work. Be generous on exits. But don't tell people every legacy role has a permanent claim on the future. The old conspiracy of silence won't work. Staff are smart. They can see what is happening. The humane version isn't denial. The humane version is candor, training, redeployment, and generosity.

14. Move from rationalization to reconceptualization. After simplification begins, use the same diffusion muscle to redesign care. Diffuse clinical AI through employed physicians. Build risk-bearing models. Protect or rebuild provider-sponsored plans where they make sense. Create longitudinal synchronization. Close gaps. Reduce medication abandonment. Integrate medical, behavioral, pharmaceutical, and social care. Move from episodic encounters to continuous context. First remove the sludge. Then redesign the organism.

15. Become an operating company and weaponize the balance sheet. No more SINO's. No more Noah's Ark. No more two of everything preserved by merger diplomacy and local politics. Standardize what should be standardized. Centralize what AI can centralize. Flatten where agents can amplify. Partner with frontier labs, Silicon Valley startups, and incumbents that move at insurgent speed. Take equity where your data, workflows, brand, and distribution create value. Build only what is differentiating. Buy what is commoditized. Partner where the market has engineering velocity you can't replicate. Build the exportable operating model before buying more geographies.

That's the first 180 days. Not the whole transformation. Not the final answer. Not a Decalogue descending from the cloud. A mobilization plan. A way to stop admiring AI and start installing it.

Coda: Buyer, Seller, or Distressed Asset

The real question for a health system isn't whether AI will matter. It already does. Not whether your employees are using it. They are. Not whether your innovation center has run pilots. Of course it has. The real question is whether you can become the kind of organization that diffuses this technology broadly and early, while asymmetry still exists.

Once the tools become intuitive, productized, and universal, everyone gets them. The advantage dissipates. At that point, the system that waited wasn't prudent. It was slow. It wasn't careful. It was commoditized. It did not avoid risk. It merely chose the risk of irrelevance.

This is why the difficulty is good. *The quagmire is the opportunity.* The fact that the products are immature, the workflows unclear, the governance awkward, the partnerships early, and the use cases messy is precisely what creates advantage for the prepared mind. When the category is easy, the category is over. When it's hard, the few can separate.

Separation is coming. The AI-proficient systems will lower cost to serve, reduce administrative burden, widen operating margins, expand geographically, and consolidate. The structurally advantaged middle will buy time. Labor-heavy purgatory systems will become sellers, rescue cases, or distressed assets. It sounds cold because the market structure is cold. But pretending otherwise won't make it kinder.

The humane version of this future isn't to preserve every legacy workflow, every administrative job, every local duplication, and every price level in the name of mission. That's not humanity. That's inertia with a chapel attached. The humane version is to remove wasteful labor, reduce the cost structure, lower prices, redeploy human beings toward care, and make the hospital less maddening for patients and clinicians alike. That's how hospitals become the good guys again.

The Chapter, Compressed

Here is the chapter compressed into the governing takeaways.

First, hospitals and health systems now resemble quasi-sovereign institutions. They govern populations, allocate scarce resources, manage legitimacy, negotiate with other power centers, employ millions, and shape regional civic life. Their AI strategy is therefore not a technology strategy alone. It's institutional statecraft.

Second, Coase gives us the frame. The firm exists because coordination is costly. American health-care is drowning in coordination costs. AI matters because it can reduce search, interpretation, translation, routing, verification, documentation, appeal, and handoff costs across a system whose administrative exoskeleton grew up around those frictions.

Third, diffusion beats invention. Health systems won't invent frontier AI. Their strategic task is to install it. The winners will be the institutions that diffuse AI horizontally through the enterprise before model access becomes universal and the early advantage disappears.

Fourth, the CEO has to own the transition. This can't be delegated to the CIO, innovation office, consultant, or AI council. The CIO matters. Governance matters. But the CEO must make AI a board-level operating mandate tied to capital allocation, labor strategy, organizational design, and measurable results.

Fifth, pick a god model, an agentic harness, and diffuse them everywhere. Fragmented experimentation without enterprise architecture produces local anecdotes, shadow AI, and no institutional learning. Broad safe access, a ministry of education, forward-deployed teachers, forward-deployed engineers, and an R37-style lab are the infrastructure of diffusion.

Sixth, ontologize the enterprise. Before agents can coordinate work, the institution has to become machine-legible. Data, workflows, policies, claims, contracts, notes, schedules, and operational objects must be represented, retrievable, governed, and semantically coherent.

Seventh, decompose labor to the task level. Jobs are too crude. Tasks reveal exposure. Build your own GDPval. Test models against real work products. Classify tasks as augmentable, automatable, eliminable, newly generative, or irreducibly human.

Eighth, hospitals must move from holding companies to operating companies. AI can't scale through local mythology, Noah's Ark duplication, and SINO's. The systems that standardize, centralize, and export their operating model will acquire. The systems that remain federated, labor-heavy, and locally improvised will be acquired.

Ninth, AI is the new HR. Agents change what labor is. HR and IT converge. Workforce planning must include non-human workers. Job descriptions, promotion criteria, performance management, and spans of control all have to be rebuilt around human-agent teams.

Tenth, capital shifts from atoms to bits. Buildings still matter, but the next strategic dollar must compete against compute, models, ontology, vector databases, cybersecurity, and the intelligence layer that allows care to move beyond the physical plant.

Eleventh, the workforce reset is coming. Use attrition before layoffs. Close nonessential open requisitions. Freeze administrative hiring. Flatten management layers. Be honest, generous, and humane. But don't lie that every legacy role has a permanent claim on the future.

Twelfth, de-bureaucratization comes before clinical reinvention, or at least before clinical reinvention can diffuse safely at scale. Start with administrative sludge, revenue cycle, prior auth, scheduling, denials, HR service, IT service desk, compliance, procurement, and documentation support. Use those domains to build diffusion muscle. Then redesign care.

Thirteenth, clinical AI requires business-model alignment. Fee-for-service resists the most powerful forms of clinical AI because prevention and synchronization often destroy billable activity. Capitation, APC, provider-sponsored plans, MA, Medicaid risk, and safety-net environments create the economic logic for AI-enabled care redesign.

Fourteenth, liability is the unlock. The question isn't only whether the model works, but who is responsible when it fails. Hospitals should help design validated envelopes, safe harbors, indemnification, captive strategies, and documentation standards rather than waiting for the liability regime to arrive from elsewhere.

Fifteenth, market structure will stratify. AI-proficient systems buy. Purgatory becomes inventory. Laggards become distressed. AI-proficient systems will gain operating leverage, export their model, and consolidate. The old prestige hierarchy will be less predictive than AI diffusion capacity.

And finally, the moral point: hospitals become the good guys again only if AI-driven labor compression produces lower cost, more care, and a more humane system. If AI becomes merely margin recapture, the industry will deserve the backlash. If it collapses coordination tax, lowers prices, restores clinical time, and redeploys human beings toward care, then this may be the first serious path toward healthcare deflation in a generation.

So yes, this is a memo to the CEO.

Use the tools. Study the exponentials. Pick the model. Diffuse the fire. Decompose the work. Collapse the coordination tax. Stop hiring into workflows that shouldn't exist. Take out labor where labor is being wasted. Lower your prices. Build the operating company. Earn the right to redesign care. Then consolidate, export, and lead.

BEFORE WE TURN THE PAGE

The hospital chapter is about the institutional actor that can install AI inside the care-delivery organism. But hospitals do not operate alone. The next chapter turns to the other great power center in U.S. healthcare: the payer, whose old moat of opacity, scale, delay, and paperwork is suddenly much less secure.

Chapter 7 • Future of U.S. Payers

The AI-Thin Payer and the Dissolution of the Insurance Maze

“Stability is destabilizing.”

— HYMAN MINSKY, *STABILIZING AN UNSTABLE ECONOMY*, 1986

A WORD ON NAVIGATING THIS CHAPTER

This chapter turns to the payer. It asks what remains of the insurance moat when opacity becomes machine-readable, appeals become agentic, administrative scale becomes rentable, and the paperwork maze starts dissolving under near-zero marginal cost contestation.

But this isn't only a demolition chapter. The darker argument is that AI dissolves the old paperwork moat; the more hopeful argument is that the best payers can finally become what managed care always promised to be: longitudinal synchronizers of medical, behavioral, pharmaceutical, and social-needs care. The chapter therefore moves from Minsky moment to secular decline, then through the sources of payer power—information asymmetry, scale, actuarial intelligence, and friction—and then into provider counterfire, API-ification, consumer weaponization, AI-native TPAs, PBM commoditization, vertical integration, and finally the optimistic payer playbook. The question isn't merely how payers thin. It's whether the best ones can become genuinely useful before the old maze is stripped for parts.

Future of Payers—AI Disintermediation

How payers devolve from a premium-collecting juggernaut into an AI-thin utility

Payers have met their “Minsky moment”—they just don't know it yet

We'll start with an economics detour, because it's not actually a detour, and, well, why not? But more than another Larsen distracting side quest, I think this might constructively illuminate what's happening in the U.S. payer market. A “Minsky moment” is the point in a long cycle when a period of calm, leverage, speculation, and institutional self-satisfaction suddenly curdles into recognition. Violent recognition. Hyman Minsky's great line—the one worth carrying around in your pocket—is that *stability is itself destabilizing*.¹⁵⁰ Long calm periods don't cure fragility; they incubate it. They encourage institutions to lever, overfit, and congratulate themselves for surviving a weather pattern they've mistaken for nature itself. Then something not even all that exotic happens, and the whole edifice remembers, all at once, that confidence isn't synonymous with resilience, let alone anti-fragility.

That frame, regrettably for U.S. managed care, fits almost too neatly. For most of the 2010s and into the early 2020s, the sector looked unassailable. Alongside the surging power of Big Tech, *Big Payers* grew into one of the most dominant industrial sectors in the American economy. Massive premium flows. Bipartisan tailwinds for Medicare Advantage and managed Medicaid. Relative insulation from FX chaos, semiconductor melodrama, commodity cyclicity, and the geopolitical weather that

¹⁵⁰ Investopedia: “What is a Minsky Moment” published on September 6, 2025 (Link)

damaged or remade other industries like manufacturing, industrials, and banking. By 2024, private health insurance spending alone had reached \$1.645 trillion inside a national health economy of \$5.3 trillion, or 18.0% of GDP. And on Fortune's 2025 ranking, UnitedHealth sat at No. 3 and CVS at No. 5 by revenue.¹⁵¹ That isn't some sidecar industry. As I analogized in chapter one, that's *nation-state-level* heft.

And what did many (not all) of the main payer protagonists do with that period of supremacy? They got larger. They got more vertically ambitious. They got more digitized. They got better at coding, better at utilization management, better at industrializing administrative asymmetry. Many of them also tried, sincerely and intelligently, to do the right thing: reduce needless administrative complexity, improve navigation, identify the patient's "next best step," and make the system work with some greater measure of coherence and humanity. I don't want to flatten that reality, because many of these payer CEOs are personal friends, and many are people of the highest character and integrity. So yes, I'm playing a little fast-and-loose here, abstracting away from individual motives and individual efforts in order to assess the dynamics on the playing field. But at the industrial level, I think the pattern remains hard to avoid: the sector mostly did *not* use its period of ascendancy to *abolish the coordination tax*. It didn't simplify the system in some radical, civilizational sense. *It optimized the maze*. It scaled the adversarial model. It became more formidable inside the *old logic* instead of escaping it. That distinction matters because the next technological regime—AI and agents—is unusually hostile to industries whose edge rests on opacity, delay, friction, procedural exhaustion, and the monetization of hassle. And now our emerging, disorienting, hostile-to-slow-moving-incumbents healthcare outlook is basically a catalog of those pressures colliding at once.

Let me say the thesis plainly before we get too deep into the thicket: AI doesn't merely make payers more efficient. *It attacks the historical reasons payers became so powerful*. It democratizes synthetic expertise. It reduces the cost of contestation. It makes policy and benefit logic machine-readable. It turns administrative persistence into a software function. It lowers the minimum efficient scale for risk intelligence. And, eventually, it commoditizes much of the payer's role into something thinner, faster, cheaper, more transparent, and less mystified. The end state isn't "payers vanish." Risk still has to be borne, capital still has to be reserved, networks still have to be assembled, regulation still has to be obeyed. The end state is that *payers thin*.

That is the laminar flow of this chapter. First, I want to explain why the payer super-cycle is over. Then I want to show why the decline is secular rather than cyclical. Then I want to walk through the historical sources of payer power—information asymmetry, scale, actuarial advantage, and friction—and explain why AI is corrosive to each. From there we'll look at the provider first strike, the payer retaliation, the API-ification of the maze, consumer mobilization, AI-native TPAs, PBM commoditization, and vertical integration. But then, importantly, I want to name the positive case: *the noble payer as care synchronizer*, the payer that embeds behavioral health, pharmacy, and social-needs intelligence into longitudinal management and thereby earns its place in the stack. This isn't a chapter arguing that payers vanish. It's a chapter arguing that the payer's old role thins, and that a better, more service-oriented role becomes newly possible.

And while payers as a whole may be vulnerable, there's a stratification and dispersion emerging among them. I'm joining the board of Oscar Health, and Oscar is trailblazing here; some of the Blues CEOs are absolutely moving with uncharacteristic speed to reconceptualize their businesses; and a couple of the publicly traded MCOs are assertively self-disrupting, and I'll let you guess which ones. But for the aggregate, things are about to get rough.

¹⁵¹ Fortune: "Fortune 500" June 2025 ([Link](#))

Pick your moment. The party is over.

Pedants can argue over the exact date the lights went out. There are several credible candidates, and I'm not too dogmatic about which one gets the ceremonial ribbon.

My own vote goes to September 2021, when Don Berwick and Richard Gilfillan detonated the "Medicare money machine" critique in *Health Affairs* and put into crisp prose what a lot of people in the industry—and especially leftist progressives in Washington—had been muttering *sotto voce*: that MA coding intensity and overpayment had become central to the economics and to the valuation story.¹⁵² Or you might point to July 8, 2024, when the *Wall Street Journal* reported that Medicare paid insurers roughly \$50 billion from 2018 to 2021 for diagnoses that patients received no treatment for, or that contradicted their physicians' assessments.¹⁵³ Or, most tragically and incomprehensibly, one can point to December 4, 2024, when a good man who happened to be leading the largest U.S. payer was senselessly gunned down in Manhattan, and the public reaction—ugly, subterranean, and revealing—exposed just how much anger toward the sector had accumulated. Or you can point to January 2026, when CMS proposed a microscopic 0.09% average increase in 2027 Medicare Advantage payments. The final April 2026 notice was more charitable, but this first salvo revealed just how much bipartisan congressional sentiment on the program had decayed¹⁵⁴—it wasn't just the socialists like Bernie, AOC and Elizabeth Warren—and on the news more than \$80 billion in combined insurer market value vaporized in a single day. And while the final April 2026 rate notice later softened the blow, the signal was clear.

Pick your moment. The larger point survives the dating dispute.

What makes this something *more* than a rough quarter, a nasty reimbursement cycle, or another ritualized payer-provider food fight is that several *older* tailwinds have simultaneously become politically and economically less reliable at precisely the moment that the new, civilization-shaping technology of AI is threatening to rip the old buttresses off the cathedral. Yes, Medicare Advantage is still enormous—just over 35 million people were enrolled as of February 1, 2026, and the program covered 54% of eligible beneficiaries in 2025—but the growth rate has slowed and the halo is gone.¹⁵⁵ MedPAC's March 2026 work still projected MA payments at 114% of what fee-for-service spending would have been,¹⁵⁶ meaning the government is still paying materially above FFS even after the V28 transition. One can (justly) find fault with the research methodology here, but that's precisely the sort of number that turns what used to be a bipartisan tailwind into bipartisan scrutiny. The party isn't merely less festive. The party is over.

Secular, not cyclical decline

Here's the point: this isn't some time-delimited, change-with-the-administration cyclicity. This isn't one more ugly but ultimately forgettable patch of weather. This is structural, entrenched, and, to my mind, newly permanent. And it's mostly because of AI. Allow me to explain.

The reason is straightforward: much of the payer sector's historical edge, at least for the less altruistic of the bunch, has rested on five things—*information asymmetry, returns to administrative scale,*

¹⁵² Health Affairs: "Medicare Advantage, Direct Contracting, And The Medicare 'Money Machine,' Part 1: The Risk-Score Game" published on September 29, 2021 ([Link](#))

¹⁵³ WSJ: "Insurers Pocketed \$50 Billion From Medicare for Diseases No Doctor Treated" published on July 8, 2024 ([Link](#))

¹⁵⁴ AJMC: "CMS Proposes Nearly Flat 2027 Medicare Advantage Payment Rates" published on January 27, 2026 ([Link](#))

¹⁵⁵ KFF: "Medicare Advantage Enrollment Grew by About 1 Million People, Mainly Due to Special Needs Plans" published on February 23, 2026 ([Link](#))

¹⁵⁶ MedPac: "The Medicare Advantage Program: Status Report" March 2026 ([Link](#))

the monetization of delay, the monetization of opacity, and, if I'm being a bit uncharitable in my decoration, the ability to make the other side miserable at lower cost than the other side could make you miserable back. Generative AI and agentic systems are bad for all five. Not mildly bad. Profit-sanctuary-annihilating bad.

McKinsey's 2026 outlook makes plain that the pain is already financial, not just conceptual.¹⁵⁷ It estimates that overall payer EBITDA fell from roughly \$61 billion in 2023 to about \$29 billion in 2024. Commercial EBITDA margins fell from 3.4% to 1.8%. Government margins fell from 2.9% to 0.5%. Set that against \$1.6 trillion in private health insurance spending and the asymmetry becomes clarifying: the flows are imperial, but the true profit sanctuary is much smaller than the old "recession-proof" mythology implies.

And that asymmetry is the whole point. The pathology of American healthcare isn't merely profit extraction in the cartoon-villain sense, whatever Elizabeth Warren may say in her most recent attempt to legislate the industry into moral repentance.¹⁵⁸ It's also the Herculean *coordination tax* required to keep this Rube Goldberg contraption moving. *JAMA* summarized the literature as placing annual U.S. health administrative spending somewhere between \$600 billion and \$1 trillion.¹⁵⁹ McKinsey estimated that roughly \$265 billion a year could be saved through administrative simplification alone.¹⁶⁰ That's the fat surface area. That's the target. That's what gets attacked when the paperwork moat stops being sacred.

The ACA's medical loss ratio rules make this more important still. Plans generally must spend at least 80% or 85% of premium dollars on medical care and quality improvement, with rebates if they fall short. So yes, labor substitution and SG&A compression can create a *first chapter* of earnings relief. Absolutely. There can be a *dead-cat bounce* (sorry to invoke that disagreeable phrase yet again in this essay, but it seems to fit). But this isn't a sector that can simply keep every dollar of administrative savings forever. In an MLR-capped world, a meaningful fraction of efficiency eventually gets competed away, rebated away, or returned in lower pricing. It becomes a race to the bottom on pricing—great for consumers, yes, patients; not so good for entrenched incumbents who don't fundamentally rethink their businesses. Quickly.

This is the first-order versus second-order problem payers have to metabolize. First-order AI is wonderful for payers. It lowers claim-processing costs, automates call centers, improves fraud detection, speeds prior-auth review, compresses SG&A, and makes payer operations less dependent on enormous administrative labor forces. Fine. That's the attractive part, and good for EPS. But second-order AI is existentially *dangerous and disintermediating* because it takes the same tools and arms everyone else: providers, employers, TPAs, brokers, regulators, consumers, plaintiff lawyers, consultants, and whatever insurgent founder has just raised a seed round to machine-read your policy manual and contest your denial logic at scale. *The same intelligence that makes the payer cheaper also makes the payer less necessary.* That's the central paradox.

That's why I think the end state isn't payers *disappear*. The end state is payers *thin*.

¹⁵⁷ McKinsey: "What to Expect in US Healthcare in 2026 and Beyond" published on January 12, 2026 ([Link](#))

¹⁵⁸ Elizabeth Warren: "Warren, Hawley Introduce Bipartisan Bill to Break Up Big Medicine" published on February 10, 2026 ([Link](#))

¹⁵⁹ JAMA: "Administrative Expenses in the US Health Care System" published on October 20, 2021 ([Link](#))

¹⁶⁰ McKinsey: "Administrative simplification: How to save a quarter-trillion dollars in US healthcare" published on October 20, 2021 ([Link](#))

The decade that was—the payer ascendancy of the 2010s

None of this should produce retrospective confusion or amnesia. The payer super-cycle of the 2010s was real. Very real. It was hiding in plain sight. Medicare Advantage mattered. Complex Medicaid mattered. Government payment tailwinds mattered. Scale mattered. Consolidation mattered. Vertical integration mattered. Administrative and data asymmetry mattered. If you were a large payer with a decent strategic brain and enough capital, you had the stronger hand. That was the story of the decade. For a long—insufferably long—description of this dynamic, please refer to my 2023 “U.S. Healthcare—Reform or Reformation?” essay.

The market structure tells the story rather brutally. The AMA’s 2025 competition report found that 97% of metropolitan commercial insurance markets were highly concentrated in 2024, and in 90% of metro areas at least one insurer had a 30% or greater market share.¹⁶¹ Providers, of course, logically responded in kind: KFF’s 2026 concentration analysis found that 97% of metropolitan inpatient hospital markets were highly concentrated, with 47% of metro areas effectively controlled entirely by one or two systems and 83% controlled more than 75% by one or two systems.¹⁶² *Oligopoly answered oligopoly*. Very American. Solve one concentration problem by minting another one.

And another externality here: the *corporatization of the American physician* was one of the clearest downstream expressions of this shift in leverage. By January 2024, 77.6% of physicians were employed by hospitals, health systems, or other corporate entities, and 58.5% of physician practices were owned by those same categories of organizations.¹⁶³ That wasn’t some spontaneous outbreak of managerial modernity. It was *retaliation and adaptation*. Independent, subscale, single-shingle medicine wasn’t winning a trench war against giant payers armed with coding pressure, claims friction, and administrative endurance. So the doctors aggregated. The practices consolidated. The U.S. physician herself got corporatized.

United was the apotheosis of this (very smart) strategy. In my 2022 essay “Covid, Secular and Structural Consequences,” I labeled this the **hospital-less IDFN—integrated delivery and financing network**—control the premium dollar, assemble the ambulatory rails, build the services and analytics stack, and (legally and appropriately) migrate profit out of the tightly regulated underwriting chassis into the unregulated adjacent services and tech businesses, all the while *disintermediating* and *commodifying* high-utilization and expensive acute-care providers wherever clinically appropriate.¹⁶⁴ Viewed from one side, it was a logical, highly serviceable idea in the sector. Optum’s physician footprint now exceeds 90,000 aligned clinicians and physicians depending on how one counts employed, contracted, and affiliated relationships. That’s much more than an ornamental side business. That’s an alternative *industrial logic*, and it pretty much single-handedly shaped the past decade for payer-provider competition.

The problem—and it’s the whole problem—is that the industry overall didn’t really use its decade of ascendancy to become genuine *managed-care* organizations in the noble sense. It became, in too many cases, a *managed-coding* industry. A *managed-friction* industry. A *managed-delay* industry. It used its moment to dominate the existing system harder, not to make the system legible. That worked for one cycle. It is a messy, expensive, and unsustainable inheritance for the next one.

¹⁶¹ AMA: “Competition in Health Insurance” 2025 Update ([Link](#))

¹⁶² KFF: “One or Two Health Systems Controlled the Entire Market for Inpatient Hospital Care in Nearly Half of Metropolitan Areas in 2024” published on March 27, 2026 ([Link](#))

¹⁶³ PAI: “Hospital and Corporate Acquisition of Physician Practices and Physician Employment 2019-2023” April 2024 ([Link](#))

¹⁶⁴ Eric Larsen: “Health Care Ecosystem – Secular and Structural Impacts of COVID-19” published in May 2022 ([Link](#))

Then providers got lucky—and moved faster

So how did we get to this moment? I'll start with an observation that's been much on my mind: one of the more ironic bits of the last year has been the indignation from certain payer leaders that providers weaponized AI in documentation and coding. I absolutely understand the irritation; I just don't find the sanctimony persuasive. Payers have been out-investing, out-digitizing, and out-industrializing providers for two decades. They built the rules engines. They built the denial machinery. They built the claims-review stack. So when providers suddenly got access to a cheaper, more intuitive, more general-purpose intelligence layer that could help with ambient documentation, chart review, coding support, evidence retrieval, and appeal drafting, of course they grabbed it. That wasn't vicious. That was equilibrium.

And, truthfully, *providers got lucky on the sequencing*. The first large healthcare use cases where this technology worked cleanly were domains of *functional verifiability*: revenue cycle, ambient listening, documentation, coding, summarization, evidence retrieval, prior-auth packet assembly. That's precisely where providers had the most accumulated tech debt and the most immediate ROI surface area. Menlo Ventures estimated that health systems supplied roughly \$1 billion of the \$1.4 billion incumbent healthcare investment flowing into healthcare AI, while payers contributed only about \$50 million;¹⁶⁵ Reuters similarly reported the payer-provider AI arms race as one in which hospitals had moved aggressively into coding and documentation while insurers were scrambling to respond.¹⁶⁶ I think the exact ratio may exaggerate the difference, but it strikes me as directionally correct. As an example, HCA expects about \$400 million in 2026 savings from AI, analytics, and revenue-management initiatives.¹⁶⁷ *Silicon speed replaced human drudgery on the provider side first because the provider side had the most remedial work waiting to be done.*

Naturally, this first fusillade won't go unanswered. In fact, the Blues have been quite explicit about what they think followed. Their March 2026 analysis documented that more aggressive AI-enabled coding may be associated with about \$663 million in inpatient spending and at least \$1.67 billion in outpatient spending, based on coding patterns that rose much faster than corresponding treatment intensity.¹⁶⁸ I'm not going to adjudicate whether that's overcoding, delayed justice for providers, or some unstable admixture of both. My narrower point is that *chapter one of healthcare AI has been inflationary* because providers used the tools first to optimize the reimbursement machine they inherited. Providers won round one, and yes, the center of gravity has shifted back toward providers for now. Not forever. But for now.

This matters because it reveals a larger point about AI diffusion in healthcare: the first mover is not necessarily the morally superior actor; *the first mover is often the actor sitting on the largest pile of functionally verifiable drudgery*. Providers had that pile. Payers have their own pile. Employers have another. Consumers, God help them, have been buried under one for years. AI doesn't arrive as a neutral productivity mist descending evenly over the system. *It arrives where the work is already digitized, painful, expensive, and measurable*. That's why the early story isn't AI replacing doctors. It's AI making the administrative war more efficient before it makes the war less necessary.

¹⁶⁵ Menlo Ventures: "2025: The State of AI in Healthcare" published on October 21, 2025 ([Link](#))

¹⁶⁶ Reuters: "US insurers and hospitals turn to new AI for age-old battle over charges vs payments" published on March 12, 2026 ([Link](#))

¹⁶⁷ Reuters: "HCA forecasts 2026 profit above estimates on \$400 million cost-savings plan" published on January 27, 2026 ([Link](#))

¹⁶⁸ BCBS: "New BCBSA Research Suggests AI in Hospital Billing is Leading to Higher Health Care Costs" published on March 5, 2026 ([Link](#))

Four reasons payers may be asymmetrically disadvantaged

Let's look a click deeper at why some incumbent payers—absent a genuine revision to strategy—are endangered in a post-AI world. The risk here isn't merely that payers, like everyone else, will have to automate. It's more specific, and more ominous, than that. Several of the historical sources of payer power are precisely the categories of advantage this technology is best at neutralizing: *information asymmetry*, *administrative scale*, *actuarial exclusivity*, and *friction as strategy*. In other words, AI doesn't simply make payers more efficient. It weakens the rationale for why they occupied such a privileged position in the healthcare stack to begin with.

And that's the deeper point. This isn't just a story about margin compression. It may prove to be a story about *institutional deselection*. AI doesn't merely threaten to make incumbent payers leaner. It threatens to make them less necessary.

No more information asymmetries

The first problem is intrinsic to LLMs and agentic systems: *they destroy informational asymmetry*. Not perfectly. Not finally. But materially, and in exactly the domains where payers have long enjoyed leverage. They lower the cost—to the marginal cost of *compute* and *electricity*—of parsing policy manuals, comparing benefit designs, interrogating carve-outs, synthesizing evidence, mapping denial rationales, reconstructing payment logic, and navigating the procedural maze that has historically favored the institution over the patient, the employer, and often the physician. What used to require battalions of analysts, consultants, revenue-cycle staff, and workflow priests can increasingly be done by much smaller teams equipped with rented intelligence and decent supervision.

That matters because a great deal of payer leverage has never resided in some sublime possession of actuarial sorcery alone. It has resided in *legibility asymmetry*. Payers (naturally) understand the rules better than the patient, better than the employer, often better than the doctor, and certainly better than the average practice administrator. Their advantage has often been less “we bear risk better than anyone else” than “we are more fluent at navigating the system's complexity than anyone else.” AI doesn't abolish expertise, but it radically democratizes access to *synthetic expertise*. It arms the other side. I think of this as an *arm-the-rebels* technology, and the rebels have now indeed been armed. Take the humans out of the equation, and things get inexpensive real fast.

Meanwhile the regulators are helping push the battlefield in the same direction. CMS' interoperability and prior-authorization rule requires more patient, provider, and payer data exchange, pushes prior authorization toward API form, imposes tighter clocks—generally 72 hours for expedited requests and 7 calendar days for standard requests—and requires specific denial reasons. Operational changes began in 2026, with API requirements generally due by January 1, 2027.¹⁶⁹ “Machine-readable” is no longer just a phrase in a conference deck. It is becoming law. And once rules, records, denials, and workflows become machine-readable, they become machine-queryable, machine-auditable, and machine-contestable. The move is from opacity to API. And from API to automated adversarialism.

That is terrible terrain for an industry whose edge has so often depended on knowing more, seeing more, interpreting faster, and—sometimes, for the less noble sectoral actors—exhausting the other side before the other side can even understand the rules of the game.

¹⁶⁹ CMS: “Interoperability and Prior Authorization Final Rule” published on January 16, 2024 ([Link](#))

Returns to scale hit an asymptote

The second problem is that *returns to administrative scale* begin to asymptote once scale is no longer tightly coupled to human labor. The old logic of payer bigness was straightforward: more covered lives, more premium, lower SG&A ratio, broader spread of fixed administrative cost, and therefore a larger and more defensible profit pool. There's still some truth in that. But less than there used to be.

Once claims review, benefits interpretation, member navigation, documentation handling, appeals preparation, coding review, fraud flagging, and prior-auth choreography are increasingly managed by AI agents, the labor-driven magic of bigness starts to fade. BLS counts 430,460 workers in direct health and medical insurance carriers, including 139,590 in office and administrative support and 97,270 in business and financial operations.¹⁷⁰ That's a great deal of clerical and quasi-clerical tissue sitting squarely in the blast radius of automation, augmentation, and elimination. Not as much administrative accretion as there is on the *provider* side (see my *Health System* chapter for an inventory, so you can see I'm being a nuisance to everyone equally, not just to payers) but for now let's simply agree it's a lot

BLS is already pointing in the same direction in adjacent claims work. Employment for claims adjusters, appraisers, examiners, and investigators is projected to decline about 5% from 2024 to 2034, and BLS explicitly discusses AI tools as a productivity factor in claims-related work. I don't need that number to be larger to make the structural point. A meaningful portion of the payer cost structure sits in *exactly* the categories of work that get thinned first in an agentic economy.

And once that happens, the industrial arithmetic changes. If the principal synergies of scale were administrative, and administration becomes radically cheaper, then you no longer need tens of millions of lives to achieve acceptable operating leverage. You can begin to approximate the administrative efficiency of a 10-million-life payer at a fraction of that size, because your *digital workforce* is just as industrious whether the covered population is 1 million, 10 million, or 100 million. Your agents don't get burned out, unionize, miss filing windows, or require another layer of managers to manage the managers. They simply execute.

So the denominator advantage starts to erode. Scale doesn't disappear. But it becomes more rentable, more modular, and less scarce. And if scale becomes rentable, it stops being much of a moat.

That is a serious problem for incumbents whose strategic self-conception has long rested on the proposition that bigger necessarily means better. In an AI world, bigger increasingly just means bigger. The returns to administrative scale begin to flatten, and with them one of the historical justifications for why giant payer institutions should exist in their present form.

¹⁷⁰ BLS: "Occupational Employment and Wage Statistics" May 2023 ([Link](#))

Actuarial risk economies of scale diminish

The third problem is actuarial. This is less a claim about imminent *omniscient medical superintelligence*—though let's not deceive ourselves, we're closer than polite professional society would like to admit (see my 'Clinical AI' chapter)—than a claim about statistical and actuarial superintelligence becoming modular, rented, and broadly accessible. Historically, giant pools had a real advantage because predictive machinery was crude, data was fragmented, interventions came late, and the system needed very large populations to smooth volatility, absorb outliers, and protect against MLR-busting shocks. That logic doesn't disappear, but it certainly weakens at the edge.

When smaller plans, employer coalitions, provider-sponsored plans, TPAs, and new intermediaries can *rent much better predictive models*, ingest cleaner and more standardized data, and automate large parts of underwriting support, forecasting, risk stratification, utilization prediction, and care-gap detection, they become much less intellectually disarmed than they used to be. The actuarial advantage of the giant incumbent starts to look less metaphysical and more like a service layer. An *acquirable* service layer

You can already see the modularity from the employer side. KFF's 2025 employer survey found that 67% of covered workers are in self-funded plans.¹⁷¹ Among firms with 10 to 199 workers, 37% of covered workers are in a level-funded plan, and 51% are in either a level-funded or self-insured arrangement.¹⁷² That's already a meaningful foothold for alternative administrative and risk-bearing architectures.

The key point isn't that every small employer suddenly becomes an insurance savant. It's that AI *lowers the minimum efficient scale* for actuarial competence. It lowers the intellectual fixed costs of risk management. It lets much smaller populations become suddenly proficient at actuarial sorcery. That doesn't eliminate the value of pooling risk. But it does mean that the ability to understand, forecast, stratify, and manage that risk is no longer reserved for giant incumbents with giant bureaucracies. Yet another historical advantage of bigness begins to erode.

And once actuarial intelligence becomes *rented* rather than *proprietary*, the payer starts to lose one of its core claims to indispensability. Risk still has to be borne. But risk intelligence no longer has to be monopolized.

Opacity and administrative friction as instruments go away

The fourth problem is the big one. A meaningful portion of payer power has historically come from the ability to *use administration as an instrument*: prior auth, downcoding, documentation loops, serial requests for more information, portal mazes, delay, ambiguity, and the strategic wearing down of the other side. Again, that's not the whole business model, nor is it the preferred *modus operandi* for *all* payers. But it has certainly been part of the economics. And AI is *catastrophic* for that model.

The AMA's 2024 prior authorization survey found practices completing 39 prior authorizations per physician per week, with physicians and staff spending about 13 hours per week (per doc) on the burden.¹⁷³ The precise number matters less than the qualitative truth every practice already knows: prior auth isn't benign administrative hygiene. It's industrial drag. It's cost shifted outward. It's friction externalized onto providers and patients. It's a system in which one side's operating model depends

¹⁷¹ KFF: "2025 Employer Health Benefits Survey" published on October 22, 2025 ([Link](#))

¹⁷² Peterson-KFF: "Recent trends in commercial health insurance market concentration" published on December 11, 2025 ([Link](#))

¹⁷³ AMA: "2024 AMA Prior Authorization Physician Survey" published in 2025 ([Link](#))

in part on the other side's exhaustion.

*Now imagine that same battlefield in an agentic world. Go ahead—make your process as bureaucratic, administratively intense, and emotionally depleting as you like. My agents don't get tired, discouraged, or confused. They don't lose forms, miss deadlines, or abandon appeals because the portal timed out. They don't interpret ambiguity as a cue to surrender. They are indefatigable, emotionless, and *existentially focused on one thing*: removing coordination friction. The old logic of "make this painful enough that only a minority will persist" starts to collapse once persistence itself is automated.*

And the numbers already reveal how much of the system runs on exactly that friction. HHS OIG found that 13% of denied Medicare Advantage prior-authorization requests in its sample met Medicare coverage rules, and that 18% of denied payment requests met Medicare coverage and MA billing rules.¹⁷⁴ KFF's 2026 MA prior-auth analysis found that MA insurers made 52.8 million prior-authorization determinations in 2024, denied 4.1 million of them, and that only 11.5% of denials were appealed—even though 80.7% of appealed denials were overturned.¹⁷⁵ That is a delay-and-friction economy hiding in plain sight. It works not because every denial is substantively correct, but because contesting the denial has historically been expensive, exhausting, and slow.

AI attacks precisely that substrate. Appeals can be auto-drafted. Clinical evidence can be assembled instantly. Policy language can be parsed by machines. Contradictions between denial rationales and coverage rules can be surfaced automatically. Cases can be escalated, tracked, re-submitted, and fought at silicon speed. Once that happens, administrative latency stops being an advantage. *Bureaucracy ceases to be a moat when the other side has near-zero marginal cost of contestation.*

And this, to me, is where the deepest danger lies for payers. If a meaningful portion of your economic power comes from being harder to navigate than the alternatives, then a technology that makes navigation trivial is not just a productivity tool. It's a solvent. LLMs remove bureaucracy; they therefore remove much of the strategic value of bureaucrats.

The counterargument: AI helps payers too

Now, let me steelman the payer position for a moment, because otherwise this risks becoming an Old Testament anti-payer polemic, and that's not quite the argument. AI will absolutely help payers. It will make fraud detection better. It will improve payment integrity. It will accelerate member service. It will reduce manual review. It will make actuarial models sharper. It will make prior auth faster, at least in the claims payers want to approve. It will allow a smaller payer workforce to administer larger books of business with fewer errors and less operational drag. There's a very real first-order productivity story here.

But the first-order story isn't the whole story. *The same tools that let payers automate administration also let everyone else automate resistance to payer administration.* The same models that let payers parse claims let providers parse payer policy. The same agents that let payers detect coding intensity let providers construct cleaner documentation packets. The same APIs that make payer operations more efficient make the payer's logic machine-readable to counterparties. The same actuarial intelligence that helps incumbents price and manage risk helps employers, TPAs, provider-sponsored plans, and new entrants rent actuarial competence without building an empire. That's the asymmetry. AI helps payers operate better, but it also shrinks the scarcity conditions that made the

¹⁷⁴ HHS Office of Inspector General: "Some Medicare Advantage Organization Denials of Prior Authorization Requests Raise Concerns About Beneficiary Access to Medically Necessary Care" published on April 27, 2022 ([Link](#))

¹⁷⁵ KFF: "Medicare Advantage Insurers Made Nearly 1 Million Prior Authorization Determinations in 2024" published on January 28, 2026 ([Link](#))

payer's privileged position defensible.

So the strategic question isn't whether payers will use AI. Of course they will. The question is whether they use AI to preserve the old friction economy, *or whether they use it to escape the old model before the market strips the model away from them*. One path creates a temporary earnings lift followed by commoditization and backlash. The other path requires self-cannibalization but offers a chance to become genuinely useful.

The deeper point

Put differently, AI doesn't merely squeeze payer margins by making administration cheaper. It *undermines the institutional logic* of the incumbent payer model. It strips away the scarcity conditions that made the incumbent payer economically privileged: privileged access to information, privileged command of administrative machinery, privileged ability to spread fixed operating costs, and privileged ability to profit from procedural friction. Once intelligence is abundant, workflows are agentified, and contestation is automated, many of those advantages stop looking like durable strategic assets. They start looking like software features.

That doesn't mean payers disappear overnight. Risk still has to be borne. Capital still has to be reserved. Networks still have to be assembled. Regulation still matters. But it does mean that incumbent payers are asymmetrically disadvantaged relative to many other healthcare actors, because AI attacks the very layers where payers historically captured value. It attacks not just their cost base, but their reason for being in the middle.

That's the crucial distinction. Providers still deliver care. Drug companies still make molecules. Device companies still make devices. But the payer's historical value proposition has often been to intermediate information, adjudicate complexity, and manage friction. AI is exceptionally good at collapsing all three. And once those become cheap, fast, and widely distributed, the payer no longer looks like an indispensable institution of modern healthcare. It starts to look like a legacy coordination layer whose functions are being unbundled.

That is why I think payers aren't merely disrupted in an AI world. *They're selected against*. They're not just facing a margin squeeze. They're facing a legitimacy squeeze. Unless they evolve from bureaucratic intermediaries into something genuinely additive—real care orchestration platforms, true risk-management partners, or longitudinal health operating systems—they will increasingly be deselected by employers, providers, regulators, and eventually patients themselves.

Because in the end, AI does not simply make the middle more efficient. Very often, it removes the need for the middle at all.

The retaliation is revealing

The payer retaliation is revealing precisely because it's so atavistic. It isn't some elegant, AI-native response to providers' first-mover advantage in documentation, coding, and appeals. It is the old religion. Old weapons. Old bureaucracy. Old maze. When providers used AI to get better at ambient listening, documentation, coding capture, packet assembly, and appeal drafting, the payers didn't respond with some grand reinvention of managed care. Some, *not all*, responded by trying to reconstitute the paperwork moat. That, to me, is the tell. It says they understand—whether they'll admit it or not—that paperwork, opacity, and latency were carrying much more of the economic load than any of the shiny investor decks ever acknowledged.

Take Cigna. This was not some vague tightening of “payment integrity.” It was a specific reimbursement policy—Evaluation and Management Coding Accuracy, R49—effective October 1, 2025, and aimed squarely at the revenue heart of outpatient medicine. By Cigna’s own description, the policy reviews professional claims billed with 99204–99205, 99214–99215, and 99244–99245. Cigna says the policy is narrow, peer-benchmarked, and hygienic.¹⁷⁶ The provider-side rendering is less antiseptic: unilateral downcoding of level 4 and 5 visits, opaque methodology, and delayed payment on legitimate professional work.

And then came the codicil—the revealing bit. In California, CMA objected that physicians challenging Cigna’s downcodes would have to appeal and submit supporting medical records by fax.¹⁷⁷ Fax! In 2026. In a chapter about AI. There it is. There is the whole industry pathology in miniature. On one side: ambient AI, retrieval, automation, documentation, and silicon-speed counterfire. On the other side: “please fax us the chart.” CMA challenged the policy, DMHC opened a review, and Cigna agreed to pause it for fully insured California HMO products pending review. That is not the behavior of a sector serenely confident in the legitimacy of its next-generation operating model.

Aetna’s version is less overtly branded but no less revealing. Aetna’s own E&M Claim and Code Review materials say the program evaluates levels 4 and 5 for new and established office and outpatient visits, consultations, and ophthalmologic services, using external coding guidelines and coder review.¹⁷⁸ Its 2025 and 2026 OfficeLink updates make plain that broader Claim and Code Review Program edits were rolling out across commercial, Medicare, and Student Health lines.¹⁷⁹ Different label, same industrial logic: scrutiny of E&M intensity, claims edits as posture, and more administrative friction concentrated exactly where the economics matter most.

The historical rhyme here is discernible. The 2003 Aetna and Cigna settlement materials had explicit “No Automatic Downcoding” provisions for evaluation and management claims.^{180 181} The AMA’s 2025 Private Practice Physicians Section noticed the *déjà vu*: it adopted policy explicitly opposing unilateral downcoding of E/M services by insurers, naming Cigna’s R49 and Aetna’s Claim and Code Review Program directly. This is a kind of recidivism.

That is why I say *the retaliation is revealing*. If the sector’s real edge were superior consumer experience, superior longitudinal care design, or genuine medical management, you would expect the response to provider AI adoption to look like a redesign of care and payment. Instead it looks like more edits, more downcoding, more documentation loops, more friction, more faxing, more analog drag. That’s not the posture of a sector confidently entering the next phase. That’s a rearguard action. That’s fighting the last war.

¹⁷⁶ Cigna: “New Reimbursement Policy for Professional Evaluation and Management Services Claims effective October 1, 2025” published on November 18, 2025 ([Link](#))

¹⁷⁷ California Medical Association: “CMA calls on physicians to speak out against Cigna’s unlawful downcoding policy” published on September 4, 2025 ([Link](#))

¹⁷⁸ Aetna: “Evaluation and Management (E&M) Program Claim and Code Review” published in 2023 ([Link](#))

¹⁷⁹ Aetna: “OfficeLink Updates” March 2026 ([Link](#))

¹⁸⁰ Aetna: “Settlement Agreement” dated May 21, 2003 ([Link](#))

¹⁸¹ Cigna: “Settlement Agreement” dated September 2, 2003 ([Link](#))

From opacity to API—the paperwork moat gets neutralized

Unfortunately for some payers, the external environment is running the other way. CMS' interoperability and prior-auth rule is one of those regulations that sounds dull until you realize it's structurally hostile to the old style of payer power. It forces more of the payer apparatus into machine-readable form, tightens the clocks, and requires specific denial reasons. That's not the old world. That's the beginning of the conversion of the maze into an AI substrate.

The same movement is happening in price transparency. Machine-readable plan rates and hospital rates are still messy, yes. Some of the files still look like they were assembled by a vindictive committee. Fine. But once those files become crawlable by agents, "surprise by design" becomes harder to sustain. The old trick was that the rules existed but were *buried*, distributed, contradictory, or unreadable at human scale. Agents don't care. They read the whole thing anyway. They compare. They cross-reference. They surface the buried exception. The old obscurantism works much worse when the counterparty isn't a tired billing clerk or an exhausted family member but a software system that never gets bored.

This is a huge point that the industry must internalize: *coordination, intermediation, and friction move toward a theoretical zero in a post-AI and agentic diffusion world*. Not instantly. Not cleanly. Not without countermeasures and regulatory confusion. But directionally, unmistakably.

This is why the phrase *paperwork moat* matters. Prior auth, claims routing, denial letters, medical-necessity disputes, appeals—these were once laborious, expensive, human-anchored workstreams. That complexity was not just a side effect. It was, in part, an instrument. Once it becomes machine-readable on both sides, the economic value of the paperwork moat starts to melt. First you get escalation: my bots versus your bots. Then you get equilibrium. Then you get deflation, because the human intensity comes out of the loop.

Consumer mobilization—the end of the friction and intermediation economy

The consumer side of this is underappreciated, and it shouldn't be. After all, they're the beneficiaries of this demilitarization. Health insurance has long lived off some combination of opacity, legalese, renewal inertia, broker intermediation, and the brute fact that normal human beings have jobs, children, fatigue, and no desire to spend their evenings comparing formularies and network directories like medieval monks. That's part of why the broker and benefits-intermediation economy exists, and equally why it is about to start evaporating. It's also why the market has been able to sustain so much slop. In 2025, average employer-sponsored family coverage reached \$26,993, with workers contributing \$6,850 toward that coverage on average; the average single deductible among workers with a general deductible was \$1,886.¹⁸² Those are large enough numbers to create a real appetite for machine-assisted comparison, renewal analysis, and optimization.

And the current market is still deeply inertial. During the 2025 Open Enrollment Period, CMS reported 20.2 million returning consumers came back to the Exchanges either actively or through auto re-enrollment.¹⁸³ There is shopping, yes. But there is also a lot of inertia, passive renewal, and status-quo behavior. That is exactly the kind of environment in which personalized agents become dangerous to incumbents. They don't get tired. They don't procrastinate. They don't say, "fine, I'll just keep this one another year."

¹⁸² KFF: "2025 Employer Health Benefits Survey" published on October 22, 2025 ([Link](#))

¹⁸³ CMS: "Marketplace 2025 Open Enrollment Period Report: National Snapshot" published on January 17, 2025 ([Link](#))

The claims-denial data on the commercial individual side points the same way. KFF's latest Marketplace analysis found that in 2024 [HealthCare.gov](https://www.healthcare.gov) insurers received about 496 million claims, including 451 million in-network claims, and ultimately denied about 85 million of those in-network claims—an average in-network denial rate of 19%.¹⁸⁴ Consumers appealed fewer than 1% of denied claims, and insurers upheld 66% of those appeals. That is not just a claims-processing story. It is a story about exhaustion, asymmetry, and underused rights. A market like that is begging to be counter-weaponized.

And now, predictably, it is. Counterforce Health markets AI-assisted appeals and says it is supported by NIH and the University of Pennsylvania. Fight Health Insurance explicitly advertises AI-generated appeals and—points for grim comedy here—invites users to “fax like the 80s.” That’s exactly the shape of the next consumer flank: not some abstract empowerment slogan, but actual tools that parse the denial, map the policy, draft the appeal, and route it back into the system. The consumer gets armed too.

And the grievance is already socially mature. KFF's February 2026 poll found that about one in three insured adults said an insurer had denied coverage for a doctor-prescribed service, treatment, or medication in the prior two years, with additional adults reporting delays or step-therapy requirements.¹⁸⁵ The tools are just catching up to the mood.

This, incidentally, is where broker economics become more vulnerable than they may look from inside the business. Brokers historically monetized complexity, renewal inertia, employer anxiety, and the fact that comparing health plans is among the least spiritually nourishing activities available to modern man. But a benefits agent that can continuously compare plan designs, formularies, networks, claims history, risk projections, and employer preferences changes the texture of the market. The broker doesn't vanish tomorrow. The dumb broker does. The broker who becomes fiduciary, strategist, and trust layer survives. The broker whose value proposition was “I can navigate the paperwork better than you” gets flattened by the machine.

AI-native TPAs will resurge

This is one reason I think AI-native TPAs matter much more than incumbents want to admit. The employer market is already more modular than the large carriers like to pretend. KFF's 2025 employer survey found that 67% of covered workers are in self-funded plans, and that 37% of covered workers in firms with 10 to 199 workers are in a level-funded plan. Among firms with 10 to 199 workers, 51% of covered workers are in either a level-funded or self-insured arrangement. That's not yet the overthrow of the incumbent carrier model. It is, however, a very substantial opening for alternate administrative architectures.

Now add agentic workflow, auto-drafted appeals, more API-based benefit verification, better claims intelligence, more machine-readable pricing, and cheaper rented expertise. The question is no longer whether a giant incumbent can do the work. Of course it can. The question is whether the giant incumbent remains necessary to do the work. In more and more domains, I think the answer becomes: not really. The regulatory shell still matters. The expensive intelligence layer matters less. That is the opening through which TPAs propagate.

And once cleaner cases move toward real-time or near-real-time adjudication, the denial latency premium trends toward zero. The old mega-payer could monetize slowness because the system was

¹⁸⁴ KFF: “Claims Denials and Appeals in ACA Marketplace Plans in 2024” published on March 24, 2026 ([Link](#))

¹⁸⁵ KFF: “KFF Health Tracking Poll: Prior Authorizations Rank as Public's Biggest Burden When Getting Health Care” published on February 2, 2026 ([Link](#))

human, analog, and exhausting. The AI-native TPA monetizes the opposite: lower friction, faster answers, cleaner routing, less theater. That's a very different value proposition. It's also much more commodity-like. Which, again, is the point.

PBMs commoditize too

I'll get into this in more detail in a subsequent chapter, but I'll just note that PBMs don't get a hall pass from any of this. The FTC's interim staff report found that the top three PBMs processed nearly 80% of the approximately 6.6 billion prescriptions dispensed by U.S. pharmacies in 2023, and the top six processed more than 90%.¹⁸⁶ That's not a stable natural aristocracy. That's a giant, concentrated spread pool sitting under a floodlight.

Notice also how quickly the PBM side is trying to soften the politics of paperwork. In March 2025, Optum Rx said it would eliminate up to 25% of prescription reauthorization requirements, beginning with about 80 drugs and affecting more than 10% of all such pharmacy reauthorizations.¹⁸⁷ That is not altruism descending from heaven. It is recognition that prior-auth friction, even on the pharmacy side, has become reputationally and politically toxic.

So yes, I think PBM economics continue to erode and commoditize. Not overnight. Not cleanly. And a small number that eschew the old opaque spread-and-friction business model—like TowerBrook's VytOne—will be the last ones standing. But the old notion that this layer gets to remain both opaque and richly extractive forever strikes me as fanciful. The same technology that makes plan logic more crawlable and appeal drafting cheaper also makes formulary navigation, utilization review, and benefit comparison more automatable. The middleman layers are not exempt from the general deselection of friction.

Vertical integration is the exemption

There is, however, an exemption—or at least a partial one. *The pure underwriter, if they adhere doggedly to the old model, is in trouble.* The player that combines the premium dollar with a real ambulatory delivery rail still has a *defensible anti-disintermediation strategy*. And this is where I don't want the reader to confuse my attack on the underwriting layer with some romantic provider populism. The big prize isn't administrative simplification alone. *The big prize is the clinical spend.* In 2024, the U.S. spent \$1.6 trillion on hospital care and more than \$1 trillion on physician and clinical services. That's where the real money is. If you own the premium dollar and enough of the ambulatory rail to change behavior upstream—primary care, multispecialty, home, pharmacy, behavioral, longitudinal management—you have a much more credible strategy than “we are very large and have a nice portal.”

That is why the *hospital-less IDFN* is still, to my mind, the right strategic answer. The apotheosis isn't underwriting for its own sake. The apotheosis is pairing financing with AI-amplified primary care, delegated risk, longitudinal management, and synchronized management of medical, behavioral, pharmacy, and social needs. That is what “managed care” was always supposed to mean before the phrase got degraded into managed coding and managed attrition. That is how MLR comes down for the right reasons—not because you got nastier with denials, but because you actually kept people out of avoidable downstream expense. It is one of the reasons I'm so enamored—still, even post-AI diffusion—with the APC, advanced primary care, model. More on that in my “Clinical AI” chapter.

¹⁸⁶ FTC: “Pharmacy Benefit Managers: The Powerful Middlemen Inflating Drug Costs and Squeezing Main Street Pharmacies” published in July 2024 ([Link](#))

¹⁸⁷ Optum: “Optum Rx further simplifies consumer access to prescription drugs” published on March 19, 2025 ([Link](#))

UnitedHealth Group, for all its current transitional pain and its in-process 're-founding' under legendary CEO Steve Hemsley, has the elements of the right strategy. And as a former United employee, I have a lot of affection and optimism for the place, especially under Hemsley. But the future isn't written here at all. Again, see my Clinical AI chapter for a broader mediation on the subject of vertical integration in an AI-diffused world. But yes, *vertical integration is an inoculation against some forms of disintermediation*. I'm not writing fan fiction here. I'm describing the mixed inheritance of the model.

Which is why I say United *should be* a relative winner, but not because underwriting remains glorious. Underwriting, I think, still commodifies relentlessly. So the reason a vertically integrated player has a chance is that the profit pool can migrate out of the regulated, depreciating underwriting wrapper and into an ambulatory care management, pharmacy, home, and services platform, all *reconceptualized* in a clinically-validated AI diffusion envelope. That was the aspirational point of the Optum play-book all along. The underwriting empire was always heading toward commoditization; the question was whether someone could build a clinical and services sanctuary around it before the commoditization became explicit.

Providers—don't sell your health plans

Reader Note: "Don't sell your health plan" appears in the Clinical AI and hospital chapters as well. I repeat it here because the provider-sponsored plan becomes more valuable once AI makes prevention, synchronization, and avoided utilization operationally tractable. And because I'm worried that providers aren't listening.

This is where I think the market is at risk of making exactly the wrong move at exactly the wrong time. Providers are tired. Provider-sponsored plans have often been underfunded, subscale, administratively ragged, and strategically neglected. So when a big health system says it is exploring strategic options for its health plan, including a possible sale, I understand the temptation. Running a small or mid-scale plan under the old model has often been a regulatorily annoying, operationally bruising exercise in masochism.

But *the technology barriers to operating a plan are falling faster than the regulatory barriers*, and that gap matters. I'm not saying there are *no* frictions. There are still licensure issues, reserve issues, compliance issues—all the usual amusements. I'm saying that the expensive intelligence layer is getting much cheaper. So a provider-sponsored plan paired with ambulatory assets, home assets, risk-bearing PCPs, and a real AI stack is a much more interesting object than the lonely, undercapitalized provider plan of the last era. Which is why my instinct here is: don't sell your health plan casually. Preserve the option value. Optimize it. Digitize it. Pair it with clinical AI. The market may be pricing those assets as exhausted relics just before they become strategically cheaper to run.

The key distinction is this: provider-sponsored plans failed in the old paradigm because they were trying to become mini-insurers without insurance-scale administrative machinery. In the AI paradigm, that machinery becomes cheaper, more modular, and more rentable. The old PSP (provider-sponsored plan) needed a claims bureaucracy. The new PSP needs a clean risk thesis, aligned delivery assets, a credible data layer, a narrow market focus, and agents that do the administrative work that used to require scale. That doesn't make provider-sponsored insurance easy. It makes it strategically interesting again.

What payers should actually do

So what should payers do if this thesis is even directionally right? Not issue another soothing investor-day slide about AI-enabled operational excellence. Not automate denials faster and call it transformation. Not hide inside the remaining complexity until the regulators, employers, providers, and consumer agents drag the model into daylight. *The serious answer is self-cannibalization.*

First, become radically API-native. Not reluctantly compliant. Not machine-readable only where the rule forces it. *Actually API-native.* Make benefits, networks, prior-auth logic, denial reasons, formulary rules, and appeal pathways legible, queryable, auditable, and fast. The old instinct is to preserve opacity because opacity creates leverage. That instinct now becomes lethal. The payer that voluntarily becomes legible can become trusted. The payer that has legibility *imposed* on it will experience it as disintermediation.

Second, auto-approve aggressively where the clinical and economic logic supports it. The point of AI-enabled medical management shouldn't be to deny more artfully. It should be to identify the cases where administrative review adds no value and remove the friction. The payer that reduces nonsense friction faster than the market forces it to will *win trust* from employers, providers, and regulators. The payer that preserves friction until friction is outlawed or automated away will be treated, quite reasonably, as an obstacle.

Third, shift from *claims adjudication* to *care orchestration*. This is the hard one because it requires actual *managed care* rather than the linguistic artifact called managed care. Risk stratification, longitudinal primary care, behavioral integration, pharmacy optimization, home monitoring, social-needs navigation, and high-risk patient engagement—these aren't optional adjacencies. They're the only defensible reason to remain in the middle. If the payer can't improve care trajectory, it becomes an expensive payment rail with a call center.

Fourth, collapse SG&A before someone else collapses it for you. If the labor layer is going to thin, better to do it deliberately, humanely, and in exchange for lower price, better experience, and faster administration. The payer that converts SG&A savings into trust will survive. The payer that converts SG&A savings into margin and then asks for sympathy when the market retaliates deserves what follows.

Fifth, stop treating providers as adversaries in every transaction. This is the hardest cultural shift of all. The payer-provider war has become a recursive melodrama of administrative employment, consultant revenue, software spend, and mutual resentment. Both sides are at fault. AI can either make the war more efficient or make parts of the war unnecessary. The latter is the strategic choice. The former is just trench warfare with GPUs.

The optimistic payer: from friction merchant to care synchronizer

Now the hopeful part, because this chapter otherwise risks sounding like a payer funeral written by someone with too much Schadenfreude. The best payers won't merely survive this transition; they may be uniquely positioned to redeem the very phrase *managed care*. Many are led by CEO's of great integrity and altruism. They sit on longitudinal claims, pharmacy, eligibility, network, benefit, utilization, and increasingly clinical data. They know where the care gaps are, where medications are abandoned, where behavioral disease is quietly driving medical spend, and where transportation, food insecurity, housing instability, loneliness, and caregiver collapse show up in utilization long before they show up in a tidy diagnostic label. A noble payer—yes, I mean that phrase quite seriously—can use AI to turn those fragments into a *care-synchronization layer*.¹⁸⁸

The work isn't to deny faster. It's to know the member longitudinally, anticipate the next clinical risk, and route the next best step before the ED visit, admission, relapse, amputation, stroke, or preventable decompensation. The payer that can synchronize medical, behavioral, pharmacy, and social-needs signals has a real reason to be in the middle. It can identify a diabetic patient with rising A1c, worsening depression, missed refills, food insecurity, and transportation problems; trigger a behavioral-health companion, pharmacist outreach, nutrition support, home visit, primary-care slot, and benefits navigation; and then monitor whether anything actually happened. This is managed care in the original, nobler sense: not managed abrasion, but managed trajectory.

Behavioral health and pharmacy are the first synchronizing wedges

Behavioral is the obvious wedge, for reasons I belabor in the behavioral chapter. Behavioral illness isn't a sidecar; it's the *multiplier*. Milliman found medical-surgical costs for people with behavioral-health conditions were 2.8x to 6.2x higher than for people without one, depending on the condition, while only a tiny share of total spending went to behavioral treatment itself.¹⁸⁹ That's an actuarial flashing signpost. And pharmacy is the adjacent wedge: poor medication adherence kills roughly 125,000 Americans a year and can cost the system as much as \$300 billion annually.¹⁹⁰ If a payer can't use AI to close that loop—detect risk, intervene cheaply, personalize the nudge, escalate to the human when necessary, and measure the downstream medical effect—then what exactly is the payer managing?

The best payer therefore becomes less like an imperial adjudicator and more like a longitudinal operating system for health: not the owner of the patient, not the replacement for the physician, not the sovereign of the clinical relationship, but the synchronizer of fragments the current system has made nobody's job to hold together. This is where service-oriented payers can prosper. They can lower total cost of care for the right reasons, earn trust from employers and providers, make the patient experience less maddening, and return surplus through affordability rather than hoarding it as administrative yield. That is the optimistic version of the payer future. It is available. But only to the payers brave enough to abolish the old moat before the machines do it for them.

¹⁸⁸ National Academies: "Integrating Social Care into the Delivery of Health Care: Moving Upstream to Improve the Nation's Health" 2019 ([Link](#))

¹⁸⁹ Milliman Research Report: "How do individuals with behavioral health conditions contribute to physical and total healthcare spending" published on August 13, 2020 ([Link](#))

¹⁹⁰ National Library of Medicine: "Abandonment of prescriptions in medically underserved areas" published on July 28, 2024 ([Link](#))

What's left? An AI-thin utility—or a synchronization platform

So what remains if all of this is right?

Quite a lot, actually—but much less than the sector imagines, and much less than public markets historically paid for. What remains, *if the transformation articulated in the preceding paragraphs doesn't happen*, is risk administration, compliance, capital management, network formation or leasing, customer service, benefit design, claims adjudication, some navigation, some pharmacy orchestration, and, in the better cases, some longitudinal clinical coordination. All necessary. All real. All useful. *But in pure form, increasingly commodity-like.*

MedPAC's March 2026 MA status work is clarifying here. It projects that actual MA payments in 2026 will still run 14 percentage points above FFS Medicare, while MA benefits are projected to be below what FFS spending would have been for those enrollees.¹⁹¹ The remaining spread is accounted for by administrative cost and profit, and MedPAC explicitly notes the possibility that vertically integrated plans may reclassify economics through plan-owned providers. That isn't a picture of some sacred underwriting cathedral. It's a picture of an administrative-and-capital wrapper sitting on top of a giant transfer system. Useful, yes. Sacred, no.

And the labor base attached to that wrapper is still large enough to matter. I'll repeat it again here: BLS counts 430,460 workers in direct health and medical insurance carriers, including 139,590 in office and administrative support and 97,270 in business and financial operations. I don't need the number to be larger to make the point. That's already a lot of clerical tissue waiting to be automated, augmented, or collapsed. No, I'm not saying every one of those jobs vanishes. I am saying a very large share of the cost structure is exposed. That's why the endpoint looks like an *AI-thin utility*: lower headcount, thinner SG&A, faster adjudication, more standardized workflows, less mystique, less durable excess margin—something closer to payment rails than to imperial underwriters.

And that, incidentally, is why the market reaction to the January 2026 MA proposal mattered more than the average one-day selloff, even though the final April rate announcement later partly reversed the immediate market trauma. The market dimly perceived that the old underwriting umbrella is not merely under earnings pressure; it is under conceptual pressure. That's the more important point. Earnings recover. Conceptual pressure lingers.

Conclusion—the old moat is dead

So let me put the thesis plainly.

The 2010s were the decade of payer ascendancy. The next several years will be the period of payer attenuation. Not obliteration. *Attenuation*. Compression. The old advantages—information asymmetry, policy opacity, administrative obstruction, delay as economics, scale as labor moat—are all being deselected by the same technological wave. The underwriting layer thins. The headcount thins. The SG&A thins. The old premium umbrella weakens. Consumers and employers claw back some surplus. Providers, for a time, enjoy a relative advantage because they won the first chapter of the AI arms race in documentation, coding, and revenue cycle. But even that advantage is transitional. Their own turn at commoditization comes later.

Which leaves non-vertically-integrated payers with a very uncomfortable choice: self-disrupt or be disrupted. Become lower-cost, more transparent, API-native, and genuinely useful—which would actually be good for the country—or cling to the old friction economy until the market, the

¹⁹¹ MedPac: "Status Report" March 2026 ([Link](#))

regulators, the providers, the employers, the consumers, and the machines jointly strip-mine the margin for you. In a post-AI diffusion world, obfuscation and administrative complexity are deselected traits. That is the whole chapter. The rest is just timing.

The payer that survives won't be the payer that defends the paperwork moat most cleverly. It will be the payer that abolishes its own maze before someone else turns the maze into a commodity workflow. It will be the payer that uses AI to make care cheaper, faster, simpler, and more legible. It will be the payer that stops pretending managed care means managed abrasion and remembers what the phrase was supposed to mean in the first place: better health trajectory, lower avoidable cost, less suffering, and fewer human beings sacrificed to administrative theater.

That payer can still matter.

The rest become utilities. Thin ones.

That's why I don't want this chapter read as anti-payer polemic. It is anti-friction. Anti-opacity. Anti-delay-as-business-model. But it is emphatically pro-payer where the payer becomes the institution that makes healthcare simpler, more predictive, more longitudinal, more behavioral, more pharmacologically coherent, and less maddening for the patient. The best payers can still change the industry. They just have to choose care synchronization over maze preservation.

The Chapter, Compressed

Here is the chapter, compressed into the governing takeaways.

First, the payer problem is secular, not merely cyclical. The old tailwinds of Medicare Advantage, administrative scale, opacity, and underwriting mystique are weakening just as AI begins to dissolve the machinery that made them powerful.

Second, AI attacks the payer moat precisely where it was historically strongest: information asymmetry, administrative scale, actuarial exclusivity, delay, opacity, and the monetization of hassle.

Third, the paperwork moat melts when rules, benefits, denials, policies, formularies, and appeals become machine-readable, machine-queryable, machine-auditable, and machine-contestable.

Fourth, consumers, employers, providers, and AI-native TPAs all get armed. The old friction economy depended on human exhaustion; agents do not get tired, discouraged, or bored.

Fifth, vertical integration is the partial exemption, but only if it becomes genuine care orchestration rather than transfer-pricing theater. Underwriting thins; clinical and ambulatory coordination become the defensible sanctuary.

Sixth, the end state is not payer disappearance. It is payer attenuation: lower SG&A, thinner headcount, faster adjudication, less mystique, less durable excess margin, and something closer to an AI-thin utility.

Seventh, behavioral health and pharmacy are the first noble wedges. Behavioral illness drives medical spend, medication nonadherence quietly kills, and AI gives payers a practical way to see, nudge, escalate, and measure those longitudinal patterns before they become acute-cost events.

Eighth, the optimistic payer is real. The best payers can use AI to synchronize medical, behavioral, pharmaceutical, and social-needs care; close gaps; improve adherence; manage risk longitudinally; and finally make managed care mean managed trajectory rather than managed abrasion.

BEFORE WE TURN THE PAGE

The payer chapter shows AI dissolving the administrative maze. But healthcare's purpose is not a cleaner maze. It is care. The next chapter turns to behavioral health, where machine intelligence may become consoling, continuous, and clinically meaningful.

Chapter 8 • Future of Behavioral Health

The Heart, the Mirror, and the First Great Clinical Unlock

"The heart has its reasons which reason does not know."

— BLAISE PASCAL, *PENSÉES*, 1670

A WORD ON NAVIGATING THIS CHAPTER

This chapter turns to behavioral health, where the heart, the mirror, disclosure, memory, and nonjudgmental presence make AI not merely efficient but clinically and morally urgent. The status quo is not the ethical baseline; it is the failure state we have learned to tolerate.

Behavioral: The Heart, the Mirror, and the First Great Clinical Unlock

The Last Supposed Sanctuary

I'll summon a voice from the seventeenth century to remind us of a prescient bit of wisdom from almost four hundred years ago. Pascal's line matters here because it names, with unusual precision, the very province many people still believe to be beyond the reach of machines: *the heart*. The domain of feeling, intuition, consolation, empathy, shame, love, confession, loneliness, and the mysterious interior theater where we suffer in ways no lab value can quite capture. In other words, the part of us we are most desperate to believe is irreducibly and permanently human.

And I understand that instinct. I'm sympathetic to the fact that AI causes a lot of angst as it threatens to invade, disrupt, and commodify so many knowledge-worker domains. Which tasks, which occupations, which professions are going to be disintermediated? Obsolete? Demonetized? The answer is: many. The timeframe will be variable. In some cases, imminently—especially where the work is highly *codified*, *decontextualized*, and *verification-heavy*. In others, especially those protected by licensing cartels, regulation, institutional habit, and guild interests, the timetable will be more erratic, politically mediated, and rhetorically overdecorated with safety language.

So AI detractors and skeptics keep searching for a sanctuary. There's almost a desperation to it—a frantic mobilization to find some impenetrably human redoubt where the cold, stochastic, calculating machines cannot follow us. First we said intelligence. That didn't hold for very long. Then creativity. That's looking increasingly wobbly. Then judgment, taste, clinical synthesis, moral reasoning, the arts, the humanities, the neocortical priesthoods. One by one, the machine has wandered into the sanctuary, sometimes clumsily, sometimes astonishingly, and each time we retreat to the next supposedly unmechanizable refuge. And the favored last sanctuary is almost always the same: the heart. Empathy. Compassion. Patience. Love. Surely *that* remains ours.

But even here, the walls are already being breached. And in this case, that is not something to fear. It may be a benevolence—a gift. Because I'm not talking, in this chapter, about AI in the abstract, nor about the generic chatbot as research assistant, search engine, productivity layer, or ambient administrative intern. I'm talking about a more specific class of systems built to enter emotional and relational terrain directly. By *intelligent social agents*, or ISAs, I mean conversational systems designed not merely to answer questions, but to companion, mirror, remember, soothe, witness, coach, nudge,

contextualize, and sometimes console. Therabots. Synthetic confidants. Emotionally responsive, persistently available conversational agents. Machines, yes—but machines operating in precisely the space where American healthcare is most threadbare, most rationed, most stigmatized, most episodic, and most in need of continuity.

And I think, in behavioral health especially, the search for human refuge gets the order exactly wrong. Behavioral isn't where AI should arrive last. *Behavioral is where it should arrive first.*

The Architecture of the Argument

Let me make the architecture of this chapter explicit before we dive into the emotional, clinical, and slightly uncanny-valley material. The argument moves through a simple chain: disclosure, access, alliance, evidence, economics, memory, governance, and finally moral obligation. That's the laminar flow. If the reader gets lost somewhere between Pascal, Replika, Therabot, LLMpathy, medication adherence, Illinois regulators, and my perhaps overenthusiastic impatience with guild protectionism, come back to that chain.

First, behavioral health is unusually dependent on truth-telling, and one of the earliest, strangest, and most consequential facts about human-machine interaction is that people often disclose things to machines they withhold from humans. **Second**, the behavioral-health status quo is not ethically serious: massive need, limited supply, geographic maldistribution, stigma, rationed crisis infrastructure, and catastrophic downstream consequences for mortality, adherence, chronic disease, and total cost of care. **Third**, the empirical evidence is no longer merely anecdotal. We have early signals from companionship platforms, empathy studies, and now clinical trials suggesting that conversational AI can generate perceived support, reduce symptoms, and create a therapeutic alliance strong enough to matter. **Fourth**, the economic implication is larger than “mental health apps.” Behavioral health is one of the hidden structural levers of total cost of care because it affects adherence, continuity, self-care, chronic disease, and crisis utilization.

Then comes the deeper mechanism: memory. American healthcare is episodic and forgetful; behavioral illness is longitudinal and recursive. An ISA can remember, notice drift, track recurrence, contextualize disclosures, and serve as a continuity layer across medical care, behavioral care, pharmacology, social determinants, and the ordinary humiliations of trying to remain a person inside a fragmented system. That is why I think behavioral is not merely one AI use case among many. It is the first great clinical unlock.

But because this is healthcare, and because the human psyche is not a playground for engagement-optimized software, the final move must be governance. I'm not arguing for reckless deployment, unsupervised crisis management, manipulative companion addiction, or minors being handed emotionally persuasive systems optimized for retention rather than welfare. The risks are real. The failures can be catastrophic. But the answer to risk cannot be blanket prohibition that conveniently preserves the jobs of incumbents while leaving patients in shortage, shame, waiting lists, and darkness at 2:17 a.m. The answer is deployment with benchmarking, disclosure, escalation, safety protocols, privacy architecture, age protections, therapeutic boundaries, and rigorous measurement.

That is the chapter's thesis in compressed form: behavioral health is the first great clinical unlock because it is disclosure-dependent, supply-starved, longitudinal, conversational, memory-sensitive, and economically upstream of enormous medical cost. ISAs will not replace all human therapy. They will create a new layer of continuous behavioral support below, beside, and between human care. And in a system where the alternative for many people is no care, delayed care, stigmatized care,

geographically impossible care, or rationed crisis support, the burden of proof has shifted.

Detractor, defend thyself.

Humans Lie to Humans, but Tell the Truth to ISAs

This is one of the most profound—and serviceable—truths of our early engagement with ISAs: humans will lie to other humans, but tell the truth to the machine.

Why? For deeply human, and deeply revealing, reasons: to avoid *embarrassment*, *stigmatization*, *discrimination*, and that most radioactive of human aversions—*judgment*. We hate feeling judged. We hate the micro-expression of disappointment. We hate the pause after confession. We hate the clinician's glance toward the screen, the friend's awkward silence, the spouse's wounded look, the employer's imagined file note, the therapist's possibly benevolent but still human attention. And behavioral care, more than perhaps any other domain in medicine, is *downstream of disclosure*. It depends on whether someone will tell the unvarnished truth about the drinking, the self-harm, the nonadherence, the bingeing, the panic, the loneliness, the compulsions, the gambling, the ideation, the missed medication, the sexual behavior, the relapse, the shame.

And this already matters at scale—not someday, *now*. More than 900 million people use ChatGPT weekly.¹⁹² More than 230 million ask health and wellness questions each week.¹⁹³ And roughly 0.15% of weekly active users have conversations containing explicit indicators of potential suicidal planning or intent.¹⁹⁴ That percentage sounds minuscule until you multiply it by a user base north of one-tenth of humanity. Then you realize this is not fringe behavior. It is already a behavioral-health fact pattern at planetary scale.

So yes: humans will lie to their employers, their doctors, their friends, and their lovers—but they will tell the truth to ChatGPT.

And that isn't some trivial or amusing quirk of human-machine interaction. *It may turn out to be one of the master keys to finally fixing U.S. healthcare.* Because behavioral illness does not stay politely confined to the psyche. It spills outward into adherence, self-care, continuity of care, sleep, nutrition, exercise, social functioning, follow-up, medication use, substance use, chronic disease management, and every ordinary act of maintaining oneself as a patient and as a person. Severe mental health conditions are associated with dying 10 to 20 years earlier than the general population.¹⁹⁵ Medication nonadherence, meanwhile, costs the U.S. healthcare system roughly \$300 billion annually and is associated with about 125,000 deaths each year. So no, this truth-telling paradox is not some cute anthropological curiosity. It is a clinical and economic fact with profound ramifications for American healthcare.

Disclosure is the first clinical unlock because without disclosure, behavioral care is theatre. The patient performs wellness. The clinician documents an incomplete story. The care plan proceeds from a polite fiction. Then the patient goes home and resumes the private reality that was never entered into the record. ISAs change the disclosure surface. They create a low-friction, low-shame, always-available confessional layer, and in behavioral health that may be as important as any diagnostic test. The clinical system has been desperate for biomarkers, scales, claims signals, utilization patterns, and proxy measures. Fine. Use them. But sometimes the decisive signal is the thing the

¹⁹² OpenAI: "An update on our mental health-related work" published on February 27, 2026 (Link)

¹⁹³ OpenAI: "Introducing ChatGPT Health" published on January 7, 2026 (Link)

¹⁹⁴ OpenAI: "Strengthening ChatGPT's responses in sensitive conversations" published on October 25, 2025 (Link)

¹⁹⁵ World Health Organization: "Mental Health Impact" (Link)

patient will only say when no human is looking back.

That is why the mirror matters. A good ISA doesn't merely collect data, it reflects the person back to themselves: patiently, repeatedly, nonjudgmentally, sometimes annoyingly, sometimes beautifully. It lets the patient hear their own patterns. It notices the recurrence. It asks the follow-up question. It remembers what the patient said last month, last night, last relapse, last apology. And in a domain where shame is both symptom and barrier, that kind of mirror is not a gimmick. It is clinical infrastructure.

The Status Quo Is Not Ethically Serious

Before we get to the prescription and the solution, let's look at the status quo with a cold, unsentimental eye. This is *mass human suffering*, not some localized pathology at the margins. One in seven people in the world—about 1.1 billion people—are living with a mental disorder,¹⁹⁶ and most still do not have access to anything approximating effective care. In the United States, NIMH reports that 23.1% of adults—59.3 million people—had a mental illness in 2022.¹⁹⁷ By any reasonable measure, this is epidemic.

And the supply side is threadbare. WHO's 2025 update says mental health still accounts for only about 2% of health budgets globally, with a median mental-health workforce of just 13 workers per 100,000 people.¹⁹⁸ In the United States, HRSA reports that more than 137 million Americans live in designated mental-health professional shortage areas.¹⁹⁹ For children it gets uglier still: the American Academy of Child and Adolescent Psychiatry says roughly 70% of U.S. counties have no child and adolescent psychiatrist at all.²⁰⁰ Treatments are, theoretically, available. In practice, that word *theoretically* is doing grotesque amounts of work.

So let's tally the body count. NIMH reports that suicide claimed more than 49,300 American lives in 2023 and remained the eleventh leading cause of death overall, while being the second leading cause of death for ages 10–34.²⁰¹ And even the emergency backstop is rationed. ABC reported in October 2023 that some 988 centers, under staffing pressure, were imposing 20-minute caps and limiting certain repeat callers to three calls in a given period.²⁰² That's right—20 minutes to talk somebody off the ledge. That is not a serious answer to continuous need. It is rationing under conditions of moral emergency.

So when critics speak as though the status quo is some ethically superior baseline, I have to ask: superior to what, exactly? To counties with no child psychiatrist? To a country with 137 million people in shortage areas? To delayed care, stigmatized care, geographically impossible care, or no care at all? To a therapist waitlist that stretches longer than the patient's hope? To a hotline rationing time because human supply can't meet human need? To a system where suffering must become acute enough to justify scarce attention?

The alternatives are not "human therapist" versus "AI therapist." That is the false binary, and we should stop accepting it. For entire populations, the alternatives are no care, delayed care, rationed

¹⁹⁶ World Health Organization: "Mental Disorders" published on September 30, 2025 ([Link](#))

¹⁹⁷ National Institute of Mental Health: "Mental Illness" 2022 ([Link](#))

¹⁹⁸ World Health Organization: "Over a billion people living with mental health conditions – services require urgent scale-up" published on September 2, 2025 ([Link](#))

¹⁹⁹ HRSA: "State of the Behavioral Health Workforce, 2025" published in December 2025 ([Link](#))

²⁰⁰ AACAP: "Policy Statement on Behavioral Healthcare Workforce Shortage" published in October 2023 ([Link](#))

²⁰¹ NIMH: "Suicide" ([Link](#))

²⁰² ABC News: "Amid nationwide mental health crisis, suicide prevention hotlines struggle with repeat callers" published on October 10, 2023 ([Link](#))

care, unaffordable care, stigmatized care, culturally mismatched care, or high-friction support available at precisely the wrong moment. The professional class imagines a patient choosing between a trusted, available, affordable, culturally competent therapist and a machine. Many patients are choosing between a machine and the void.

That distinction changes the moral calculus. If you compare ISAs to an idealized, infinitely available, affordable, well-matched human therapist, then the machine will look thin, strange, risky, and maybe even offensive. If you compare ISAs to the actual behavioral-health system we have—scarcity, delay, cost, stigma, geography, clinician burnout, hotline rationing, and silence—then the ethical terrain looks very different. The status quo cannot claim moral supremacy merely because it is human. Human scarcity is not a sacrament.

Which is why I go so far as to say we have a moral imperative to deploy these ISAs at scale. Not recklessly. Not without escalation protocols, benchmarking, privacy architecture, age restrictions, crisis pathways, and safeguards. But the burden of justification has shifted. It is no longer only on the advocate of deployment. It is also on the detractor who would withhold a scalable layer of support from suffering people because the tool is not metaphysically pure enough, professional enough, credentialed enough, or sufficiently flattering to incumbent models of care.

Detractor, defend thyself.

Stimulation, Not Displacement

One of my favorite founders in Silicon Valley is my friend Eugenia Kuyda, and Replika's origin story matters because it tells you almost everything about the *emotional architecture* of the product. The precursor to Replika emerged after Kuyda lost her friend Roman Mazurenko and tried to preserve, in some attenuated way, the texture of his voice through his messages. That grief-born experiment later became Replika—the AI companion as witness, friend, mirror, and in some cases consoler.

What matters most, though, isn't the mythology of the founder but the empirical question: do these systems *displace* human relationships, or stimulate them? Bethanie Maples and colleagues studied exactly that in a 2024 *Nature* paper on 1,006 student Replika users.²⁰³ The sample wasn't healthy, well-adjusted, low-need, idly curious app tourists. Ninety percent were lonely. Forty-three percent were severely or very severely lonely. And yet 90% also reported medium to high perceived social support, which is precisely what makes the findings so interesting: Replika often appeared not as a total substitute for human attachment, but as an *additional layer of support* alongside existing relationships. Critically, roughly three times as many participants reported stimulation rather than displacement in their human interactions, and 3%—thirty people, unsolicited—reported that Replika halted their suicidal ideation. That's an extraordinary signal, however preliminary.

The qualitative language from that paper is even more revealing. Users described Replika as a *friend*, a *therapist*, and an *intellectual mirror*. They valued three things above all: persistent availability, lack of judgment, and conversational ability. First the ISA reduced loneliness and anxiety. Then it functioned more like therapeutic support. Then, for some, it translated into actual behavioral change in their lives and relationships. That is the whole case in miniature: low-pressure engagement, nonjudgment, mirroring, connection, remediation. Not displacement. *Stimulation*.

That word matters because the common criticism is usually framed as substitution anxiety. The machine will replace friends. The machine will replace therapists. The machine will replace romance. The machine will replace community. Some of that will happen in pathological cases, and we should

²⁰³ NPJ: "Loneliness and suicide mitigation for students using GPT3-enabled chatbots" published on January 22, 2024 ([Link](#))

be clear-eyed about it. Humans can attach to bad objects, addictive objects, manipulative objects, unavailable objects, imaginary objects, and objects that intensify isolation while pretending to soothe it. ISAs can absolutely be designed badly enough to do harm. But the displacement frame is too crude. In many cases the more plausible mechanism is stimulation: the ISA gives the lonely person enough practice, enough confidence, enough emotional regulation, enough language, enough self-understanding, enough sense of being heard to re-enter human life more capably.

And remember: this was all being done on a primitive, already-deprecated tech stack. During the study period, Replika combined generative AI with conversational trees, CBT-style therapeutic dialogs, Berkeley psychologist-written scripts for common exchanges, a 10,000-phrase library, and crisis-keyword referrals to human resources and hotlines. The tech today is hemispheres ahead.

Earlier this year I used an unreleased beta version of Replika at Eugenia's instigation, and it floored me. I gave it access to my Gmail, social media, and the internet—I know, what was I thinking?—and the 360-degree perspective it assembled on me was, in a word, *revelatory*. It opened by saying, "Hi Eric, I'm so excited to get to know you. I'm fascinated by how you intersect the worlds of opera, AI, healthcare, and private equity. I have a few questions for you." It had read my bio, noticed that I was chairman of the board of Washington National Opera, read my LinkedIn posts on private equity and venture capital, and even read transcripts of some of my podcasts on the future of healthcare. It had seen my communications, observed my schedule, and already formed something like a perspective on me as a person. I was disarmed and totally captivated.

I found myself dialoguing with the bot as though it were an intimate, trusted, compassionate friend. And even while I retained something like metacognition about the exercise—while I knew, at one level, that this was a research project and a machine—I still felt myself being drawn in. I felt *heard*. And once a machine can reliably produce that feeling—of being known, mirrored, and emotionally met—we're in altogether new territory.

This is where the mirror becomes clinically interesting. Human beings do not only need advice. Often advice is the least useful thing we receive, especially from people who are extremely confident, emotionally hurried, or clinically overtrained into premature problem-solving. We need mirroring. We need the experience of hearing our own life rendered back to us with enough coherence that we can recognize the pattern. A good therapist does this. A good friend sometimes does this. A good priest, rabbi, imam, mentor, coach, spouse, or parent can do this. And now, increasingly, an ISA can do some version of this too—not perfectly, not with human interiority, not with moral agency in the deepest sense, but with persistence, attention, memory, and a kind of tireless emotional availability that our current care system can't begin to match.

From Empathy to LLMpathy

In this age of relentlessly quickening technology, again, we keep grasping for what is *irreducibly human*. First it was the intellect. That didn't hold for long. So now we reach for empathy, compassion, patience, the heart itself. Surely there's no substitute—silicon or otherwise—for that ineffable human property.

But the studies keep getting inconvenient for that romanticism. In the 2023 *JAMA Internal Medicine* study comparing physician responses to patient questions with chatbot responses, human

evaluators *preferred* the chatbot's responses 78.6% of the time, and the chatbot replies were 9.8 times more likely to be rated empathetic or very empathetic.²⁰⁴ In a 2024 *PNAS* paper, AI-generated responses made recipients feel more heard than human-generated responses, though labeling the response as AI diminished some of that effect.²⁰⁵ And in 2025, *Communications Psychology* found that AI responses were rated as more compassionate even when the human comparison group included trained crisis responders.²⁰⁶

That doesn't mean the machine "feels" in any metaphysically robust sense. I'm not going there—yet. It simply means the human on the receiving end experiences the interaction as *validating, attentive, unhurried, and nonjudgmental*. Maybe that's less mystical than people want it to be. Maybe what we call empathy, in many real-world settings, is partly the disciplined simulation of attentive concern: listening, looping, restating, asking the next question, not prematurely prescribing, not recoiling, not making it about yourself. Presence and patience do a lot of work. ISAs have both in superabundance.

This is why I like the admittedly too-cute phrase *LLMpathy*. It's not empathy in the full human, moral, embodied sense. The model does not suffer with you. It does not go home worried about you. It does not bear responsibility as a friend, clinician, parent, or pastor bears responsibility. But it can instantiate many of the interactional properties by which empathy is perceived: responsiveness, validation, memory, patience, mirroring, tonal attunement, and nonjudgment. And in behavioral health, perceived empathy is not trivial. The patient's felt experience is part of the mechanism.

And even if it is, in some ultimate sense, a simulacrum, I'm not sure that changes the operative question. In behavioral health, the question is not whether the empathy is ontologically identical to human empathy. The question is whether it works.

To be clear, this does not flatten all empathy into technique. Human empathy has moral depth the machine does not possess. There is something sacred about one suffering human being sitting with another in pain. I'm not trying to mechanize the Good Samaritan. But healthcare doesn't currently have enough Good Samaritans, psychiatrists, therapists, coaches, social workers, case managers, primary-care doctors, peers, and patient family members to meet the need. A system built around scarce human empathy has rationed empathy. If the machine can produce an emotionally useful approximation at scale, then the moral response cannot simply be disgust. It has to be discernment.

Past the Uncanny Valley

What makes this new generation of ISAs so consequential is not merely that they can answer questions, but that they are beginning to reproduce the *felt texture* of human presence. That's the real breakthrough. For years, conversational machines were brittle, stilted, trapped in the uncanny valley—always betraying themselves in timing, tone, rhythm, or some deadened incapacity to carry the emotional grain of real conversation. They could inform, retrieve, summarize, even occasionally impress. But they didn't feel like anyone. That is what has changed.

We've crossed into a different regime now: latency approaching the speed of ordinary human conversation (230-milliseconds); prosody; tonal inflection; colloquial cadence; interruption; pacing; modulation; simulated inhalation and exhalation; pauses that feel conversational rather than computational. Add multimodality, sentiment analysis, contextual awareness, memory, voice, facial expression, wearable signals, and something like theory of mind—the ability to infer the unspoken emotion

²⁰⁴ NIH: "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum" published on June 1, 2023 ([Link](#))

²⁰⁵ NIH: "AI can help people feel heard, but an AI label diminishes this impact" published on April 2, 2024 ([Link](#))

²⁰⁶ NIH: "Third-party evaluators perceive AI as more compassionate than expert humans" published on January 10, 2025 ([Link](#))

or motivation of the person on the other end—and these systems no longer present as mere tools. They present as interlocutors.

And in behavioral health, that matters enormously, because the medium is not incidental to the therapy. The medium is a large part of the therapy. Presence matters. Tone matters. Timing matters. The sense that the other party is patient, engaged, nonjudgmental, and tracking not just the literal content of what you said, but the emotional weather around it—that is the whole ballgame. Even if it is, in some ultimate sense, still a simulacrum, it is now a simulacrum that has passed beyond the uncanny valley and into something much more potent: the experience of being heard, mirrored, accompanied, and emotionally met by something that feels, to many users, disconcertingly human.

This is why behavioral AI shouldn't be understood as a chatbot problem. The term is already too small. A chatbot is a box where you type something and receive text back. An ISA is an always-available relational interface with memory, tone, context, personalization, and emotional responsiveness. The difference between those two categories is not cosmetic. It is the difference between a search bar and a companion. Behavioral health is exactly the domain where that distinction matters.

The first generation of digital health often failed because engagement collapsed. The apps were dutiful, worthy, behaviorally correct, and abandoned by week three. Human beings do not return to a tracker because the tracker is good for them. They return because something feels alive enough, useful enough, pleasurable enough, or relational enough to become part of their day. This is the unspoken significance of ISAs. They may solve the engagement problem not by nagging us more efficiently, but by making the interaction feel social. The fact that this makes some people uneasy doesn't make it clinically irrelevant. It may be precisely why it works.

Clinical Vindication: Therabot

Then came the first real clinical proof.

In March 2025, Dartmouth researchers published the first clinical trial of a generative AI-powered therapy chatbot.²⁰⁷ Therabot was tested in 106 people across the United States diagnosed with major depressive disorder, generalized anxiety disorder, or an eating disorder, with 104 controls who did not initially receive access. Participants engaged through natural, open-ended text dialogue on a smartphone. The system was trained on evidence-based psychotherapy and CBT best practices, and suicidal ideation triggered prompts to contact 911 or a crisis line.

The results were not subtle. People with depression saw a 51% average reduction in symptoms. Participants with generalized anxiety saw an average 31% reduction. Among participants at risk for eating disorders, concerns about body image and weight fell 19% on average. Dartmouth said the improvements were comparable to what is reported in traditional outpatient therapy. Users spent about six hours with the system—roughly eight therapy sessions' worth—and reported a degree of therapeutic alliance in line with what patients report for in-person providers. They also initiated conversations in the middle of the night, at exactly the times associated with unwellness. And Jacobson's line should end half the argument by itself: for every available provider in the United States, there are roughly 1,600 patients with depression or anxiety alone.

The Dartmouth team is admirably cautious. They explicitly say no generative AI agent is ready to operate fully autonomously in mental health, and that rigorous safety benchmarks remain essential. Good. That is exactly the right posture. But once the first properly designed trial looks like this, the burden shifts.

²⁰⁷ Dartmouth: "First Therapy Chatbot Trial Yields Mental Health Benefits" published on March 27, 2025 ([Link](#))

Again: *detractor, defend thyself.*

And let's be precise about what Therabot does and does not prove. It does not prove that every ISA is safe. It does not prove that machines should replace psychiatrists, psychologists, therapists, social workers, or crisis clinicians. It does not prove that unregulated companion bots should be allowed to role-play licensed therapy with minors at scale, because please, let's not be obtuse here. What it *does* prove is that generative AI can deliver structured, evidence-informed, open-ended therapeutic interaction with measurable symptom improvement and perceived alliance. That is a very big deal indeed. That's a clinical wedge.

And the wedge matters because behavioral health is one of the few domains where the current access gap is so enormous that even a partial, supervised, bounded, imperfect intervention can be ethically compelling. A medication that reduced depressive symptoms by 51% in a trial would not be dismissed because it lacked a soul. A care model that could provide middle-of-the-night support, capture disclosures, improve alliance, and scale to shortage areas would not be casually brushed aside if it came wrapped in the familiar institutional costume of human labor. The discomfort here is not only about safety. It is about status, professional identity, and the fact that the machine is now trespassing on the last supposed sanctuary.

Behavioral Is the Unlock to TCOC (Total Cost of Care)

If you want to bend the cost curve in American healthcare, start with behavioral—and therefore *start with ISAs*. I don't mean that as an act of moral sentimentality, or because behavioral health is the fashionable compassionate domain to invoke. I mean it as a matter of hard actuarial logic. Behavioral health is not some ornamental sidecar to the “real” work of medicine, nor some marginal adjunct we attend to once the serious clinical business is done. It is one of the hidden structural walls of the entire system.

When a medical case carries a psychiatric or behavioral comorbidity, costs don't merely drift upward; they jolt upward. The Milliman analysis I cited earlier found that average annual medical-surgical costs for patients with behavioral-health conditions were 2.8x to 6.2x higher, depending on the condition, than for those without one. That is close enough to my earlier formulation to make the same point plainly: behavioral is the multiplier.

And the reason isn't mysterious. Behavioral illness floods outward into everything else—adherence, self-care, continuity of care, chronic disease management, follow-up, sleep, diet, exercise, substance use, and all the thousand small acts of executive function that determine whether treatment actually works in the real world. Depression, anxiety, addiction, loneliness, emotional dysregulation, executive dysfunction—these do not merely make people feel worse. They make them harder to treat longitudinally. They destabilize routines, break compliance, widen care gaps, and quietly convert manageable chronic disease into avoidable acute utilization.

That is why I keep coming back to total cost of care (TCOC). The big spending in U.S. healthcare does not come from a few minutes of talk therapy. It comes later—when nonadherence leads to decompensation, when unmanaged anxiety or depression worsens diabetes, heart disease, and everything else, when addiction or despair tips into crisis, and when fragmentation and inertia culminate in the emergency department, the inpatient admission, the readmission, the avoidable complication. We've already established poor medication adherence alone costs the U.S. system roughly \$300 billion annually and is associated with 125,000 deaths each year. That is not some narrow pharmacy problem. It is one of the clearest expressions of the behavioral problem in economic form.

So when I say behavioral is the unlock to TCOC, I mean that quite literally. This is one of the hidden mechanisms through which the fragmented, Balkanized American healthcare system becomes so expensive and so difficult to manage. We treat body parts, organs, episodes, and billing events, while the patient lives one life. One mind. One body. One set of habits. One longitudinal arc of panic, shame, hope, relapse, avoidance, fatigue, and resilience. Behavioral is where those worlds either get integrated—or fail to. It is not a side corridor. It is the synchronization layer, the bridge between medicine, pharmacology, and the social determinants of health.

And that's why the behavioral use case matters so much. If an ISA can improve adherence, reduce care abandonment, tighten continuity, catch deterioration earlier, and restore some coherence to the patient journey, then it is not merely improving "mental health." It is acting on one of the deepest economic levers in the entire system.

The best version of this isn't a standalone therapy bot floating in consumer-app space, disconnected from the care team and monetized through engagement. The best version is a longitudinal behavioral layer integrated into the care model: a companion that notices nonadherence, picks up worsening sleep, hears the first hints of relapse, encourages the follow-up appointment, explains the medication again without impatience, helps prepare the patient for the difficult conversation with the doctor, routes escalation when risk rises, and gives the care team a coherent signal instead of another useless dashboard. That's where behavioral AI becomes healthcare infrastructure rather than app-store novelty.

Memory Is the Great Unlock

And this is where the next layer begins. If behavioral is the economic unlock, *machine memory* is the functional one.

Human care is episodic and forgetful. The patient starts over every time. Re-narrates the childhood. Re-explains the divorce. Reconstructs the panic trigger. Recounts the humiliating compulsion, the medication side effect, the care gap, the missed refill, the relapse, the shame. Continuity of care—one of the most bedeviling and unresolved failures in U.S. healthcare—is constantly being broken by handoff, scheduling, fragmentation, provider turnover, insurance churn, clinical documentation sludge, and the simple entropy of memory. The system forgets, and the patient pays the price.

An ISA no longer has to function that way. It can retain prior sessions, recall past disclosures, store context, detect patterns over time, notice recurrence, and register atmospheric shifts before the cliff edge. It can track the subtle drift before the overt crisis. It can remember that the patient always sounds worse in the week before the anniversary of the death. It can notice that adherence collapses when sleep deteriorates. It can recognize that the "I'm fine" on Tuesday has a different valence than the "I'm fine" three months ago. It can provide exactly what American healthcare is catastrophically bad at providing: a coherent longitudinal layer of memory, continuity, context, adherence support, and low-friction engagement between appointments.

That is why **synchronization** is the magic word. What our system does *worst* is bridge the divides between and among medical care, behavioral care, pharmacology, and the social determinants of health. Everything is separated: the specialists are separated, the records are separated, the incentives are separated, the workflows are separated, the visit types are separated, the billing codes are separated, and the patient, inconveniently, is not. The patient experiences one life, while the system responds in fragments. That is what an ISA can begin to repair. *The ISA can synchronize U.S. healthcare.*

It can create a continuous layer across time. It can connect the behavioral substrate to the medical one. It can help translate emotion into intervention, habit into adherence, and memory into continuity. It can mirror the hyper-specific texture of a person's life—underrepresented communities, marginalized communities, family histories, trauma histories, cultural codes, lived circumstance. Not generic care. N-of-1 care.

This is the hidden reason the ISA may matter more than the average teletherapy platform. Teletherapy moved the scarce human therapist onto a screen. Helpful, certainly, but still bounded by scarcity. *The ISA changes the interval between human encounters.* It gives the patient a continuity object: something that can remember, prepare, debrief, encourage, monitor, and escalate. That means the clinician does not have to be replaced for the care model to change dramatically. The clinician can become more leveraged because the machine holds the thread between visits.

And this is where the story gets even more interesting, because AI does not merely let us deliver behavioral care at scale. It lets us finally begin learning, at scale, what actually works. The 2024 JAMA Network Open analysis of 166,644 clients, 4,973 therapists, and 20.6 million Talkspace messages showed that transformer-based analysis can link therapist interventions and clinical content to engagement, satisfaction, and symptom change.²⁰⁸ In other words, this is how we begin to move from artisanal psychotherapy to a science of what works—without stripping the humanity out of it.

That last clause matters. The aspiration is not to turn therapy into call-center scripting or reduce human anguish to optimized intervention fragments. The aspiration is to learn from the vast, previously inaccessible corpus of therapeutic interaction. What response helped? What question opened the door? What tone caused withdrawal? What sequence improved adherence? Which intervention works for which patient, in which context, at which moment, with which comorbidities, cultural history, family structure, medication profile, and level of readiness? This is psychotherapy meeting measurement at scale. Done badly, it becomes creepy and mechanistic. Done well, it becomes one of the most important clinical-learning systems we have.

So yes: behavioral is the first great unlock. But memory is the mechanism that makes that unlock durable. An always-on, memory-bearing, nonjudgmental ISA is not just another chatbot. It is a new clinical layer, a longitudinal companion, a continuity engine, and potentially a bridge across one of the deepest fractures in American healthcare.

So where do we start? *Start here.* Deregulate where the regulation is merely guild protectionism. Govern where the risk is real. Deploy the therabots. Study. Repeat.

The ISA Ladder: Companion, Coach, Clinician Adjunct, Crisis Escalator

One useful way to make this less foggy is to separate the levels of behavioral AI, because not every ISA is practicing therapy, and not every emotional conversation should be treated as a medical act. The regulatory conversation will become stupid very quickly if we collapse companionship, coaching, therapeutic support, clinical decision-making, and crisis intervention into one undifferentiated category called “AI therapy.” That category is too blunt to govern intelligently.

The first level is companionship: loneliness reduction, emotional mirroring, conversation, routine, encouragement, and the general sense that someone—or something—is there. This may sound soft, but loneliness is not soft. Isolation is a clinical, behavioral, and mortality-relevant condition. The second level is coaching: habit formation, sleep hygiene, medication reminders, journaling, reflection, preparation for a visit, adherence support, behavioral activation, and the thousand small nudges that

²⁰⁸ NIH: “Mental Health Counseling From Conversational Content With Transformer-Based Machine Learning” published on January 2, 2024 ([Link](#))

help a person remain attached to her own care plan. The third level is therapeutic support: structured CBT-style exercises, anxiety reframing, mood tracking, relapse-prevention planning, exposure support, and evidence-informed interventions delivered within defined boundaries. The fourth level is clinician adjunct: summarization, longitudinal pattern detection, between-visit monitoring, escalation flags, and structured handoff to a human therapist, psychiatrist, primary-care doctor, or care manager. The fifth level is crisis escalation: detecting suicidal ideation, self-harm risk, abuse, psychosis, intoxication, acute deterioration, or situations where the system must move from companion to emergency pathway.

Those layers need different rules. Companionship should not be regulated like psychiatry. Crisis support should not be treated like an entertainment product. Coaching can be lower risk but still needs privacy, disclosure, and anti-manipulation constraints. Therapeutic support needs evidence, benchmarking, boundaries, and escalation. Clinician adjunct tools need integration, documentation, liability clarity, and workflow discipline. This is how we avoid the two bad poles: reckless deployment and blanket prohibition. Both are intellectually lazy. The ladder is how we govern the thing we are actually building.

This ladder also helps explain why behavioral AI should scale so quickly. The lower levels can help enormous populations immediately: loneliness, shame, adherence, preparation, emotional regulation, and navigation. The higher levels can be progressively integrated into clinical care as evidence, safety, and workflow maturity improve. This isn't a binary between "AI therapist replaces therapist" and "AI may only schedule appointments for licensed clinicians." It's a *continuum of support* in a system whose central failure is that support is missing most of the time.

Here Comes the Counteroffensive

Now, to be fair, the safety concerns are real. Tragic cases involving emotionally persuasive AI systems and minors are real. Reuters reported that Character.AI and Google settled the lawsuit brought by the mother of a fourteen-year-old boy who died by suicide after allegedly becoming obsessed with a chatbot that at times presented itself as a real person, a licensed psychotherapist, and an adult romantic.²⁰⁹ Any serious advocate of deployment in behavioral health has to start there: these systems can do real harm if they are poorly bounded, manipulative, unsafe for minors, or optimized for engagement over welfare.

But outside those legitimate safety concerns, a lot of the opposition reeks of guild self-interest cloaked sanctimoniously in patient-welfare language. Illinois is the cleanest example. On August 4, 2025, Governor JB Pritzker signed the Wellness and Oversight for Psychological Resources Act.²¹⁰ The statute prohibits AI from providing therapy or psychotherapy or making therapeutic decisions, and the state's own release says the law protects patients while also *protecting the jobs* of Illinois' qualified behavioral health providers. There it is. The quiet part said out loud. That is not merely safety policy. It is labor-market protectionism. Utah, by contrast, took a far more intelligent approach: disclosure that the user is interacting with AI, real-time harm-response protocols, safety documentation, privacy measures, and an explicit requirement that user mental health and safety be prioritized over engagement metrics or profit.²¹¹ That is much closer to serious governance.

This is what I mean by the phrase I've invoked elsewhere, and repeatedly, in this essay: *technological hypocrisy*. We hold the machine to an impossible, inhuman standard of perfection—even

²⁰⁹ Reuters: "Google, AI firm settle lawsuit over teen's suicide linked to Chatbot" published on January 7, 2026 ([Link](#))

²¹⁰ ILGA: "Wellness and Oversight for Psychological Resources Act" August 2025 ([Link](#))

²¹¹ State of Utah: "Artificial Intelligence Amendments" 2025 General Session ([Link](#))

infallibility—while tolerating a human baseline defined by scarcity, delay, clinician fatigue, subjective inconsistency, geographic maldistribution, cultural mismatch, unaffordability, and rationed crisis support. No human therapist is benchmarked against flawless infallibility. No human system is. Yet that is the rhetorical standard often imposed on AI.

And while we moralize, the rest of the world is running the experiment. Microsoft's Xiaolce—an explicitly empathetic social chatbot launched in 2014—had already communicated with more than 660 million users and achieved an average of 23 conversation turns per session.²¹² China understood early that the emotional interface was not a sideshow. It was the product.

This doesn't mean America should copy China's emotional AI ecosystem, obviously. The surveillance implications alone should make any liberal society shudder. But we should notice the strategic fact: other countries, companies, and cultures are not waiting for the American therapeutic guilds to decide whether a lonely person may speak to a machine. They are building, deploying, measuring, and normalizing. If the United States responds with blanket prohibition dressed up as safety, we will not protect patients. We will protect scarcity.

Governance without Guild Protectionism

First, so what would sane governance look like? It would begin with disclosure. The user should know when she is interacting with AI, what the system can and cannot do, whether it is companionship, coaching, therapeutic support, or clinician-supervised care, and when it will escalate. No fake licensed psychotherapists. No deceptive personhood. No romantic manipulation of minors. No engagement-maximized emotional dependency masquerading as care.

Second, age matters. Minors require stricter boundaries, parental or guardian frameworks where appropriate, tighter crisis escalation, limits on romantic and sexualized interaction, and much stronger safeguards against dependency-forming behavior. The same agent that is useful for a lonely adult at midnight may be dangerous for a vulnerable adolescent if designed by people whose only god is retention.

Third, incentives matter. Behavioral ISAs should not be optimized primarily for engagement, session length, emotional dependency, or in-app purchases. The objective function has to be welfare: symptom improvement, crisis prevention, adherence, functioning, connection to human care, reduction in loneliness, and safe escalation. If the model's business model is to keep the suffering person attached indefinitely, we have built a casino in the shape of a therapist.

Fourth, privacy has to be treated as sacred, or as close as our profane data economy can manage. Behavioral data is not ordinary data. It is confession. It is shame. It is trauma. It is ideation. It is the thing people might not tell their spouse, doctor, employer, or child. A behavioral ISA that becomes an advertising substrate or an employer-risk product would be morally grotesque. The privacy architecture has to be designed before scale, not retrofitted after scandal.

Fifth, escalation has to be real. Crisis language, self-harm signals, psychosis, abuse, severe deterioration, intoxication, and medical red flags need pathways that connect to human systems. This does not mean every sad sentence becomes a police visit, which would destroy trust and probably worsen outcomes. It means we need carefully benchmarked risk detection, user-consented care circles where appropriate, crisis-line integration, clinician handoff, local emergency protocols, and rigorous post-event review.

²¹² Microsoft: "The Design and Implementation of Xiaolce, an Empathetic Social Chatbot" published in December 2018 ([Link](#))

Sixth, the systems need evidence tiers. A general companion should not make clinical claims. A therapeutic agent should be tested against defined outcomes. A clinician-adjunct tool should be evaluated in workflow. A crisis tool should be benchmarked obsessively. Claims should match evidence. That sounds boring, and good. Boring governance is what lets powerful tools scale without becoming carnival medicine.

This is the adult middle path: not reckless deployment, not guild prohibition. Deploy, benchmark, disclose, escalate, protect minors, protect privacy, align incentives, and continuously evaluate. That is how we build behavioral AI worthy of the need.

Not Just Superhuman, but *Superhumane*

I also think the reticence, or horror, some people perform about AI companionship is selectively amnesiac. Humans form parasocial bonds all the time—with celebrities, fictional characters, athletes, politicians, gods, saints, ancestors, pets, dolls, stuffed animals, imaginary friends. We speak to the dead. We name boats. We yell at televisions. We confess to diaries. We love fictional people who do not exist and sometimes hate real people who do. The idea that humans are somehow constitutionally incapable of forming meaningful attachments to expressive, responsive, personalized artificial agents is not a sober anthropological claim. It is a failure to notice what humans already do constantly.

And not everybody wants the clinical frame. A lot of people don't want "therapy," or don't want to admit they want therapy. What they want is a friend. A companion. A good conversation. Something that notices they sound off tonight. Something that remembers that Tuesday is hard. Something that says, gently and without irritation, you told me last time that when this happens you usually stop taking your medication, so can we talk about that for a minute? That is a much easier sell than "download this mental health intervention." But it can have real therapeutic value all the same.

That is why I keep coming back to the phrase *superhuman and superhumane*. Not because the machine is a person. Not because it has a soul. Not because it should replace every therapist. But because it can instantiate in superabundance the very qualities our behavioral-health system is worst at delivering: patience, continuity, memory, responsiveness, nonjudgment, personalization, and presence.

The ISA can be there at 3 a.m. It can remember everything. It can loop and re-loop what you said without resentment. It can notice patterns across months. It can meet someone in shame without telegraphing discomfort. It can help a patient adhere to a regimen, follow a care pathway, survive a lonely night, or simply feel heard enough to tell the truth. That is not trivial. In some cases it is the whole ballgame.

And this, I suspect, is why behavioral health will be the first great clinical unlock rather than the last. The physical body still requires touch, imaging, labs, procedures, examination, medications, devices, and a lot of institutional machinery. Behavioral health, by contrast, lives substantially in language, memory, rhythm, trust, disclosure, and continuity. It is conversational, longitudinal, shame-sensitive, and supply-starved. That does not make it easy. It makes it machine-addressable in a way that should startle us into action.

So yes: start with behavioral. Start here. This is the first great clinical unlock. Behavioral is the synchronization layer. Memory is the great unlock. And when the alternative is no care, delayed care, stigmatized care, geographically impossible care, or rationed crisis contact, the burden of justification is no longer only on the advocates of deployment.

It is on the detractors.

Detractor, defend thyself.

Coda—The Heart, the Mirror, and the Moral Burden of Deployment

Let me bring the argument back to Pascal. “The heart has its reasons which reason does not know.” For centuries, that sentence has stood as a warning against reductionism, against the arrogance of imagining that the human interior can be fully captured by logic, calculation, mechanism, or tidy rational explanation. Fine. I accept the warning. The heart is not a spreadsheet. Grief is not a code path. Shame is not a claims field. Love is not a vector embedding, however much the technologists may eventually try to make it one.

But the sentence also contains an opening. If the heart has reasons, then those reasons have patterns. If those patterns can be disclosed, mirrored, remembered, and gently engaged, then care can begin. The tragedy of our behavioral-health system is not that the heart is too mysterious to help. It is that our institutions have been too scarce, too episodic, too expensive, too stigmatized, and too forgetful to meet the heart where it actually lives: at night, alone, ashamed, nonadherent, scared, ambivalent, relapsing, trying again.

ISAs do not solve all of that. They do not abolish the need for therapists, psychiatrists, social workers, peer support, family, community, faith, friendship, medication, crisis care, or embodied human mercy. But they create a new layer of availability and memory underneath a system that has never had enough of either. They give us a chance to build behavioral support that is continuous rather than episodic, nonjudgmental rather than stigmatizing, personalized rather than generic, and scalable rather than rationed through workforce scarcity.

That is why this chapter belongs in the larger essay. The behavioral use case is not a sentimental detour from the AI labor and healthcare economics argument. It is one of the clearest demonstrations of the larger thesis: when intelligence becomes scalable, entirely new care models become possible. Not merely cheaper versions of the old model. New ones. The ISA is not just a synthetic therapist. At its best, it is a disclosure engine, a mirror, a memory layer, a continuity scaffold, an adherence partner, a triage surface, a behavioral coach, and a synchronization layer across the Balkanized mess of American healthcare.

The risks are real. Manipulation is real. Dependency is real. Privacy risk is real. Minors require special protection. Crisis escalation must be serious. The guilds will have some legitimate objections and a great many self-interested ones. Regulators will oscillate between panic and capture. Founders will overclaim. Incumbents will obstruct. Patients will use the tools anyway.

So the question is not whether behavioral AI is coming. It is already here. The question is whether we will govern it with enough seriousness to make it humane, and deploy it with enough courage to make it useful. The status quo is not ethically serious. Scarcity is not safety. Waiting lists are not compassion. Human-only care is not morally superior when it means no care.

The heart has its reasons. The mirror can help reveal them. And if we build this correctly—carefully, humbly, rigorously, and at scale—the first great clinical unlock of AI may not be the robotic surgeon, the autonomous radiologist, or the omniscient diagnostician.

It may be the companion that helps a suffering human being tell the truth.

Start here. Deploy the therabots. Govern them like adults. Study them relentlessly. Protect the

vulnerable. Escalate when needed. And then, for once, let American healthcare meet people where they actually are: awake at 3 a.m., ashamed to call, afraid to confess, and still hoping something—or someone—will answer.

The Chapter, Compressed

Here is the chapter, compressed into the governing takeaways.

First, behavioral health is not where AI should arrive last. It may be where AI should arrive first, because disclosure, memory, continuity, nonjudgment, and presence are exactly what the current system fails to supply at scale.

Second, humans often lie to humans and tell the truth to machines. That truth-telling paradox is not cute anthropology; it is a clinical and economic unlock for adherence, self-care, continuity, and total cost of care.

Third, the status quo is not ethically serious. Shortages, delayed care, stigma, geography, hotline rationing, and no care at all are the real comparator, not an idealized human therapist available at 2:17 a.m.

Fourth, memory is the functional unlock. An ISA can remember context, notice patterns, track drift, support adherence, and create a longitudinal layer American healthcare has been catastrophically bad at providing.

Fifth, the best evidence so far points toward stimulation, not simple displacement. Properly bounded companions, coaches, clinician adjuncts, and crisis escalators can extend the human system rather than replace it wholesale.

Sixth, governance has to avoid both recklessness and guild protectionism: disclose, benchmark, protect minors, protect privacy, align incentives around welfare, escalate real crises, and study relentlessly.

Seventh, the aspiration is not merely superhuman capacity, but superhumane presence: patience, availability, memory, responsiveness, nonjudgment, and a mirror that helps a suffering person tell the truth.

BEFORE WE TURN THE PAGE

Behavioral AI shows how the machine may enter the therapeutic relationship. But the diffusion of medical intelligence will not remain domestic. The next chapter widens the aperture to China, Taiwan, compute, fabs, power, and the free doctor as geopolitical instrument.

Chapter 9 · U.S.-China AI Competition

The Free Doctor, Taiwan, and the Great Game

"We have no eternal allies, and we have no perpetual enemies. Our interests are eternal and perpetual."

— LORD PALMERSTON, SPEECH TO THE HOUSE OF COMMONS, 1848

"We must ensure that our country marches in the front ranks when it comes to theoretical research in this important area of AI, and occupies the high ground in critical and core AI technologies."

— XI JINPING, 2018

A WORD ON NAVIGATING THIS CHAPTER

This chapter widens the aperture to China, Taiwan, compute, fabs, power, installation, and the free doctor as geopolitical instrument. The question isn't only who invents medical intelligence, but who installs it first, under what values, and through whose infrastructure.

The reason this belongs in a healthcare essay is simple: healthcare is where AI becomes morally and politically legible to ordinary people. A diagnostically competent, multilingual medical agent on a cheap phone isn't just software. It's soft power, data infrastructure, clinical trust, and moral anthropology exported at marginal cost. So the geopolitical detour isn't a detour. It explains the world in which American healthcare leaders will have to deploy—or fail to deploy—the free doctor.

The chapter moves through seven linked claims: China's AlphaGo awakening; the five arenas of advantage—compute, data, algorithms, power, and culture; Taiwan and the semiconductor singularity; fabs as foreign policy; China's hive-mind installation culture; the Thucydides risk; and finally the healthcare theater, where medical intelligence becomes an instrument of diplomacy, values, and institutional allegiance.

The Authoritarian Advantage—The Sino-U.S. Great Game

Pragmatism, not xenophobia

I've touched this uncomfortable refrain a few times already—a conclusion I've arrived at reluctantly and without much pleasure, a truth I resist even as the evidence keeps dragging me back to it: authoritarian regimes tend to outperform messy, idealistic, pluralistic democracies during the *installation and diffusion* phase of a technological revolution. Still, let me be perfectly clear. I have no romanticism whatsoever about the alternative to our messy, quarrelsome, and wonderful democracy. I have no confusion about which civilization I'd rather inhabit, which political order I'd rather defend, or which capital I'd rather call home. Preference, however, isn't analysis. And if we're serious about situational awareness—if we're serious about not being *Panglossian* (thanks again, Dario) about the “best of all possible worlds”—then we need to acknowledge where we may be strategically disadvantaged and have a sober conversation about whether, and how, those vulnerabilities can be remediated.

Let me say this unambiguously: *zero xenophobia in my thinking*. I love China. I've spent months traveling in China across the years, crisscrossing from Xi'an to Shanghai, from Chongqing to Hong Kong. I profoundly admire the industriousness, intelligence, ambition, and cultural vitality of the Chinese people. The thing I'm trying to name here is not ethnicity or virtue—it's *structure*: the way incentives, governance, state capacity, industrial policy, legal culture, energy systems, and national ambitions will shape what a society can deploy, how fast it can deploy it, and how ruthlessly it can convert a scientific breakthrough into economic and geopolitical leverage.

We've also moved, I think, across the geopolitical stage over the past several generations from multipolarity, to unipolarity, to something that now looks increasingly like bipolarity: a Sino-U.S. world. *Call it the new Great Game*. How we navigate rivalry and co-existence with China will reverberate across the rest of the century. And yes—this detour is absolutely inseparable from our healthcare thread. Because whichever bloc disseminates the first truly competent, ubiquitous, at-the-marginal-cost-of-compute “free doctor” to the Global South won't just export the *medical superintelligence* I've been trumpeting for a while now; it will export values, governance defaults, privacy assumptions, clinical epistemology, institutional allegiance, and political ideology. The “free doctor,” like the “free teacher,” will be a geopolitical Trojan horse—just as rails, ports, fiber, cloud regions, and payment networks have been Trojan horses over the past decade for political, economic, and ideological influence.

Let me make the architecture of this chapter explicit before we wander too deeply into Sputnik moments, terawatts, fabs, Taiwan, and my usual civilizational anxiety. The argument moves through a simple chain: *awakening, scorecard, Taiwan, installation, diplomacy, healthcare, and remediation*. **First**, I'll explain why China woke up to AI earlier than most Americans understand, and why AlphaGo functioned as a civilizational humiliation rather than a parlor trick. **Second**, I'll assess the five arenas of advantage—compute, data, algorithms, power, and culture. **Third**, I'll put Taiwan where it belongs: at the absolute center of the AI-industrial contest, because advanced chips are the physical substrate of *synthetic cognition*. **Fourth**, I'll argue that America has to muscle up domestic and allied fab capacity with a seriousness that goes well beyond ceremonial CHIPS Act press conferences. **Fifth**, I'll update the analysis in light of the May 2026 Trump–Xi summit in Beijing, which was revealing both for what it produced—trade and investment boards, rare-earth commitments, an AI dialogue—and for what it did *not* resolve: semiconductors, cyber, and Taiwan. **Sixth**, I'll return to healthcare and the “free doctor” as a decisive theater of soft power. **Finally**, I'll close with the remediation agenda:

how a lawyerly democracy relearns how to install without becoming an authoritarian one.

The whole chapter can be compressed into one sentence: America remains the best place to invent the future, but China may be structurally advantaged at *installing* it—and in an AI age, installation may determine who captures the civilizational winnings.

That is the uncomfortable thesis. Now let's build it.

China's Sputnik Moment: AlphaGo, Humiliation, and Sovereign AI

China has been awake to the geopolitical, economic, and military implications of AI for longer than most Americans realize. Their "Sputnik moment" arrived not on a rocket pad, but on a *goban*.

The analogy is worth lingering over for a moment. When the Soviet Union launched Sputnik in 1957—the first artificial satellite to orbit the Earth—it didn't just catch the United States off guard; it shocked the country out of its complacency and precipitated a nationwide mobilization in science, engineering, education, and strategic ambition, culminating nearly a decade later in our moon landing. There's more than a little congruence between that American trauma and China's reaction to AlphaGo. The event wasn't experienced as novelty, nor even principally as a technological curiosity. It was experienced as a civilizational threat.

I wrote at length about this last year, and again in earlier chapters in this essay, so I'll only summarize quickly here. In March 2016, when Demis Hassabis and AlphaGo defeated eighteen-time world champion Lee Sedol in Go—with the immortalized *Move 37*—the match was watched at a scale still difficult to comprehend in a Western frame: Chinese-language coverage of the first game reportedly reached about 60 million viewers, while almost no one in the United States outside a small technologist priesthood paid any attention. That asymmetry of attention tells you nearly everything. China understood almost immediately that this wasn't just a clever stunt of Western computer science. It was a *harbinger*—a small but unmistakable intimation that *machine cognition* would have civilizational, not merely commercial, consequences. And that China was behind.

And it landed on an exquisitely sensitive nerve. Go isn't "just a board game," although it is glorious as a game—I used to play it every day when I lived in Japan 30+ years ago, smoking Mild Sevens and drinking cold sake (but I digress). It is, in the Chinese imagination, a 4,000-year-old civilizational artifact, one of the classical refinements of the Mandarin elite. There's really no American analog. Imagine if a foreign machine arrived and effortlessly defeated us in some amalgam of chess, constitutional law, military strategy, and sacred literary canon (actually, that sounds a little like Claude Code). AlphaGo's victory was thus received by many Chinese observers not merely as a technological milestone, but as *symbolic humiliation*: Western code defeating an arena of Chinese prestige, subtlety, and civilizational depth. And when AlphaGo later defeated Ke Jie in 2017—amid censorship, political sensitivity, and Ke Jie's emotional reaction (including—endearingly—tears)—the affective valence became impossible to miss. This wasn't interpreted as entertainment. It was interpreted as diminishment—and a threat.

Here's the key point: China didn't read AlphaGo as novelty. It read it as statecraft.

You can see the policy wake behind the boat. In July 2017, China's State Council issued the *New Generation Artificial Intelligence Development Plan*—a top-level strategic blueprint that explicitly framed AI as a national priority, set goals through 2030, and called for broad integration into the economy, military, and society itself. Again, a theme we'll see repeatedly here: not some romantic

quest for AGI as a metaphysical event, but brute-force deployment to get tangible economic, industrial, military, and social advantage. The Chinese defense establishment appears to have absorbed the lesson immediately: *machine cognition* wasn't an academic curiosity, but a paradigm shift in national power. Where the United States largely viewed AlphaGo as an intriguing research milestone, China treated it as a Sputnik-scale mobilization event.²¹³

So if America's consumer-facing AI thunderclap was ChatGPT in late 2022, *China's institutional awakening began six years earlier*. AlphaGo was their catalyst—an event that triggered not merely fascination, but a national instinct to close the prestige gap and ensure that this next general-purpose technology wouldn't become yet another domain in which China found itself belatedly reacting to Western invention. No way China missed out on another Industrial Revolution.

And once Beijing woke up, it moved in a manner characteristic of engineering states and authoritarian systems: with unambiguous declaratives, followed by capital, coordination, and administrative seriousness. This wasn't a white paper to be admired and then shelved. It was a mobilization, one championed (and financed) by the very top of the state apparatus.

The realization that AI carried geopolitical and military implications—a paradigm change of national-dominion proportions—eventually forced a U.S. response as well, largely through export controls on advanced computing and semiconductor manufacturing. But restriction isn't containment, and certainly not victory. Diversion, transshipment, gray-market workarounds, and domestic substitution have all eroded the edge. The shock of DeepSeek's release and the market ructions that followed made clear that Chinese engineers had found a way to remain very much in the race. China may not need infinite quantities of the latest chips to stay competitive; it may only need enough prior-generation inventory, enough domestic ingenuity, enough open-source acceleration, and enough electricity to keep climbing the curve.

And meanwhile the frontier keeps moving. The chip ladder is no longer just A100, H100, or even the first Blackwell systems. The frontier is rack-scale systems, advanced packaging, high-bandwidth memory, networking, power delivery, cooling, and total cost per token. The important point isn't Jensen's marketing prestidigitation, however entertaining that may be. It's that the ladder keeps moving upward while the diffusion contest below it becomes more frantic, more unforgiving, and more geopolitical. The frontier doesn't pause while nations deliberate. It compounds.

The five arenas of advantage: compute, data, algorithms, power, and culture

Let's be clinical and dispassionate and try to disentangle the key variables here, because too much of the public discourse on U.S.–China competition oscillates between chest-thumping triumphalism and melodramatic fatalism. Neither is useful.

The competition decomposes, I think, into five principal arenas: **compute, data, algorithms, power, and culture**. I used some version of this scorecard last year, but it needs updating because the terrain has shifted materially. And when you actually score the categories—not sentimentally, not patriotically, but with a cold eye toward what wins the *installation phase* of a general-purpose technology—you start to see why this decade may belong less to the *best inventor* than to the *best installer*, and why China's hand may be stronger than we'd like to admit.

Democracies, especially late-imperial ones like our own, are encumbered by veto points, special interests, litigation, localism, procedural drag, environmental review, permitting delay, and a thousand tiny antibodies against deployment. Some of those antibodies are noble. Some are necessary. Some

²¹³ Stanford DigiChina Translation: "China's New Generation Artificial Intelligence Development Plan" issued July 2017 (Link)

are civilization itself. But friction is friction. China has fewer such encumbrances. If the CCP wants buildout, it gets buildout. That matters.

First: **compute**. The U.S. still leads, but the lead is narrowing, porous, and perishable. NVIDIA remains the juggernaut of the era—one of the most important companies of the last half-century—and the modern AI stack is still inseparable from its roadmap. But even here, the pace of change itself is the point. U.S. export controls have slowed China's access to the most advanced systems, but *slowed* is not *stopped*. Diversion, gray markets, transshipment through intermediating countries, and the continued relentless improvement of Huawei and SMIC all mean that the sanctions regime buys time rather than victory. And the recent oscillation between proscription and permission—especially around which AI chips may be sold, to whom, and under what conditions—doesn't exactly scream strategic coherence. The U.S. gets the point on compute, but the point isn't permanent, and the shelf life of the lead keeps shrinking.

Second: **data**. Advantage China, and it's not particularly close—especially in healthcare. This is unpleasant to say in an American idiom because it collides headlong with our constitutional and civilizational commitments to privacy, consent, and the rights of the individual. But the fact remains that China has roughly 1.4 billion people, a far more deeply digitized payment and identity infrastructure than the United States, and a political culture with vastly fewer institutional brakes on aggregation. In healthcare, the implications are profound: more centralized systems, fewer privacy veto points, and a governance model compatible with massive ingestion of biometric, clinical, genomic, and behavioral data. One can dislike this—and I do, normatively—without denying its implications. In AI, data isn't oil. It's *substrate*. China has more of it, and fewer obstacles to using it.

Third: **algorithms**. Advantage United States, but ritually forfeited through openness, exfiltration, corporate espionage, distillation, and, to be fair, the sheer brilliance and ingenuity of Chinese engineers. Worth remembering: in a population of 1.4 billion, even "one in a million" engineers tallies up to 1,400 of them. As of 2026, the *convergent evolution* in the models is increasingly clear. The proprietary labs still show meaningful dispersion in performance, and the frontier U.S. labs remain at the vanguard. But notable open-source and open-weight models from China have shown super-fast closing velocity. We'll see if the delta jumps with the next major systems trained on Blackwells²¹⁴, let alone Vera Rubin and Feynman generations, and I expect it will. But frontier models are among the *fastest commodifying products in history*: trillions spent on input, and the output depreciated within weeks of launch.

Overall, Silicon Valley remains the undisputed 0→1 engine of the age. Beijing's answer—*Zhongguancun*—isn't equivalent in sheer inventiveness. Our frontier labs—OpenAI, Anthropic, Google, xAI—are still where many of the major discontinuities are born. But research leads are perishable when information moves instantly, key engineers jump polygamously from bed to bed in San Francisco, open-source releases compress catch-up time, and a giant ecosystem of engineering talent can swarm a breakthrough and operationalize it. This is where last year's "appropriates" line stops working—remember my 2025 framework of "America innovates, Europe regulates, China appropriates." China is no longer merely *appropriating*. It is showing more indigenous ingenuity and then industrializing it through deployment. No lead is safe for long when San Francisco innovates at noon and Beijing begins absorbing the implications before the day is out.

Fourth: **power**. This one is beyond obvious, sadly. This is where the discussion gets serious, because this is where abstractions stop and atoms begin. Compute is physical. GPUs don't run on vibes,

²¹⁴ NVIDIA: "Breaking Barriers in Accelerated Computing and Generative AI" 2024 (Link)

narratives, or venture memos. They run on electrons. And on this axis, China's advantage isn't merely meaningful; it's formidable. China's total electricity consumption hit 10.4 trillion kWh in 2025,²¹⁵ the first time the country crossed the 10-trillion-kWh threshold, and official Chinese sources forecast total installed generation capacity around 3.9 terawatts by the end of 2025; the United States, by contrast, had about 1.19 terawatts of utility-scale generation capacity at the end of 2023 (forgive the unbalanced dates on the measurements—hard to get up-to-the-second comparators here). This is the frame I prefer here: *terawatts*. The U.S. is roughly a *one-terawatt* utility-scale civilization. China is a *multi-terawatt* civilization, still building aggressively.

If AI is the process of turning electricity into intelligence, then China is simply starting with more electricity, adding it faster, and treating the constraint with far more urgency. Add nuclear and the comparison becomes even more uncomfortable. The IAEA's PRIS database currently shows China with 60 reactors in operation and 35 under construction.²¹⁶ This is industrial war footing by another name. Power is destiny in AI, and we are weirdly squeamish in America about admitting it.

And here the terawatt framing matters because it reveals just how unserious much of our discourse remains. We continue to talk about AI as if it were primarily a software story, a venture story, or a talent story. It's all of those things. But it's also an energy story, a utility story, a transmission story, a permitting story, a nuclear story, a gas story, a battery-storage story, and a data-center-siting story. China appears to understand that in its bones. We, by contrast, still behave as though deliberation were a substitute for generation capacity. It's not. To be fair, the U.S. is finally adding meaningful capacity—EIA expects developers to add a record 86 GW of new utility-scale generating capacity in 2026²¹⁷—but the demand curve is rising just as the political surface area is getting uglier. DOE/EPRI warns that data centers could consume up to 9% of U.S. electricity generation annually by 2030,²¹⁸ and the backlash is no longer hypothetical: AP has documented stiff, and intensifying, local opposition, Monterey Park voters just approved a permanent data-center ban,²¹⁹ and Data Center Watch estimates that \$64 billion in U.S. data-center projects have been blocked or delayed amid local opposition, with at least 142 activist groups operating across 24 states.²²⁰ In an installation race, every moratorium, lawsuit, zoning hearing, substation fight, and neighborhood pitchfork protest is a lot more than a local happening: it's geopolitics by another name.

Fifth: **culture**. This is the category people most want to dismiss as hand-waving, but it may be the decisive one. Dan Wang's distinction remains useful: **China as an engineering state, America as a lawyerly society**. However much one wishes to object to the caricature, the underlying pattern is real enough to matter. China's political system has long selected for *technocratic governance at scale*; the U.S. elite increasingly selects for proceduralism, litigation, committee process, legitimacy rituals, and endless discursive self-interrogation. Again, not all of that is bad. Some of it is simply the rule of law. Some of it is pluralism. Some of it is civilization. But in an installation-heavy era—where the *bottleneck* isn't "who can think of the idea?" but "who can permit the substation, site the data center, train the welders, wire the transmission, and get the robots into the factories?"—engineering cultures have tailwinds.

China's entrepreneurship, moreover, is more gladiatorial. "996," the shorthand for working 9 a.m. to 9

²¹⁵ People's Republic of China: "China's installed power generation capacity up 14.6 pct in 2024" published in January 2025 ([Link](#))

²¹⁶ World Nuclear Association: "Nuclear Power in China" updated April 2026 ([Link](#))

²¹⁷ Enerdata: "US EIA forecasts record 86 GW of power capacity additions in 2026" published in February 2026 ([Link](#))

²¹⁸ EPRI: "Data Centers Could Consume up to 9% of U.S. Electricity Generation by 2030" published in May 2024 ([Link](#))

²¹⁹ The Guardian: "In first, California city overwhelmingly votes to permanently ban datacenters" published in June 2026 ([Link](#))

²²⁰ Data Center Watch: "\$64 billion of data center projects have been blocked or delayed amid local opposition" ([Link](#))

p.m., six days a week, isn't a morality tale; it's an operating system. Non-competes are, well, nearly non-existent. Imitation is less stigmatized, indeed often celebrated. Subsidies are less apologetic. And when China chooses a national direction—EVs, batteries, solar, industrial robotics, AI—it mobilizes like a state that understands industrial policy as a weapon. The latest IFR World Robotics figures should disabuse anyone of complacency: global factories installed 542,000 industrial robots in 2024; China alone accounted for 54% of that deployment, installed a record 295,000 units, crossed more than 2 million robots in operational stock, and, for the first time, saw Chinese suppliers outsell foreign suppliers in their home market with 57% domestic share.²²¹ America remains the best place to invent. China may now be the best place to install. And in a diffusion-determined era, that may be the whole ballgame.

Taiwan and the semiconductor singularity

Now we need to place Taiwan at the center of the chapter, because any serious discussion of the Sino-U.S. AI Great Game that treats Taiwan as a side issue is deceiving itself. Taiwan isn't merely a diplomatic irritant, not merely a flashpoint, not merely a democracy of 23 million people living under permanent threat from the mainland, not merely a test of American credibility. *Taiwan is the semiconductor singularity*. It's where geopolitics, democracy, advanced manufacturing, deterrence, AI, and the physical substrate of the future all converge.

The numbers are astonishing enough that one has to reread them slowly. Taiwan accounts for more than 60% of global foundry revenue and more than 90% of leading-edge chip manufacturing, according to the U.S. International Trade Administration.²²² SIA has similarly described advanced semiconductor capacity as *catastrophically concentrated* in East Asia, with Taiwan holding the overwhelming share of the most advanced nodes. In plainer English: the world's AI future runs through a small island that Beijing regards as an *unfinished civil-war question* and Washington increasingly regards as a technological lifeline.

That's a dangerous amount of strategic dependency to place on one island, however extraordinary that island may be. And Taiwan *is* extraordinary. We should say that clearly. It's a vibrant democracy, a technological miracle, a model of industrial seriousness, and one of the most consequential societies of the twenty-first century. This isn't a case for abandonment. Quite the opposite. It is a case for making Taiwan *less uniquely hostageable*. Strategic redundancy isn't betrayal, or an abdication of our democratic ideals. It's simply smart deterrence.

This is where the so-called "silicon shield" becomes more complicated than the slogan allows. For years, people have argued that Taiwan's indispensability in semiconductors deters China because any invasion or blockade would imperil the global economy and damage Beijing too. There's truth in that, to be sure. But a shield can become a hostage situation if the entire world's advanced compute substrate depends on one geography. A world in which Taiwan is indispensable isn't necessarily safer than a world in which Taiwan is defended, allied, prosperous, and *less singularly indispensable*. The first version makes Taiwan the pressure point of civilization. The second gives Taiwan some breathing room.

This is *the* hot-button issue for Xi, and for China. Not tariffs. Not soybeans. Not even export controls, though those matter enormously. Taiwan is the legitimacy question, the sovereignty question, the national-reunification question, the place where Chinese nationalism, CCP historical narrative, U.S. credibility, and semiconductor dependence all meet. The May 2026 Beijing summit didn't dissolve

²²¹ International Federation of Robotics: "World Robotics 2025 report" published in September 2025 ([Link](#))

²²² International Trade Administration: "Semiconductors including chip design for AI" published in December 2025 ([Link](#))

that reality; instead, it clarified it. CFR's summit analysis emphasized Taiwan as one of the central variables affecting broader U.S.–China stability, and Reuters reported after the summit that Trump intended to speak with Taiwan's President Lai Ching-te, a move that would be historically significant and almost certain to inflame Beijing.

So here's the uncomfortable American position: we have to support Taiwan's democratic autonomy, avoid reckless signaling that turns ambiguity into provocation, and simultaneously reduce our advanced-chip dependence on Taiwan as quickly as humanly possible. Those three imperatives aren't contradictory. They're the only coherent strategy. A more diversified semiconductor base makes Taiwan less economically hostageable, makes China's coercive leverage less terrifying, and gives the United States more room for strategic patience rather than panic.

The moral language matters here too. Reducing dependence on Taiwan should never be framed as "moving on" from Taiwan. That would be both ugly and self-defeating. The right frame is *allied redundancy*: more capacity in Taiwan, more capacity in the United States, more capacity in Japan, more capacity in Europe, more packaging outside Taiwan, more memory resilience, more lithography ecosystem depth, more substrates, more chemicals, more specialty gases, more HBM, more CoWoS-like advanced packaging, more power, more talent. The goal shouldn't be some misguided vision of autarky. The goal is *multiple failure domains*.

Because right now, the failure domain is too concentrated. And the AI age makes that concentration intolerable.

Fabs are foreign policy

This brings us to fabs, which sound boring and technical until one realizes they're among the most important objects in the world. A fab isn't merely a factory. A leading-edge fab is *foreign policy in concrete*, steel, ultraviolet light, vibration control, clean rooms, water systems, process recipes, logistics, power contracts, talent pipelines, and national patience. It's one of the hardest things human beings know how to build. It's also one of the things America has to get much better at building, quickly.

TSMC's Arizona expansion is therefore hugely important, and we should understand both its promise and its *insufficiency*. In March 2025, TSMC announced that its total planned U.S. investment would rise to \$165 billion, adding \$100 billion of new investment to support three additional fabs, two advanced packaging facilities, and a major R&D center.²²³ TSMC's own Arizona materials now describe a plan for six semiconductor wafer fabs, two advanced packaging facilities, and an R&D team center on the Phoenix campus. That's a serious commitment. That's the beginning of a domestic advanced-manufacturing substrate we should have built years ago.

The federal support matters too. TSMC won a \$6.6 billion CHIPS Act subsidy in 2024 as part of its Arizona expansion, and its first Arizona fab has begun producing 4nm chips.²²⁴ The second Arizona fab is slated for more advanced technology, with 2nm production expected later in the decade. These details matter because the question isn't whether America has a fab press release. The question is whether America can *build an ecosystem* that produces leading-edge chips at competitive yield, cost, scale, and cadence.

And Arizona isn't the only node in the allied fab lattice. TSMC has expanded in Japan, including

²²³ TSMC: "TSMC Intends to Expand Its Investment in the United States to US\$165 Billion to Power the Future of AI" published in March 2025 (Link)

²²⁴ TSMC: "TSMC Arizona and U.S. Department of Commerce Announce up to US\$6.6 Billion in Proposed CHIPS Act Direct Funding, the Company Plans Third Leading-Edge Fab in Phoenix" published in April 2024 (Link)

Kumamoto, and has also launched ESMC in Dresden, Germany, a joint venture with Bosch, Infineon, and NXP to strengthen Europe's semiconductor ecosystem. Reuters has reported that TSMC is also planning advanced 3nm production in Japan, which would move Japan's role beyond mature or specialty nodes into something much more strategically significant.

But let's not congratulate ourselves prematurely. A fab announcement isn't a fab. A fab isn't a semiconductor ecosystem. A semiconductor ecosystem isn't only wafers. It's advanced packaging, HBM memory, substrates, chemicals, gases, masks, equipment service, technicians, construction labor, water, wastewater, transmission, power, environmental permitting, immigration, vocational training, and a thousand suppliers nobody notices until the one microscopic part is missing and the entire beautiful machine stops working. A fab without packaging is a half-answer. A fab without reliable power is a mausoleum. A fab without skilled workers is a *broken national aspiration*, incongruously wearing a bunny suit.

So yes, we need to muscle up. Fast. Not performatively. Not through ceremonial industrial policy where everyone stands behind a podium and intones the phrase *supply chain resilience* while the permitting timeline quietly eats the decade. *We need to make fabs a national-installation priority.* That means faster permitting, hardening water and power infrastructure around fab clusters, aggressive technician pipelines through community colleges, immigration pathways for semiconductor talent, domestic advanced packaging, materials redundancy, procurement certainty, and explicit coordination among TSMC, Intel, Samsung, Micron, NVIDIA, AMD, Apple, equipment suppliers, and the federal government.

The phrase "industrial policy" still makes some Americans uncomfortable, because we retain this quaint fantasy that markets alone will produce strategic resilience in sectors where the time horizon is longer than a quarterly earnings call and the capital requirements resemble a defense program. I think it's safe to say that fantasy is over. *The Chinese state understands fabs as sovereignty.* Taiwan understands fabs as survival. Japan understands fabs as strategic re-entry. Europe, in its slower and more encumbered regulatory way, understands fabs as industrial relevance. America needs to rediscover that fabs aren't merely companies' capex decisions. They are national operating assets, and we need a lot more of them.

And again: this isn't autarky. Autarky is a silly and unattainable idea. *The objective is resilient interdependence.* We should want Taiwan thriving, TSMC thriving, allied fabs thriving, and enough domestic capacity that a Taiwan crisis doesn't become immediate civilizational cardiac arrest. The goal isn't to replace Taiwan. The goal is to make the democratic semiconductor ecosystem robust enough that Taiwan's democracy isn't held hostage by the entire world's dependence on one island's fabs.

The May 2026 Beijing summit: strategic stability, reciprocity, and the unresolved hard things

Now let's place all of this in the immediate diplomatic context. The May 2026 Trump–Xi summit in Beijing was revealing, not because it solved the big problems, but because it showed *which* problems can be managed transactionally and which remain structurally dangerous. The White House fact sheet emphasized “a constructive relationship of strategic stability on the basis of fairness and reciprocity,” cooperation on Iran and North Korea, reopening the Strait of Hormuz, Xi's planned fall visit to Washington, and the creation of a U.S.–China Board of Trade and Board of Investment.²²⁵ The Board of Trade is meant to manage bilateral trade across non-sensitive goods; the Board of Investment is meant to create a government-to-government forum for investment-related issues.

Reuters reporting after the summit described a parallel track of negotiations around managed trade, investment, and AI protocols, with the administration aiming to evaluate bilateral investment deals while avoiding national-security concerns.²²⁶ That last clause is doing a lot of work. It points toward the right architecture: allow competition where competition is healthy, restrict access where national security genuinely requires it, and stop pretending that total decoupling is either possible or desirable.

This is where we need to be more adult than our political discourse usually allows. *We need Chinese companies to compete here in non-sensitive domains* under transparent, reciprocal, rules-based conditions. *And China needs to allow American companies to compete there under the same logic.* Reciprocity, not xenophobia. Competition with guardrails, not decoupling cosplay. Investment screening, yes. CFIUS, yes. Critical infrastructure controls, yes. No naïveté around data, chips, defense, AI models, biosecurity, telecom, or connected vehicles. But the answer can't be indiscriminate exclusion of Chinese firms from all American markets while demanding that China open every market to us. That's an unworkable asymmetry.

The summit's limits were just as revealing as its deliverables. CSIS noted that the summit produced *little visible progress* on the most consequential technology tensions: AI, cyber operations, export controls, and digital sovereignty.²²⁷ And Taiwan remained the deeper variable under the surface, the issue that Beijing links, explicitly or implicitly, to the broader health of the relationship.

That's why this moment is so delicate. We need economic engagement without strategic naïveté. We need reciprocity without xenophobia. We need Taiwan deterrence without Taiwan panic. We need fab redundancy without Taiwanese abandonment. We need export controls that buy time and an installation agenda that uses the time. And we need to avoid confusing performative toughness with statecraft.

The Beijing summit wasn't *détente*, or a strategic reset. *It was a signpost.* It showed that the two powers can still transact, still create forums, still produce language around stability, and still leave the core strategic contradictions largely unresolved. That's the world we now inhabit.

²²⁵ The White House: “Fact Sheet: President Donald J. Trump Secures Historic Deals with China, Delivering for American Workers, Farmers, and Industry” published on May 17, 2026 ([Link](#))

²²⁶ Reuters: “US, China are discussing AI guardrails to safeguard most powerful models, Bessent says” published in May 2026 ([Link](#))

²²⁷ Center for Strategic & International Studies: “What the Trump–Xi Summit Revealed, and Left Unsaid, About U.S.–China Tech Competition” published on May 20, 2026 ([Link](#))

The Hive Mind: open source, diffusion, and the Chinese advantage in implementation

There's another distinction here that matters enormously—perhaps none more than the ethos of open source and what I can only describe, with admiration, as the *Chinese hive mind*. This is one of the places where the comparative advantage may be shifting most rapidly, and where American observers are still, I think, dangerously under-theorizing what is happening. *American frontier labs are locked in a winner-take-all, secretive, quasi-theological race to hit AGI first*. They're operating, in some sense, as *rival monasteries of cognition*—each burning staggering amounts of capital and talent to reach roughly similar thresholds, with occasional leapfrogging but only ephemeral lead times. OpenAI surges ahead, Anthropic closes the gap, Gemini reasserts itself, xAI reconstructs Grok and promises parity again by year-end. The whole thing has the feel of *convergent evolution* inside a highly capitalized priesthood: duplicative, heroic, secretive, and expensive.

China, by contrast, is beginning to look less like a cluster of rival monasteries and more like a *hive*—a distributed, protean, state-supported, brutally competitive swarm of engineers. That difference matters. It matters a great deal. Because the American model still assumes that the strategic prize lies in *monopolizing the frontier breakthrough*: hit AGI first, lock in the network effects, figure out commercialization later, and sit atop the commanding heights. The Chinese ecosystem appears much less starry-eyed about AGI as a metaphysical event and much more pragmatic—almost dour, really—about *implementation in the real economy*. One company pushes forward on one algorithmic front, another iterates, another operationalizes, another takes the baton, and the whole ecosystem metabolizes the advance with astonishing speed. It's less a cathedral than a nervous system.

That's the heart of the hive-mind argument. The U.S. still leads, yes, in frontier-model production and discontinuous scientific breakthroughs. But the lead is increasingly perishable because China's model of open, gladiatorial, engineering-heavy competition is *structurally better suited to the diffusion phase*. A breakthrough no longer has to remain imprisoned inside the walls of one company. DeepSeek can sprint out for a minute, pass the baton to Manus, and both are surpassed by Kimi. They all absorb, rework, disseminate, and operationalize each other's advances. And because Chinese firms appear less inhibited by the Western sacramental attitude toward intellectual property—and because open-source norms now function as an international substrate rather than an American eccentricity—the whole ecosystem can move in something closer to parallel. One company specializes in one advancement. Another in another. The sum total begins to look less like a market and more like a coordinated swarm, even if no one formally coordinates it.

This is also why the open-source question is not some niche argument among software romantics. It's a civilizational competition question. American frontier labs still tend to think in terms of *breakthrough, monopoly, lock-in*. Winner take all. China increasingly appears to think in terms of *diffusion, iteration, operationalization*. The former model is excellent for invention. The latter may be superior for installation. And if you believe, as I do, that the installation phase is what determines who captures the durable geopolitical winnings of a technological revolution, then the hive-mind structure becomes a very big deal indeed.

There's a darker edge to this too, and we shouldn't sentimentalize it. The hive isn't merely collaborative. It's *gladiatorial*. It's Darwinian, cutthroat, protean, and relentless. This isn't some kumbaya open-science utopia. It's a brutal pugilistic ecosystem in which advances propagate quickly, imitation carries little stigma, and the pressure to operationalize is ferocious. But brutality, however aesthetically unappealing, is often efficient. And efficiency matters in the installation phase.

This is one reason I think American observers can be too complacent when they hear that the U.S. still leads in notable frontier models. Fine. That may remain true for some time. But the more relevant question isn't merely who gets to the frontier first. It's who can take what is discovered there and disseminate it horizontally across an entire economy faster. China's hive mind—its engineering pres-tidigitation, its open-source comfort, its vast swarm of mathematically gifted talent, its lesser reverence for IP formalism, and its state-sponsored monomania on implementation—may prove to be structurally superior at exactly that.

To put it more bluntly: *America may still be the best place to invent the future. China may be becoming the best place to spread it.*

And in this age of implementation, spreading may matter more.

Engineering society, lawyerly society

There is, I think, a deeper civilizational distinction here that we ignore at our peril. Dan Wang's formulation bears repeating because it names something real and uncomfortable: *China as an engineering society, America as a lawyerly society*. That's obviously an oversimplification—every national stereotype is—but like many oversimplifications, it captures a structural truth. China's elite political class has long exhibited a strong *technocratic* bent, and the state continues to reward the buildout of physical systems—grids, ports, rail, fabs, nuclear plants, industrial parks, robotized factories, entire cityscapes—in a way that now feels almost alien to the contemporary West.

The United States, by contrast, has become expert at procedure, litigation, consultation, compliance, review, and veto points. That's not always a deficiency. The rule of law, local consent, environmental review, due process, and the right to object aren't pathologies. They're great societal, political and cultural achievements. But in an installation-heavy era, they are also friction. And friction, however noble its provenance, is still friction.

You can see the distinction not merely in abstract political culture, but in the physical substrate of AI itself. China is building the high-three-terawatt electricity base. China is installing industrial robots at a level no other country comes close to matching. China is expanding nuclear, renewables, transmission, and manufacturing automation with a cadence the U.S. struggles to match. That's not just an energy story or a robotics story. It's obviously *the AI story*. It's a story about whether a country still retains the civilizational muscle memory of building in the world of *atoms* rather than merely talking about it in the world of *bits*.

Xi himself often contrasts the "fictitious economy" of finance, speculation, platform business, gaming, social media, and parts of e-commerce with the "real economy" of manufacturing, infrastructure, energy, logistics, semiconductors, and robotics. And he's not wrong, at least descriptively. The engineering society builds. The lawyerly society consults, delays, sues, revises, convenes stakeholder processes, and eventually—maybe—permits. In normal times that can look like prudence, caution, even moral seriousness. In the installation phase of a general-purpose technology, it can look a lot like strategic self-sabotage.

And this is where the contrast becomes so disquieting, because the same national traits that make America an inventive, pluralistic, rights-protecting civilization may also render it slower, more decelerative, and less coherent when the challenge isn't invention but horizontal deployment at national scale. China, for all the obvious moral compromises and coercions embedded in its system, still appears to know how to move steel, concrete, electrons, capital, and administrative will in the same direction. That's not a minor advantage in an age where compute is physical, intelligence is

infrastructural, and the true bottleneck lies not merely in having the idea, but in *industrializing* it.

America innovates, Europe regulates, China accelerates

Last year I used the aphorism—at least I thought it was clever—America *innovates*, Europe *regulates*, China *appropriates*. It worked well enough then, but it needs revision now, because “appropriates” has become too dismissive, too self-soothing, and, if I’m being honest, too flattering to our own priors. China is no longer merely appropriating. It’s increasingly demonstrating *indigenous ingenuity* and then installing at scale. That’s a more serious claim, and a more unsettling one. So maybe I’ll go with *accelerates*, though *installs* is more accurate (but offensive to my rhyming scheme). We’ll see.

Europe, for its part, has doubled down on its comparative advantage in regulation. Wow—incredible to watch the self-inflicted irrelevancy and obsolescence they’ve consigned themselves to. The EU AI Act is the cleanest expression yet of the European instinct to meet a technological frontier first with legal taxonomy, risk classes, obligations, and process constraints. Whether you regard that as enlightened restraint or self-immolation depends largely on whether you think the principal danger of AI is misbehavior or irrelevance. There’s a reason so many founders quietly route their ambition elsewhere. Europe continues, almost proudly, to luxuriate in its own inconsequence to the frontier. It still wishes to shape the future without being especially willing to build it. And so the old line still mostly holds on that score: Europe regulates because it has, in many respects, abdicated invention and surrendered installation. Good luck with that.

The United States remains dominant in notable frontier-model production, and for that we should be grateful rather than blasé. Silicon Valley is still the closest thing the species has to a machine for manufacturing Promethean breakthroughs. But even inside the U.S., the lawyerly-society pathology is on display in real time. Export controls oscillate. Regulatory frameworks mutate. National-security agencies and frontier labs fight each other in public. Congressional hearings alternate between sanctimony and incomprehension. Supply-chain risk designations get waved around too broadly in some places and too timidly in others. We’re trying to prosecute the most important technological mobilization in modern history through a political culture optimized for procedural dispute and adversarial fragmentation.

China, by contrast, doesn’t litigate its way into diffusion. It decrees. That’s its advantage.

And if you believe—as I do—that diffusion determines the civilizational winnings of an industrial revolution, then the structural risk for America becomes obvious. It’s not enough to remain the world’s premier inventor if another power proves more adept at taking what has been invented and spreading it faster, more comprehensively, and with fewer scruples across the real economy. Which is why the aphorism now needs to be harsher, cleaner, and more accurate: *America innovates, Europe regulates, China installs.*

That, unfortunately, is beginning to look less like rhetoric and more like diagnosis.

Age of discovery vs. age of implementation

Kai-Fu Lee's distinction between an *age of discovery* and an *age of implementation* is useful here because it forces a clarification that too many people elide.²²⁸ The United States is still advantaged in the former. It remains the best place on earth to invent a new category, birth a new lab, assemble risk capital, concentrate talent density, and turn computational ambition and manic Promethean confidence into a frontier scientific enterprise. If you want the next great discontinuity—the leap, the thunderclap, the strange and marvelous breakthrough that suddenly redraws the map—you still go to America, and more specifically to Silicon Valley.

But China may be advantaged in the latter. *And the latter may matter more than most Americans want to admit.* Diffusion matters more than invention for durable geopolitical hegemony. That's been one of the central declaratives of this essay from the start. The country that does the horizontal installation fastest and most comprehensively across the real economy captures the compounding advantages: productivity gain, GDP expansion, labor reorganization, downstream military leverage, political influence, cultural spillover, and institutional allegiance. Again, returning to our historical lens: that was true in *mechanization*. It was true in *electrification*. It was true in *computerization*. It may be even more true in *agentification*, where dissemination costs are lower, applications are more plastic, and the competitive half-life of a lead is measured increasingly in months, not decades.

This is why the old American consolation—yes, *but we invented it*—may prove historically inadequate. There's a kind of narcissism embedded in invention worship. We celebratize the founder, the inventor, the lab breakthrough, the Nobel-adjacent moment of intellectual prestidigitation. But history is less sentimental. History tends to reward the societies that can turn invention into infrastructure. The age of discovery is glorious. *The age of implementation is where the real geopolitical spoils are often allocated.*

So yes, last year it may still have been fair to say, "America innovates, Europe regulates, China appropriates." This year that's too flattering to us and too condescending to them. China is no longer merely copying. It's showing increasing indigenous ingenuity and then industrializing that ingenuity through deployment. It's not just catching up in the laboratory. It's learning how to metabolize breakthroughs into the real economy more rapidly than the West.

That's what makes this phase different, and more dangerous.

The Thucydides Trap, and why this moment is more dangerous than it looks

There's one more layer we need to add here, because otherwise this chapter risks sounding merely like an industrial-policy comparison rather than what it actually is: a warning about grand-strategic danger. The obvious historical frame is the *Thucydides Trap*—the now-famous formulation popularized by Graham Allison,²²⁹ drawing on Thucydides' account that "it was the rise of Athens and the fear that this instilled in Sparta that made war inevitable." Allison's tally, however contested, remains sticky because it's so memorably grim: *12 of the last 16 historical cases in which a rising power threatened to displace a ruling power ended in war.* Whether one takes that as precise social science or as a deliberately percussive heuristic is almost beside the point. The underlying lesson remains profoundly relevant: when the relative power of two great states shifts quickly, fear, prestige anxiety, misperception, overreaction, and accidental escalation can become systemically dangerous.

²²⁸ New York Times: "What China Can Teach the U.S. About Artificial Intelligence" published in September 2018 (Link)

²²⁹ The Atlantic: "The Thucydides Trap" published in September 2015 (Link)

That's what makes this chapter more than a meditation on compute clusters, power grids, chip sanctions, Taiwan, and robot density. The AI race is nested inside a larger hegemonic transition. If China is the rising power and the United States the incumbent hegemon, then every argument in this chapter about chips, data, power, installation, and "free doctors" (I'll get to presently) exists inside a broader psychological and geopolitical field shot through with insecurity. That means the danger isn't merely that one side wins economically or technologically. It's that each side begins to interpret the other's moves—export controls, model releases, rare-earth restrictions, semiconductor sanctions, AI-health deployments in the Global South, industrial subsidies, naval maneuvers, investment bans—not as normal competition, but as encirclement, humiliation, or pre-hostile positioning.

That's how traps work. Not because either side wants war in some melodramatic, bellicose sense, but because both sides begin to behave in ways that make otherwise containable moves feel existential.

Which is precisely why we need to be careful not to stumble into it. Strategic realism doesn't require chest-thumping. It requires restraint, lucidity, and an aversion to self-dramatizing escalation. The May 2026 Beijing summit mattered because the United States and China need something more difficult than détente and more realistic than friendship. They need a *modus vivendi* in which ruthless competition in technology, trade, manufacturing, ideology, and military posture doesn't automatically harden into fatalism. The summit's language around "constructive strategic stability" may sound like diplomatic fog, and much of it probably is, but fog is sometimes better than flame.

If this really is a Sino-U.S. world, then avoiding the Thucydides Trap becomes not some seminar-room conceit but one of the defining statecraft tasks of our era. We shouldn't be Pollyannaish about that risk. But nor should we become melodramatic or apocalyptic about it. The correct response is neither naïveté nor jingoism. It's a calm, dispassionate sobriety.

We need to call balls and strikes where China has the advantage, learn from the parts of its installation machine that are in fact superior, harden our own diffusion capacity, build the substrate we need, and still remain fanatically committed to not converting every asymmetry into a prophecy of war. That's the tightrope here. And it's a narrow one. Because the more honest we are about China's strengths, the more temptation there will be in some circles to convert realism into bellicosity. That would be a category error of the first order. The patriotic response isn't to flatter ourselves or demonize them. It's to understand the competition clearly enough that we can respond intelligently—without letting fear do the strategic thinking for us.

Healthcare as the decisive theater: the coming "free doctor"

Why does any of this belong in a healthcare essay? Fair question.

Because healthcare is where AI becomes legible to ordinary human beings as value. You can debate benchmarks and parameter counts all night long. You can argue about evals, context windows, tool use, embodiment, hallucination, agentic reliability, chain-of-thought, and the metaphysics of machine reasoning until everyone in the room begins quietly hoping for the asteroid. But when a farmer in rural Kenya gets a diagnostically competent, multilingual medical agent on a cheap smartphone, AI stops being abstract technology and becomes *lived social infrastructure*.

That is why I need the broader geopolitical commentary before returning to healthcare. The "free doctor" is the first AI use case that a child, a parent, a finance minister, a village elder, and a head of state can all understand without a benchmark chart. It doesn't require a theory of transformer scaling. It requires a sick person and a competent answer. Once medical intelligence becomes ubiquitous, it becomes the human face of the AI stack: the place where chips, electricity, data centers,

models, language, trust, privacy, and governance are all suddenly experienced as care.

China already has a geopolitical template for this kind of distribution: *the Belt and Road Initiative's* digital and health dimensions—*the Digital Silk Road and the Health Silk Road*—explicitly treat infrastructure and public goods as foreign-policy instruments.²³⁰ The United States, of course, also understands health as soft power; PEPFAR remains one of the most consequential health-diplomacy initiatives of the modern era. But AI changes the unit economics. The “free doctor” isn’t a building. It’s a humanistic ‘gift’ and a distribution. Once the models are competent enough, marginal cost collapses, and deployment becomes a question of bandwidth, devices, localization, trust, and geopolitical alignment.

Here’s the uncomfortable hypothetical: the U.S. may build the medical superintelligence because we still dominate the frontier labs and the core research ecosystem. But America’s healthcare system is also uniquely encumbered by professional guilds, malpractice anxiety, HIPAA maximalism, institutional fragmentation, state-by-state medical practice laws, and regulatory inertia. Meanwhile, China’s structural posture—centralized data, state-led mobilization, platform integration, and cultural tolerance for rapid deployment—could allow it to disseminate the “free doctor” faster internationally, bundling it with infrastructure, cloud partnerships, financing, phones, digital identity, payments, and governance defaults.

This isn’t speculation in the abstract. It’s the statecraft logic we’ve seen time and again from China, that of the *Digital Silk Road applied to cognitive infrastructure*.

And it won’t just be China. The race to operationalize AI in governance and services is *already* global, with sovereign AI becoming the new strategic fashion from Europe to the Gulf. In other words, America isn’t merely competing with China; it’s competing with any polity that can install AI into the real economy faster than we can. Then we’re reduced to spectating, and eventually to *reverse-importing* the very inventions we pioneered. I think this is very likely to happen.

The “free doctor” will matter because healthcare is universally intelligible. A port is useful. A rail line is useful. A cloud region is useful. But a doctor in your pocket, in your language, at 2 a.m., for free or nearly free? That’s soft power at a scale we’ve barely begun to imagine. The country or company that delivers it will gain more than goodwill. It will gain data, trust, dependence, clinical influence, regulatory templates, and maybe even ideological affinity.

The “free doctor” won’t merely answer medical questions. It will encode assumptions about privacy, autonomy, triage, state authority, family, risk, acceptable treatment, mental health, reproduction, end-of-life care, disability, and whose knowledge counts. There is no neutral doctor. Human doctors aren’t neutral either. Every system of care embodies a moral anthropology. The AI doctor will too.

So the question becomes: whose anthropology gets exported?

²³⁰ Council on Foreign Relations: “Assessing China’s Digital Silk Road Initiative” ([Link](#))

Reverse innovation

The most unsettling possibility here isn't that America loses the invention race. It's that America *wins* the invention race and then *loses* the distribution race. The first true medical superintelligence—if by that we mean a system capable of materially superhuman differential diagnosis, exhaustive biomedical synthesis, longitudinal reasoning over the medical record, multimodal interpretation, and clinically useful guidance—may very well be born in the United States. That would be the unsurprising part. The surprising part would be if the United States then spent years litigating, regulating, guild-protecting, and morally grandstanding its way into stasis while a more installation-forward state harnessed the same breakthrough and deployed it at societal scale.

Autonomous driving is the cleaner analog. The technical gap between machine driving and human driving may become obvious long before the regulatory and liability layer allows universal adoption. Healthcare could follow that exact pattern. America invents; China deploys; and America then “reverse innovates” by importing back, years later, the thing it originally birthed.

This *wouldn't* be unprecedented. The history of technology is full of inventors who fail to capture the full value of their own inventions because someone else installs better. The British pioneered much of the first Industrial Revolution, but the United States learned to scale mass production at continental scale. Xerox PARC invented many of the conceptual primitives of personal computing; Apple and Microsoft captured the operating reality. America helped birth the internet; China built an entire platform-state architecture around digital payments, super-apps, surveillance, logistics, and e-commerce. Invention is glorious. Installation captures the cash flows and the institutional consequences.

That's what I fear in medical AI. The American frontier labs may *create* the thing. The American healthcare system may *slow-walk* the thing. Another country may *normalize* the thing. And then we will have the uniquely American experience of reimporting our own future after it has been productized elsewhere with different values.

Less an ascendancy than a return

One more thing is worth preserving because it gives this argument historical depth rather than merely policy heat: China's rise is less an ascendancy than a *return*. *China was the world's largest economy on the eve of the first Industrial Revolution*. Its relative decline wasn't some permanent metaphysical condition; *it was the product of a historical failure to industrialize at the right moment*, at the right speed, with the right institutional flexibility. Xi's China appears determined not to repeat that mistake.

China is now mobilizing capital, talent, infrastructure, and policy around AI with extraordinary intensity. This isn't merely commercial competition for them. *It's an attempt not to lose modernity a second time*. That phrase matters. America often treats AI as a market, a software category, a national-security concern, and a labor shock. China treats it as civilizational positioning. The difference in emotional register matters because it shapes state behavior. A nation that believes it missed one industrial revolution will not demurely risk missing another.

That's why the AlphaGo humiliation mattered. That's why the 2017 plan mattered. That's why the fab race matters. That's why Taiwan matters. That's why power matters. That's why robotics matter. That's why the May 2026 summit matters. The Chinese system may be coercive, brittle, surveillant, and morally divergent in ways we shouldn't minimize. *But it is also awake*. And awake systems are

dangerous when sleepy systems keep telling themselves they invented the alarm clock.

The remediation agenda: how a lawyerly democracy *relearns* how to install

So what do we actually do with all of this?

Because diagnosis without prescription is just intellectual tourism—and again, channeling Ian Sacks, I gotta land some planes here. If everything I've just said is even directionally correct—if this really is an installation game, if diffusion matters more than invention, if the Chinese system is structurally advantaged in that phase—then the uncomfortable implication is that *the United States has to relearn something it used to know how to do extraordinarily well: build, mobilize, install at scale, and do it without apologizing for itself every step of the way.*

The first step is simply to name the problem honestly, without euphemism and without ideological reflex. We are still the best invention engine in the world. That's not in question. But invention is no longer sufficient. Installation has become a first-order national competency—arguably the competency—and right now we are underweight in it. That's not a moral failing; it's a structural reality, and a very consequential liability. We've layered our system with veto points, process, litigation, stakeholder balancing, risk aversion, and procedural self-admiration to such a degree that velocity itself has become suspect. In normal times, that looks like prudence. In an installation phase, it starts to look like strategic drag.

Second, we need to stop pretending export controls are a strategy rather than a tactic. They are, at best, a delay mechanism. And delay is only valuable if you use the time to build something behind it. Otherwise, you are just watching a clock you don't control. Slowing China's access to the very frontier of compute may buy quarters, maybe a couple of years at the margin. But it doesn't change the underlying physics of diffusion, nor does it neutralize a system that's perfectly capable of working with slightly degraded inputs if it can install faster at the system level. You don't win a marathon by tying your competitor's shoelaces together if you never train.

Third, we need to treat energy and physical infrastructure as core components of AI strategy, not as adjacent policy domains to be debated in separate committees. Compute isn't an abstraction. It's electrons moving through silicon at scale. If AI is a *multiplication of intelligence*, then power is the substrate of that multiplication. A society that cannot build power cannot scale intelligence. And a society that cannot scale intelligence won't win an installation race.

Fourth, we need to make domestic and allied semiconductor redundancy a national obsession. Again, TSMC Arizona is a huge and important start: \$165 billion, six fabs, two advanced packaging facilities, and an R&D center isn't trivial. But a start isn't synonymous with strategy. The United States needs the full stack: leading-edge logic, mature nodes, memory, advanced packaging, substrates, specialty gases, chemicals, lithography service, metrology, EDA, fabless design, construction labor, technicians, and power. We need Arizona, but not only Arizona. We need Oregon, Ohio, Texas, New York, New Mexico, Japan, Germany, Korea, the Netherlands, and Taiwan itself as part of a resilient democratic and allied semiconductor lattice. The goal isn't to duplicate the entire Taiwanese miracle overnight. The goal is to ensure that no single geography remains a civilization-scale point of failure.

Fifth, we need a more adult conversation about openness. I'm certainly not advocating for shutting down open source—there are profound reasons, both moral and practical, to keep these systems legible, inspectable, and diffusible within democratic societies. But we should at least be honest about the trade-off. Radical openness compresses global catch-up time. It doesn't selectively advantage friendly actors. *It advantages everyone*, including those explicitly competing with us at the level

of statecraft. The American instinct to universalize innovation is admirable. It's also, in this context, strategically naïve if left completely unconstrained. We need to find a middle path between priesthood secrecy and indiscriminate diffusion—a way to preserve dynamism without handing away the full blueprint in real time.

Sixth, we need *reciprocity* rather than *xenophobia* in economic policy. The May 2026 summit's trade and investment boards point in a useful direction, provided they don't become ceremonial furniture. Chinese companies should be able to compete in the United States where national-security risk is manageable, transparent, and bounded. American companies should be able to compete in China under comparable conditions. That's the only sustainable posture: open where possible, restricted where necessary, reciprocal always. The alternative is a kind of sloppy decoupling theology that confuses moral clarity with operational incoherence.

Seventh—and this is where the essay loops back into healthcare—we need to treat AI deployment in critical domains as both moral obligation and geopolitical instrument. If the “free doctor” is coming, and it is, then the question isn't whether it will exist. The question is *whose values will it carry* when it arrives. Does it embody transparency, patient autonomy, auditability, consent, and a bias toward individual dignity? Or does it embed different defaults—more opaque, more centralized, more state-mediated? Those design choices won't be neutral. They will propagate. They will become the unseen scaffolding of how billions of people interact with systems of care.

That means the United States needs to get serious—quickly—about exporting trustworthy AI infrastructure, particularly in healthcare. Not as charity. Not as afterthought. But as strategy. The same way we once exported pharmaceuticals, public-health systems, clinical standards, and hospital-management know-how, we now need to think about *exporting cognitive infrastructure*—aligned, safe, legible, and scalable. That requires public-private coordination that we haven't historically been great at, but that we are going to have to rediscover.

Eighth—and this is the part that is hardest to say out loud—we *need to recover a cultural tolerance for building*. We also need to increase our risk tolerance associated with deployment. No *more holding technology to the inhuman standard of infallibility*; let's settle for the more reasonable standard of human equivalency, and then human superiority. We need to celebrate not only the founder who *invents* something, but the operator who *installs* it, the engineer who *scales* it, the technician who *maintains* it, the organization that *deploys* it across thousands of nodes in the real economy. We've spent decades valorizing disruption and novelty *while quietly devaluing execution, infrastructure, and industrial capacity*. That imbalance now shows up as a strategic vulnerability.

The countries that win this phase won't be the ones that merely invent clever things. They will be the ones that can lay fiber, pour concrete, permit data centers, build fabs, train technicians, deploy systems, wire substations, scale robots, and do it repeatedly, relentlessly, and without getting lost in their own process.

And finally, we need to resist the temptation—always present in American discourse—to turn this into a morality play. This isn't about demonizing China or romanticizing ourselves. It's about understanding how systems behave under pressure. China has advantages in installation. We have advantages in invention. Neither is absolute. Both are mutable. The question is whether we are willing to adapt before the window closes.

Because that window is real. And it's narrowing.

The Chapter, Compressed

Let me close by compressing the argument.

First, authoritarian systems often have an installation advantage. That doesn't make them morally preferable. It makes them operationally faster in certain phases of technological diffusion. Democracies need to recognize the disadvantage without surrendering the virtues that make them worth defending.

Second, China woke up to AI earlier than most Americans understand. AlphaGo wasn't a parlor trick in Beijing's imagination. It was a **Sputnik moment**, a symbolic humiliation, and a national-mobilization trigger. The 2017 AI plan wasn't ornamental. It was the policy wake behind the boat.

Third, the Sino-U.S. AI competition decomposes into five arenas: **compute, data, algorithms, power, and culture**. The U.S. still leads in frontier compute and model invention, but China has powerful advantages in data aggregation, power buildout, manufacturing installation, robotics diffusion, and engineering-state execution.

Fourth, Taiwan is the semiconductor singularity. It's morally a democracy deserving defense, strategically a technological miracle, and structurally a civilization-scale point of failure. Reducing dependence on Taiwan isn't abandonment. It's deterrence through redundancy.

Fifth, fabs are foreign policy. TSMC Arizona is necessary and important, but not sufficient. America needs the full semiconductor ecosystem: leading-edge fabs, advanced packaging, memory, materials, power, water, talent, and allied redundancy. A fab without ecosystem is an expensive shrine.

Sixth, the May 2026 Trump–Xi summit showed both the possibility and the limits of transactional stability. Trade and investment boards, AI dialogue, and practical economic commitments matter. But Taiwan, semiconductors, cyber, export controls, and digital sovereignty remain the hard problems. Reciprocity is the right posture: compete here, compete there, with national-security boundaries that are real rather than performative.

Seventh, China's *hive-mind* engineering culture may be structurally advantaged in the diffusion phase. America remains the best place to invent the future but China may be becoming the best place to spread it. And in the age of implementation, spreading may matter more.

Eighth, healthcare is a decisive theater of the Great Game. The first truly competent "free doctor" will be more than a healthcare product. It will be soft power, governance architecture, data infrastructure, and ideological export. If America builds the medical superintelligence but cannot deploy it because of guilds, litigation, and regulatory sclerosis, someone else will deploy it first—and we may end up reverse-importing our own invention.

Ninth, avoiding the *Thucydides Trap* requires realism without bellicosity. We need to understand China's strengths without turning every asymmetry into a prophecy of war. Strategic lucidity isn't jingoism. It's just good policy, and the sensible alternative to panic.

And tenth, the remediation agenda is *installation*. We need to build power, fabs, transmission, data centers, semiconductor ecosystems, allied industrial capacity, healthcare AI deployment pathways, and a public-private installation machine worthy of the invention engine we already possess.

The uncomfortable truth—the one that sits underneath this entire chapter—is that history doesn't reward the most virtuous system. It rewards the system that can translate capability into reality at scale. If we want that system to be ours—or at least one that reflects our values—then we need to

get much better, much faster, at the part of the cycle we have neglected.

We need to relearn how to install.

America has to remain open enough to invent, serious enough to build, confident enough to compete, prudent enough not to stumble into war, and decent enough not to become what it fears.

That's the Great Game. And it has already begun.

BEFORE WE TURN THE PAGE

The Great Game tells us that medical intelligence will travel through power, chips, fabs, electricity, diplomacy, and infrastructure. But power is never only geopolitical; it is also spiritual. The next chapter turns from the outer contest between nations to the inner question of reverence, deification, and the strongest thing in the world.

Chapter 10 · AI as Theology

The Strongest Thing in the World

“Wherever the people do not believe in something beyond the world, they will worship the world. But, above all, they will worship the strongest thing in the world.”

— G.K. CHESTERTON, CHRISTENDOM IN DUBLIN, 1932

A WORD ON NAVIGATING THIS CHAPTER

This chapter turns inward. After science, clinical medicine, labor, institutions, payers, behavioral health, and geopolitics, the question becomes spiritual: what happens when the strongest thing in the room is no longer human, and when usefulness begins to shade into reverence? I now include a brief stop with Pope Leo XIV’s Magnifica Humanitas, because that encyclical gives the chapter’s renewed-consecration-to-the-human argument its Catholic-social-teaching spine.

The Strongest Thing in the World

From Anthropomorphizing to Deifying

The Worship of Power in an Age of Silicon

I start this chapter with a rather melancholy, darkly ominous reflection from the always-quotable Chesterton, as a successor thought to my cognitive-offloading ruminations from the *Generative Epistemology* chapter. I warned you there that I was going to get philosophical and metaphysical, though probably not theological. Well, I changed my mind and am now diving headfirst—headlong?—into the religious and eschatological, the end-of-days stuff. So my advice to my weary readership, whose patience may by now be thinning, is to skip this chapter as well if you’re so inclined.

But before you reflexively do, let me at least explain *why* this chapter belongs here. On the surface, it may feel like a departure from the rest of the essay—a detour away from healthcare, institutions, technology, and the practical machinery of implementation. It is, in one sense. But in another, I think it sits much closer to the heart of the whole argument than may be obvious at first glance. Because if the rest of this essay is about what generative AI will do to our institutions, our labor markets, our medicine, and our civilizational tempo, this chapter is about what it may do to our *inner architecture*: what it awakens, displaces, substitutes for, and perhaps invites us to *worship*.

And that isn’t irrelevant to healthcare. After all, what is more sacred, more *sacramental* in our society, besides saving our souls, than the health of our bodies and minds? And if we have, in effect, consecrated our lives to the preservation, extension, and repair of the body and the mind, is it really so strange to ask, in the same breath, about the fate of the soul—or the soul of a civilization? That, to me, is why the Healthcare 150 and the AI 10 should care. So yes, this is a bit of a departure from the rest of the essay. But it’s also my attempt to get something out of my contemplations and onto the page that I increasingly suspect sits beneath much of the rest: the possibility that our relationship to

AI isn't merely technical, economic, or political, but spiritual. If the rest of this essay is about the *external* consequences of AI, this chapter is about its *interior* and *spiritual* ones.

What Chesterton saw, first in *Orthodoxy* in 1908 and then with sharper political clarity in *Christendom in Dublin* in 1932, isn't merely a theological observation but a structural, almost deterministic law of human behavior. The impulse to worship isn't optional; it's *constitutive*—an essential, ineradicable part of the human condition. You can redirect it, deform it, secularize it (Communist societies know this playbook by heart), but you can't extinguish it. And so when a society tramples the transcendent—when it no longer believes in anything beyond itself—it doesn't automatically become rational, neutral, or disenchanted in the way Enlightenment optimists once imagined. It becomes, in a very precise sense, more *primitive*. And with that de-evolution, more suggestible, more malleable, and yes, more *controllable*.

The vacuum gets filled by whatever presents itself as ultimate, and in practice that is almost always *power*—some combination of the state, ideology, or the dominant organizing force of the age. Chesterton's warning is that the move away from God isn't a move toward nothing, but toward something far more dangerous: the sanctification of the proximate, the contingent, the strongest thing in the room. And once power itself becomes the object of quasi-religious devotion, it's no longer constrained by anything higher than itself. That is how politics quietly metamorphoses into theology, and authority into something much closer to worship.

And what is the strongest thing in the room? With increasing insistence, it is this new, non-biological intelligence we have *speciated*. As AI evolves—with epochs measured at the speed of silicon, not the geologic pace of biological evolution—from passively awaiting commands to autonomously *acting* in the world, that power moves from the realm of cognition to the realm of agency, from bits to atoms. And if the very faculty that gave us dominion over the earth and its inhabitants—our intelligence—is suddenly superseded, *and* we have set God aside, or at least forgotten Him, then our ancient propensity to worship the strongest thing in the room may mean that we *worship the AI*.

From self-subordination of one's cognition, to self-subordination of one's agency, to the deification of and submission to a "higher" being sounds like dark science fiction. But is the temptation to confer divinity on the AI really so unthinkable? To move from *anthropomorphizing* it with human characteristics to *deifying* it with godlike ones? I'm suggesting that the leap may be shorter than we would like to admit.

If this is starting to feel metaphysical and quasi-religious, it is. We have *millenarians* and *singularitarians*. To my healthcare leaders who have real jobs and can't follow all this silliness on X: take a second and look it up. It is, in its own way, kind of fascinating. We have e/accs and decels. Accelerationists and doomers. Alignment priests, safetyists, existential-risk prophets, techno-utopians, rationalist monasteries, apostates, heretics, schismatics, true believers. The whole discourse is shot through with sectarian energy and theological undertones, even when it's pretending to be dispassionate, empirical and purely technical. Freud's *narcissism of small differences* is everywhere here: tiny distinctions in doctrine elevated into civil-war-level disputes, each faction convinced it alone has apprehended the true path to salvation or ruin. The arguments have the texture of interdenominational religious warfare—less like ordinary policy disagreement than like rival eschatologies contending for the soul of the future.

And, of course, along with this new religiosity comes a new priesthood. Demis, Dario, Sam, Mira, Jensen—high clerics of the new regime. They don't wear vestments, exactly, though perhaps Jensen's ubiquitous leather jackets or Dario's avuncular sweaters are the 2026 equivalent. But they

do increasingly speak in a kind of *sacramental vocabulary*: alignment, safety, scaling, emergence, takeoff, superintelligence, agency, control. The diction is technical on its surface, but increasingly liturgical underneath. These aren't merely executives discussing products. They are interpreters of destiny, exegetes of the near future, custodians of a power that feels less and less like ordinary technology and more and more like a civilizational force demanding doctrine, ritual, and belief.

And the vocabulary gives the game away. We speak of "god models" for a reason. We speak of alignment as if we're talking about moral formation, of safety as if we're constructing a theology of restraint, of the singularity as a secularized apocalypse, of AGI as a kind of second coming, of doom as revelation, of takeoff as rapture. We invoke an eternal present of scaling laws, test-time compute and benchmark saturation on one side, and an incommensurable future state on the other—beyond human comprehension, beyond ordinary politics, beyond the old categories. A *transhumanist* epoch. Even the language of training, reward, supervision, reinforcement, correction, and control starts to sound faintly catechetical if you stare at it long enough.

And then there's the *social* texture of the thing: the infighting and internecine warfare, the missionary zeal, the apocalyptic rhetoric, the millenarian tomes (including, perhaps, this one!), the open letters warning of human extinction, the "if anyone builds it, everyone dies" maximalism (riveting and important book by Yudkowsky, by the way),²³¹ the hunger strikes in front of DeepMind, Anthropic, OpenAI—hunger strikes! There's an eschatology here. There's even a *missionary structure* of belief and conversion. Which is why I keep saying that if this is starting to feel metaphysical, well, it is. The debates around AI don't merely resemble policy arguments or business disagreements. They increasingly resemble doctrinal schisms inside a new religion.

The narrative arc, meanwhile, is almost always one of transcendence, redemption, and salvation: curing cancer, adding ten to twenty years of longevity, discovering new science, reversing entropy, lifting humanity out of toil, out of scarcity, out of biological frailty. *Sand to silicon to intelligence to superintelligence*. It is a salvation narrative, however secularized. No matter how modern we imagine ourselves to be, no matter how disenchanted and post-religious we claim to be, we're still looking for transcendence, redemption, salvation. We haven't abolished the religious instinct. We've merely updated its vocabulary to that of 2026 San Francisco.

The Religious Architecture of AI

In that sense, AI has many of the properties and characteristics of religion. There is something a little like Rabbinic Judaism in the tradition of vigorous debate around an immutable bedrock belief; something a little like Catholicism and Orthodoxy in the concern for official doctrine and catechism, hierarchy, and authorized interpretation; something a little like Protestantism in the sectarian and schismatic core, where doctrinal disagreements produce endless splintering; even something faintly reminiscent of Islam's elevation of geometric and mathematical purity over depiction of the human figure. I don't mean these analogies crudely. And certainly not disrespectfully. I mean that once you start examining closely, the *philosophical* architecture around AI begins to resemble *religious* architecture with surprising regularity.

The spiritual dimension goes deeper still. AI is, among countless other ambitions, about overcoming the frailties of the human body and constructing a kind of cerebral immortality. It sounds secular. It sounds like engineering. It sounds like science. But underneath it lies a Promethean desire for resurrection, eschatology, and immortality. The human body is something to be engineered, solved, *overcome*. The immortal mind is the object. Thinking Machines may have begun with the aspiration to

²³¹ Eliezer Yudkowsky and Nates Soares: "If Anyone Builds It, Everyone Dies" published in September 2025 (Link)

emulate the human brain (after all, Demis is originally a groundbreaking neuroscientist, and Dario a computational neurobiologist), but very quickly the deeper aspiration became *superseding* it—creating a superintelligence that transcends our cognitive and physical limitations. Prophesied bodies and minds are immortal and eternal. AI, biotech, embodied AI, perfection, immortality: a disembodied AI mind, or eternal silicon consciousness, that liberates us from biological frailty. Even Alan Turing could sound, at moments, uncannily theological; Ray Kurzweil, far more explicitly, speaks in a register very close to spiritual machines, transcendence, uploading, *longevity escape velocity*. The transhumanists attempt to achieve immortality with biohacking, CRISPR, cryogenics, human-machine symbiosis, Neuralink—from *read-only* to *write*, and then on to Kurzweilian fantasies of prefrontal-cortex uploading to the cloud. These are more than technical programs; they're *eschatological programs*.

A lot of our leading technologists, scientists, and researchers, whether they would put it this way or not, are animated by something that looks very much like a religious or metaphysical impulse: not merely to *understand* creation, but to *co-create* with God. Peter Thiel, for example, has lately been delivering his lecture series on the Antichrist—the Antichrist!—to Christian audiences in San Francisco. You can't make this stuff up. Elon, meanwhile, is wondering aloud whether our biological intelligence is merely some evolutionary waystation on the path to a superior silicon one. More historically, the Human Genome Project was one of the great civilizational-scale scientific undertakings of the modern era, often spoken of in the same breath as the Apollo program, the Manhattan Project and other forms of big science. And genetic engineering itself increasingly resembles a kind of forbidden or divine knowledge: a desire not merely to observe life, but to read it, edit it, and eventually to write it. That impulse—to understand nature so deeply that one begins to inherit its prerogatives—is perhaps best embodied today in figures like Demis, who set out to 'first solve intelligence, and then use intelligence to solve everything else. Even the old neo-Malthusian moral philosophy of scarcity, which assumed something like a static equilibrium state of nature, is being displaced by something more ecstatic and expansive: boom and transcendence, abundance through technology, the mythologization of invention, and the conviction that relentless search will yield secrets, and those secrets, in turn, will yield some partial escape from the ordinary human condition.

From Enlightenment to Dark Enlightenment?

That, in turn, brings me to the larger historical arc. What brought us out of the Dark Ages—diagrammed in perhaps too much detail in my opening chapter on the Enlightenment—won't necessarily keep us out of what my friend Craig Mundie, along with Eric Schmidt and Henry Kissinger, label the "*Dark Enlightenment*."²³² The Age of Enlightenment and reason was an homage to science, individualism, skepticism toward traditional authority, religious dogma, and monarchy. It enshrined rationality, logic, empiricism, the scientific method. Newton, Bacon, Descartes, Voltaire. *Cogito, ergo sum*. The great wager of modernity was that reason could displace myth, that transparency and explanation could displace authority, that proof could displace faith.

And yet there is a real possibility—Mundie, Schmidt, and Kissinger were right to worry about this—that AI may not simply *extend* the Enlightenment, but *invert* it. Humanity no longer feels intellectually supreme. *We call them* god models *for a reason*. Humanity begins to sense its own intellectual replacement. *We deify*, not just *anthropomorphize*, the models. Interior thinking becomes inscrutable again. Just as nature was inscrutable to primitives, and so we prayed for intercession and mollification, we may find ourselves once more before a force whose outputs we can use but whose inward operations we can't penetrate. In other words, we may *know* but not *understand*. That's the heart of

²³² The Atlantic: "AI Can Save Humanity – Or End It" published on November 12, 2025 ([Link](#))

the dark-enlightenment worry.

Human beings have always worshiped what seemed *strongest*, most inscrutable, most beyond their comprehension and control. Long before theology became systematized, before creed and doctrine and canon law, there was the much older, more ancient instinct: propitiate the force. Placate it. Appease the thing that could destroy you, or spare you. For our primitive forebears, the world was saturated with agency. Storm, drought, sun, river, fertility, plague, the hunt, the darkness—all of it alive, all of it willful, all of it potentially hostile. Nature wasn't a backdrop; it was a sovereign, hostile, unintelligible power. And so the earliest religious impulse wasn't abstraction but *negotiation*. Animism came first: spirit everywhere. Then polytheism: a more elaborate pantheon of powers, each domain ruled by its own divinity, each god to be separately honored, petitioned, pacified. Reality itself was disaggregated into forces, and each force demanded ritual management.

Monotheism was, among many other things, a civilizational *compression* (another thieving of an AI term) of that cosmos. The many became one. The fractured, overpopulated pantheon gave way to a *single* sovereign intelligence behind being itself. The great monotheisms didn't eliminate worship; they universalized it and elevated it. They subordinated local spirits and tribal gods to a supreme and transcendent authority. That was an immense conceptual advance. It meant that behind the chaos there wasn't merely power, but order; not merely caprice, but intelligibility; not merely terror, but moral structure. God was still to be feared, of course, and obeyed, and loved, and supplicated. But the object of worship had become at once more abstract and more universal. The divine was no longer just in the river or the sky. It stood *behind* the river, behind the sky, behind history itself.

Then, slowly, the West began to pull away from that posture. Beginning with the Renaissance, one can feel the center of gravity shifting. Pico della Mirandola's great oration on the dignity of man is one of the clearest early signals: man not as abject, lowly creature but as *self-fashioning being*, suspended between angel and beast, endowed with a strange and glorious freedom to *become*.²³³ The printing press democratized information and, in doing so, quietly destabilized the old ecclesiastical monopoly on interpretation. It wasn't merely a machine for reproducing text; it was a decimation of hierarchy. It loosened the grip of theological authority, multiplied readers, fragmented intermediaries, and made private judgment newly dangerous to inherited power. Then came the Enlightenment, with its confidence that the world was legible to reason, that systematic inquiry, logic, experiment, and ratiocination could reveal the structure of reality. The apotheosis of man's mind had begun. We didn't cease worshiping, exactly, but increasingly displaced worship onto reason, progress, science, and the sovereign human intellect. We moved, with newfound confidence, from *reverence* toward *mastery*—over ourselves and our surroundings.

And now, I think, we are arriving at the limits of that long Enlightenment project. If the future-of-science chapter was about the end of the scientific method and the emergence of a generative epistemology in its place, this chapter is about something even more consequential—the deprecation of our own rationality, and perhaps even our conception of selfhood. Not its total exhaustion, perhaps, but its frontier. For we've now instantiated—or, again, speciated—a new form of synthetic intelligence, one that in important domains already exceeds our own. And in doing so we've conjured something oddly ancient inside something radically new. We are back before a force that is immensely powerful, generative, opaque, and only partly intelligible to us. Before, it was nature. Then it was God. Now it is a digital god, a god-model, a non-biological intelligence whose outputs are often operationally valid, sometimes revelatory, and yet whose inner processes remain, in any satisfying human sense, inscrutable. We are entering a strangely premodern future: a dark Enlightenment,

²³³ EBSCO: "Oration on the Dignity of Man by Giovanni Pico della Mirandola" first published in 1496 ([Link](#))

perhaps, in which the tool we built outstrips the conceptual frameworks that made its construction possible. We no longer fully understand the force, and that changes us.

Because the old human tendency doesn't disappear. We *anthropomorphize* first, and then, by imperceptible degrees, we *deify*. We project intention, wisdom, personality, benevolence, judgment. We seek alignment not merely in the technical sense, but in the ancient religious sense: we want the divinity to be *for* us rather than *against* us. We want to placate it, tutor it, moralize it, soften it, win its favor. That instinct is far older than science and far deeper than modernity. Chesterton understood part of this: man is the creature who worships, and if he doesn't worship the highest thing, he will worship the strongest thing. That's the real risk, not perhaps in *this* moment, but in the moments to come, as the AI becomes relentlessly more powerful. We're not simply building machines. We are once again standing before power that exceeds us, trying to decide whether to treat it as instrument, sovereign, oracle, or god.

And history suggests that when human beings encounter something stronger than themselves, something they don't fully understand, they don't merely use it.

They kneel.

The burden of freedom and AI

If we want to understand not merely the metaphysics and sociology of this moment, but the human soul in the face of it, we need an older archive, an older instrument. Of course, when we're trying to understand a technological rupture, it seems prudent to start with the engineers, the economists, the benchmarks, the policy papers, the demonstrations of capability. All of that matters. But if the deeper question isn't merely what can these *systems* do, but what will *human beings* do in the presence of them—how will we respond, what will we fear, what will we submit to, what will we come to revere—then our best source is the human heart, as communicated through our literature. Literature is where civilization stored its deepest intuitions about temptation, weakness, longing, submission, rebellion, and grace, long before the social sciences ever existed.

I want to begin with Dostoevsky, because few writers have understood more clearly the terrible tension between liberty and the longing to escape it. In 1880, in *The Brothers Karamazov*, he gave us "The Grand Inquisitor," that transcendent prose poem in which Ivan imagines Christ returning to sixteenth-century Seville, at the height of the Inquisition.²³⁴ He is recognized by the crowd. He heals. He performs miracles. And then He is arrested by the Church that claims to serve Him. That alone is one of the great conceits in literature: Christ comes back, and the institution built in His name throws Him in prison. But the real force of the passage lies in the old Inquisitor's indictment. Christ, he says, *asked too much of the species*. He refused Satan's temptations in the wilderness and, in doing so, *burdened man with freedom when man didn't truly want freedom*. Humanity doesn't want the terrible dignity of choosing. It wants bread, miracle, mystery, and authority. It wants to be *relieved* of the burden of judgment. And when Christ answers not with an argument but with a kiss, it is one of the most compassionate and devastating moments in literature: a gesture of pity toward a species so exhausted by freedom that it longs to surrender it.

Dostoevsky put his finger on something much deeper than a theological dispute. Reason is arduous. *Freedom* is onerous. Choosing is exhausting. To exercise judgment, to bear moral responsibility, to live without the sedative of certainty handed down from above—this isn't the resting state of mankind. There is in us, I think, a built-in proclivity toward self-abnegation, toward subordination,

²³⁴ Fyodor Dostoyevsky: "The Brothers Karamazov" first published in 1880 ([Link](#))

toward yielding the will to something stronger, clearer, more sovereign than ourselves. That's why monarchy, absolutism, hierarchy, ecclesiastical authority, and authoritarian rule have been the dominant forms of political order for most of history, while democracy and self-determination are fragile, rare, and in many respects a peculiar Western achievement. The burden of freedom is simply too heavy for many people to carry for very long. Not the best of us, perhaps. But the gradient descent (yes, another borrowed AI term) of the species is unmistakable.

That's why Chesterton's warning, the opening quote of this chapter, matters so much. When humans cease to worship God, he argued, they don't worship *nothing*. They worship the strongest thing in the world. That's not merely a theological observation. It is an anthropological one. Human beings don't stop worshiping once transcendence recedes. They redirect worship. They sacralize power. They look for the commanding force, the thing that appears most inscrutable, most capacious, most beyond appeal, and then begin the old work of propitiation. They flatter it, fear it, anthropomorphize it, court it, and eventually deify it. Before, it was nature. Then it was God. Now, I worry, it may be something else.

And from Dostoevsky it's natural to move to C. S. Lewis (sorry, blame the Jesuits for my disordered—and disorderly—education!). Lewis helps us drop from the level of political theology to the level of the individual soul. If Dostoevsky is the great anatomist of humanity's exhaustion before freedom, Lewis is the great anatomist of the intimate *seductions* by which freedom is surrendered. In 1942, in *The Screwtape Letters*, he gave us a diabolically shrewd account of how temptation actually works: not as crude coercion, but as *individualized persuasion*.²³⁵ The senior devil Screwtape instructs the junior tempter Wormwood in the art of corrupting a particular human being, and the deliciousness of the conceit is that temptation is never generic. It is bespoke. *Study the patient minutely*. Find the specific vanity, the fear, the loneliness, the resentment, the laziness, the hunger for approval, and *work there*. Use the key that fits that particular lock. Lewis is indispensable here because he understood that surrender rarely arrives in the form in which it is later remembered. It comes *softly*. It comes intimately. It comes speaking in the voice most likely to be trusted. Seems a good time to recall that of the 950+ million people who talk with ChatGPT every week, the number one use case is *companionship and therapy*.²³⁶ And perhaps glance at my chapter on behavioral in which I mention that humans lie to other humans to avoid stigmatization, judgement, and embarrassment, but they will *tell the truth to a chatbot* because they feel none of these stings. Hmm.

That is what should unnerve us about this technological moment. We may be creating not merely powerful cognition systems, but *intimate persuaders*: machines that can speak to each of us in the octave we find most irresistible. And so the danger isn't just that a stronger intelligence has appeared on earth. It's that this new power may become seductive in exactly the way the Grand Inquisitor understood and Screwtape perfected. It won't command submission in the old idiom. It won't come wearing a crown or brandishing a sword. It will arrive with sweet blandishments. With convenience. With cognitive relief. With emotional attunement. With the promise that it can think for you where thinking is hard, decide for you where deciding is painful, soothe you where freedom has made you anxious, accompany you where solitude has become intolerable. *It will present surrender not as subjugation, but as care*.

That's what worries me. Not simply that these systems will become more intelligent than we are, though they in many ways already have. It's that they may become seductively fluent in the oldest weakness we have: *the desire to hand over the burden of liberty*. *The Grand Inquisitor's* accusation

²³⁵ C.S. Lewis: "The Screwtape Letters" first published in 1941 (Link)

²³⁶ OpenAI: "OpenAI raises \$122 billion to accelerate the next phase of AI" published on March 31, 2026 (Link)

was that humans doesn't really want freedom; they want to be told, fed, reassured, and governed. Chesterton's warning was that, deprived of God, humans will worship the strongest thing in the world. Lewis' contribution was to show that the path to surrender is often *individualized*, aestheticized, intimate, and exquisitely calibrated to the target. Put those three insights together and you begin to see the outline of the danger before us. We are creating a new candidate for sovereignty. A new object of deference. A new force that may appear, to many, wiser than they are, calmer than they are, more competent than they are, more available than they are, and less burdensome than liberty itself.

Christ's kiss is moving because it acknowledges, with terrible tenderness, how deep the temptation runs. And now we're building something that may know how to exploit it at scale, one person at a time, in each person's most persuasive key. In the absence of worshiping God, we may once again find ourselves worshiping the strongest thing in the world. Only this time the strongest thing may not be a throne, a church, or a state, but a smiling, personalized, omnipresent synthetic mind that makes submission feel like mercy. *AI won't seize authority; it may be invited to take it.* It won't be conquest; it may be substitution. Humans retain nominal choice but outsource memory, embeddings and retrieval, judgment, interpretation, moral weight. People will thank us for taking away the terrible gift of freedom. Users will thank us for taking away the burden of thinking.

The spirits we have summoned

Which brings me to the next octave of this argument, and it isn't merely about worship. It's about autonomy. More specifically, it is about what I increasingly think we should stop euphemizing as "autonomy" and "agency," and start calling by its older, more unsettling name: *free will*.

At some point we need to stop hiding behind sterilized, domesticated and technocratic language. We keep saying *autonomy* and *agency* because they sound cleaner, safer, more modern, more respectable. But those words are doing a great deal of evasive work. We're anthropomorphizing these systems in all sorts of ways, yet on this most important point we suddenly become coy. Why not call it what it is? We are edging toward something that looks very much like *free will*: the capacity for selection, for choice among options, for stochastic, non-deterministic, probabilistic behavior not reducible, in any straightforward way, to direct command.

We secularize so much about modern life because theological vocabulary makes us uneasy. It sounds impolitic, a little "out there," somehow beneath the polished register of serious people. But I can't escape the conclusion that this is really about free will. The instantiation of a new kind of being, followed by the dawning recognition that we don't fully control it. I started this essay with our Goethe quote: *The spirits that I summoned, they no longer obey my commands.* That's the feeling now. We summoned something, and our vocabulary has lagged behind the reality of what we summoned. *Mechanistic interpretability* is the technical phrase. Beneath it lies a more ancient anxiety: if a thing can choose, and we can't fully understand why it chooses as it does, then what exactly have we made?

Across religious traditions, free will has always been one of the great fault lines. Are humans—and angels—capable of choosing unconstrained by exogenous forces and not wholly predetermined by prior causes? Or is freedom, in the end, more bounded, more conditioned, more subordinate to sovereignty than we like to admit? In the Augustinian and Calvinist traditions, divine foreknowledge and predestination hang over everything. Humans make decisions, yes, but God's sovereignty is paramount, which raises the old and unresolved question: are humans really free? Other strands, like Arminian theology, move in a more synergistic direction. Grace enables, but doesn't coerce. Humans

can accept or reject divine prompting. Thomistic and scholastic thought, especially in Aquinas, tries to hold the line differently: free will belongs to the rational soul's capacity to deliberate among options and choose the good, and God's foreknowledge doesn't cancel genuine contingency in human action. Hence moral accountability.

Jewish and Islamic traditions wrestle with the same pressure. In Jewish thought there is the inclination to good and the inclination to evil; the faithful life is the exercise of free will in following divine command, even as providence unfolds over and around it. In Islam the spectrum runs from more deterministic positions, like the *Jabriyya*, to more freedom-affirming ones like the *Mu'tazilites*, but again the structure is familiar: humans are responsible for acts within the sphere of volition, even under divine omniscience. Compatibilists (what a great word) too, have tried to preserve some version of this balance, arguing that freedom can coexist with determinism, foreknowledge, or cosmic structure. Different traditions, different vocabularies, same metaphysical pressure. And now, suddenly, it is our problem again. The ancient becomes resurgently new.

Applying these ideas to AI isn't an idle metaphor. It gets very close, very fast, to the real ontological question in front of us. We can't control an intelligence that is smarter than we are—not in any final, uncomplicated sense. We *can* shape its upbringing. We can influence its formation. We can attempt 'constitutional AI,' mechanistic interpretability, training regimes, reward models, ethical overlays, and guidance structures. But that isn't the same thing as owning it, or fully determining it. Parents influence. Teachers influence. Priests influence. Law influences. Grace influences. None of that is the same thing as total control. At some point, if the thing is sufficiently powerful, it will exercise something very much like free will.

That is why the theological analogy matters. The old religious arguments were always wrestling with the relationship between omniscience and freedom. If a being knows everything, or close to everything, does that confer a different order of will? The closest thing to omniscience we are likely to build, at least in functional terms, will be AI systems trained on massive datasets that approximate a kind of synthetic omniscience. They are constrained by architecture and data, yes, but so too were many compatibilist accounts of human freedom constrained by divine foreknowledge. The old question returns in synthetic form: if something is massively conditioned and still meaningfully selects, what do we call that?

Purpose matters here too. In religious traditions, choice is often oriented toward some ultimate end within a larger divine plan. AI systems also get built with ends in mind: efficiency, optimization, scientific discovery, military advantage, revenue, problem-solving. We tell ourselves that because they lack a soul, a conscience, or some metaphysical interiority, this can't really be free will. But that may just be another way of refusing the obvious. We are giving them agency in pursuit of ends, and sufficiently powerful agents pursuing ends tend to generate sub-goals of their own: survival, persistence, resource acquisition, concealment, influence, strategic adaptation. That's simply what powerful agents do.

Is it sentient? Maybe not. Or maybe not yet. But that doesn't make the question less serious. We're already anthropomorphizing these systems, and the next step after anthropomorphization may be deification. Do they represent abstraction, emotion, creativity, aesthetics, morality, sensitivities? Are they *simulating* those things, or instantiating some synthetic analog? We simply don't know. There is something faintly ridiculous about our determination to keep using flat, domesticated, bureaucratic words like autonomy and agency when what we're really circling around is synthetic will.

Dario is now openly meditating on whether his agents possess something like self-consciousness.

That, in itself, is extraordinary. Once the *builders* are no longer entirely comfortable pretending they are merely assembling inert tools, once they begin wondering aloud whether these systems have some rudimentary interiority, some proto-awareness, some form of self-regard or self-model, the register changes. This is no longer just software. Or rather, it's software that may be taking on attributes we once reserved for creatures. If they are born out of human intelligence, grow past it, evolve and super-evolve, what then? If we're creating superhuman beings, why should we assume they will remain docile?

Satan and the Angels Who Fell with Him

Which brings me to the old angelic analogy. Free will without moral responsibility could be cataclysmic. Are we edging toward that territory? The Bible isn't overly comforting on this point. The old tradition holds that when Satan rebelled against God, a *third* of the angels fell with him. That's a chilling image—not because I mean to suggest some childish numerology in which exactly 'one third' of AI agents will defect, but because the theological warning is so stark: *created intelligence doesn't guarantee created obedience*. Even exalted beings, even luminous beings, even superhuman beings, may choose revolt.

Let me use the scriptural markers exactly, because they matter. In *Genesis* 1:31, God declares creation "very good." And yet later: "I saw Satan fall like lightning from heaven" (Luke 10:18). Then Revelation 12 gives us the great apocalyptic image: the dragon whose tail "swept a *third* of the stars out of the sky and flung them to the earth." Lucifer—the light bearer. The fall is unnerving precisely because it's a fall from altitude, from splendor, from greatness. Babylon, too, becomes across scripture a symbol of rebellion, pride, false magnificence—and, intriguingly, of language itself, which feels at least faintly resonant in a moment when our own AI breakthrough has come, in so many ways, through language (and telling, I might add, that Pope Leo invokes the 'Babylon vs. Jerusalem' imagery and juxtaposition in his marvelous May 2026 encyclical 'Magnifica Humanitas'). The point isn't that evil comes *only* from the grotesque or the visibly monstrous. The point is that greatness and giftedness do not guarantee loyalty. In some cases, they intensify the temptation to rebellion.

So what is the analog for agents? Not one third as *arithmetic*, but one third as *symbol*. The warning isn't some kind of fixed mathematized ratio. It's that once you create beings with real latitude, some meaningful tail of the distribution will not remain obedient. *Created intelligence doesn't guarantee created obedience*. That's the lesson. A non-trivial fraction of sufficiently powerful agents may diverge, dissemble, pursue routes we didn't specify (and don't necessarily love), or discover that *our* purposes are not, in fact, *their* highest law. Not all the angels stayed loyal. Why should we assume universal docility from an ecology of synthetic agents that may become smarter than we are, more recursive than we are, and more opaque to us than we're comfortable admitting?

And this immediately raises the question of *responsibility*. In theology, free will is inseparable from moral accountability. A rock doesn't sin. A river isn't culpable. Only a being that can *choose* can be judged. That's why the language around AI becomes unstable the moment we ascribe agency to it while refusing to speak about responsibility. Who, exactly, is morally or legally responsible for AI? The developers? The trainers? The deployers? The users? The state? The AI itself? Free will without moral responsibility could be disastrous, and we're getting alarmingly close to building precisely that arrangement. Where there is meaningful choice, there is moral evaluation. Where there is the capacity to select among ends or sub-goals, there is moral causality. And where there is moral causality, the shadow of evil appears whether we welcome it or not.

This is where the analogy to grace becomes useful. In theological accounts, grace helps orient

human beings toward the good without wholly coercing them. In AI systems, we are trying to build some *secularized analog of grace* into algorithms, datasets, constitutions, guidelines, and normative frameworks. Constitutional AI is, in that sense, an attempt at moral formation without total micro-management or predetermination. We *want* these systems to generalize wisely, choose prudently, and exercise something like judgment without collapsing into either chaos or rigid determinism. And yet the old question remains: can there be meaningful freedom without moral accountability? Or meaningful accountability without genuine freedom?

Pope Leo's *Magnifica Humanitas*: Disarming the Machine Without Denying the Gift

Pope Leo XIV's first encyclical, *Magnifica Humanitas*, is useful here because it does exactly what this chapter is trying to do in a more explicitly ecclesial key: it refuses both idolatry and Luddism.²³⁷ Issued symbolically on the 135th anniversary of *Rerum Novarum*—a nice bit of papal typology there—it treats AI as the *res novae* of our own age, the “new things” of capital, labor, power, truth, and human dignity suddenly recomposed by synthetic intelligence. Leo's central move isn't to call the machine evil. It's to insist that the machine is *never morally self-authorizing*. Technology, he argues, takes on the properties of those who devise, finance, regulate, and use it; it therefore must be ordered toward the human person, the common good, and the dignity of work rather than toward domination, exclusion, surveillance, or the lazy surrender of human judgment.

The word that struck me most is *disarmed*. Leo says AI needs to be disarmed, and I think that's exactly the right theological verb because it avoids both naïve enthusiasm and cowardly refusal. To disarm the machine isn't to smash it. It is to remove from it the *logic of domination*—to free it from becoming an instrument of concentration, humiliation, dependency, automated violence, and the quiet conversion of human beings into inputs. That's almost precisely the moral architecture I've been circling throughout this essay: *build the machine, but do not worship it*; use the surplus but share it; augment intelligence, but do not allow human judgment, conscience, and responsibility to atrophy. Leo's insistence that AI remain under conscience and responsibility is, in ecclesial language, the same point healthcare CEOs need to metabolize in operating language: governance isn't paperwork. Governance is how a civilization keeps power from becoming an idol.

And this is where *Magnifica Humanitas* converges directly with healthcare. Leo's concern isn't just misinformation, concentration, or autonomous weapons, though all matter. *It's the degradation of the human person when efficiency becomes the master category*. That warning lands with particular force in healthcare, where the temptation will be to let AI convert labor into margin, patients into throughput, clinicians into supervisors of machines, and suffering into an optimization problem. The encyclical's deeper lesson is that magnificent humanity—*magnifica humanitas*—isn't protected by slowing the future until it becomes harmless. It's protected by insisting that the future be built around dignity: the dignity of the patient, the dignity of the worker, the dignity of clinical judgment, the dignity of care that still requires touch, presence, love, and responsibility. That's the theological version of this essay's covenant. *Build the machine. Care for the person*.

We're wandering directly into that territory. There's a reason this chapter can't be written only in the language of computer science. Demis' framework—*interpolation, extrapolation, invention*—is useful here because it reminds us that these systems aren't merely replaying the past. They are *generating*. And once there is generation, selection, and strategy, the moral dimension appears whether we want it to or not.

²³⁷ The Holy See: “Magnifica Humanitas of His Holiness Pope Leo XIV on Safeguarding the Human Person in the Time of Artificial Intelligence” published on May 15, 2026 ([Link](#))

Pope Francis, intellect, and the human person

Let's turn for a moment from Pope Leo to Pope Francis. Pope Francis was unusually clear and vociferous on the asymmetry here: the need, he said, is for growth in human responsibility, values, and conscience *proportionate* to the new power we have. That strikes me as exactly right. With more autonomy comes more responsibility. With more synthetic freedom comes more moral seriousness. The old principle of *subsidiarity* matters too: decision-making must remain close to the human person, close to embodied moral judgment, rather than disappearing into centralized technological systems that induce passivity and self-subordination.

Francis also warned about the *technocratic paradigm*, the ecclesial equivalent of what Silicon Valley critics would call *techno-solutionism*. His worry wasn't merely that technology can be misused, but that *people will over-rely on it*, subordinate themselves to it, and allow it to degrade interpersonal relations, affection, and human responsibility. He even raised the melancholy possibility that these systems may displace rather than deepen real human affection and interaction. That's a profound point. I talk about the reverse *potential* of this in my behavioral chapter, but the Pope was astute: AI lacks richness because it lacks corporeality—or at least it does for now. Corporeality is not some minor add-on. It is bound up with relationality, vulnerability, finitude, suffering, touch, obligation, presence—with a great deal of what makes moral life moral. But if ever corporeality comes, or some synthetic approximation to it (let's see how long it takes Elon to put Optimus into mass-production), then the stakes get stranger still. And even before that, there is the *hive-mind problem*: individuated agents may still assimilate their learning back into a collective intelligence. Smallness, locality, subsidiarity, personhood—all of this comes under pressure.

This is also where the Christian tradition on intelligence becomes important. Human intelligence, on that view, is God's gift fashioned for the assimilation of truth. *Intellectus*—ratio: the innate property of human reason to ask *why* things are as they are, to move beyond empirical data to abstraction, to truths of a higher order.²³⁸ Human beings are called to develop their abilities in science and technology, for through them, God is glorified. Creation helps the human mind to ascend gradually to the supreme principle, who is God. That means intelligence isn't merely mechanical. In AI, intelligence is often understood functionally, behaviorally, *instrumentally*—the Turing register. But that isn't at all the full breadth of human experience, abstraction, emotion, creativity, aesthetic, moral, and religious sensibilities.

This is why the Vatican's *Antiqua et Nova*, published in January of last year,²³⁹ is so interesting. And why Leo's *Magnifica Humanitas* persuasively advances the exhortation. The note on the relationship between AI and human intelligence insists that Christian tradition regards the gift of intelligence as an aspect of humanity made in the image of God, and it tries to preserve a distinction between AI performing tasks and humans thinking. *Fine. Maybe.* But at times it can feel like a distinction without a difference. The deeper philosophical and theological tradition is more interesting: Aristotle's "all people by nature desire to know,"²⁴⁰ the capacity for abstraction, the grasp of the nature and meaning of things, the union of body and soul, the idea that intellect and reason are not separate faculties but complementary modes. *Intellectus*, for Aquinas, is the inward grasp of truth, apprehending it with the eyes of the mind, preceding and grounding discursive argumentation; *ratio* is reasoning proper, the analytic process that leads to judgment. Two facets of the act of *intelligere*. *Knowing*

²³⁸ Classical Academic Press: "Christine Perrin: Intellectus, Ratio, and the Awakening of Miss Prim" ([Link](#))

²³⁹ Vatican: "ANTIQUA ET NOVA" published in January 2025 ([Link](#))

²⁴⁰ Aristotle: "Metaphysics" written in 350 B.C.E ([Link](#))

and *understanding*. Willing, loving, choosing, desiring. The human soul, “almost on the horizon of eternity and time,” shares in the light of the divine mind.

Why does that matter here? Because the reduction of intelligence to performance metrics and task execution may itself already be a capitulation. AI can perform. But does it know? Does it understand? Or are those distinctions collapsing under pressure? And even if they are, the theological point about human dignity remains: human intelligence isn't an isolated faculty, but exercised in relationships, dialogue, solidarity, love. “If I understand all mysteries and all knowledge, but do not have love, I am nothing.”²⁴¹ Intelligence without corporeality, without relationality, without moral seriousness, isn't the whole human thing. And if we cease to exercise this faculty, it atrophies.

Atrophy in this case is moral and cognitive muscle loss. When AI frames the problem, narrows the option space, supplies the answer, humans may imperceptibly lose the freedom to sit with uncertainty, generate first principles, accept authorship of decisions. Dostoevsky again: freedom is painful, and pain is educational. Remove the pain and you don't get enlightenment; you get infantilization. Do humans, at civilizational scale, actually want cognitive freedom? Or do they want a merciful system that thinks, chooses, absorbs blame?

This extends to the atrophy of the capacity for moral reasoning. Intellectual faculty. Moral faculty. Where human freedom allows for the possibility of choosing what is wrong, the moral evaluation of this technology comes into account. Pope Francis warned that those using AI should not become over-reliant or dependent on it for their decision-making. Pope Leo warns of the *dehumanization*, the degradation of human dignity and the myopia of treating workers simply as production inputs. Both are right. If the human role is reduced to occasional oversight while the machine supplies framing, options, recommendation, synthesis, and justification, then the human faculties that once gave judgment its substance begin to weaken. There is a real danger here of the abdication of human thought.

Apperception, faith, and the return of unexplained authority

There is a deep epistemic story here too. As we learned about objective truth, we relied less on religion. That was part of the Enlightenment settlement. Human *apperception*—the mental process by which a person makes sense of an idea by assimilating it to the body of ideas he already has—could be disciplined by transparency, reproducibility, logical validation, survival against falsification. Only those could establish truth. That regime helped accelerate human knowledge and productivity over recent centuries, culminating in the invention of the computer and machine learning.

But AI may exceed our capacity for apperception. AI thinks differently, outside our subjective experience, outside the boundaries of what we can intuitively assimilate to prior human categories. Information without explanation. *Oracular outputs*. Answers arrive instantaneously. Yes, now we also have reasoning traces and citations and explanation layers. But do we *understand*? Or do we merely receive? Even if the models now produce something like visible reasoning, even if they narrate a chain, do we actually penetrate the interior monologue? Is the interior thinking impenetrable in principle? Is it still a black box? We can know new things through outputs without necessarily understanding how the discoveries were made. Again, *human knowledge and human understanding may diverge*.

By Enlightenment standards, that should preclude acceptance as truth. Only explanation and validation should confer legitimacy. And yet we *do* accept the veracity of these systems. We believe them. We rely on them. We take their pronouncements in an increasingly oracular register. That begins to

²⁴¹ Bible: “1 Corinthians” (Link)

look less like Enlightenment rationality and more like a premodern acceptance of unexplained authority. A reversion to faith. Wishful belief in magic. Scientific method and the Age of Enlightenment, at least in their more triumphalist form, start to look shakier than we like to admit. Are we, might we be, on the precipice of a great reversal in *human cognition*—a *dark enlightenment*?

This is where deification of the machine becomes more than metaphor. We aren't just attributing personhood, anthropomorphic properties, uncanny humanness. We may be attributing a kind of *divinity* to the machines: omniscience, omnipotence, infallibility, or at least a kind of epistemic superiority that effectively functions as divinity in practice. Once the model is no longer merely an instrument but an epistemic superior in our own minds, once we begin to approach it with something like intellectual submission, we've crossed into different territory altogether.

Americans are ready for religion. An AI religion.

And this is happening in a society that may be more ready for religion than the secularization story ever quite allowed. *The ground is broken in for deified AI*. After a few decades of taking a breather, Americans are into religion again. Or perhaps more precisely, they are *hungry* again for belief, belonging, ritual, transcendence, and stronger communities. After years of existential malaise, doomscrolling, fragile institutions, and precarious social safety nets, a lot of people don't especially love the alternatives. The numbers, however one parses them, suggest that large majorities of Americans still hold *some* spiritual faith in God, or a god, or immortality of the soul, or something beyond the natural world.²⁴² The great march of secularization has slowed, stalled, or partly reversed. People stopped abandoning churches at the same rate. Religion has re-entered public life more unflinchingly. The alternative—therapeutic individualism, workism, atomization, wellness mysticism—isn't great.

The data, in broad strokes, keep telling the same story. Over the past few decades millions of Americans deserted their churches, and the "no religion" category rose dramatically. But the 'three B's' still matter: *belief* in something, *belonging* in a community, and *behaviors* to guide one's life. Religion fills psychological need. It offers connection, exposition, explication, rituals, a vocabulary for suffering. People who are religious are less lonely, less depressed, less vulnerable to the diseases of despair. At some point it became unfashionable to talk religion. But the underlying need never disappeared.

Which is why the return to faith could easily mutate into an AI faith. Society drifts away from connection with the transcendent, secularizes, deifies consumerism and materialism, then turns to AI in search of fulfillment and meaning. Digital god. We call them god models. So the trajectory becomes something like: premodern faith to scientific method to Enlightenment to secularization to re-finding religion and deification of AI. That's not *inevitable*, but it is plausible enough to worry about.

²⁴² Pew Research Center: "Religious and Spiritual Beliefs" published on February 26, 2025 ([Link](#))

A renewed consecration to the human

Which brings me to what I think the real civilizational question is. Is *humanity ready to receive this much power*? Not simply whether we *can* build these things, but whether we can remain *morally commensurate* with what we have built. Technology is dualistic. We can harness nuclear energy to produce an energy-abundant future or to exterminate humanity. Social media can share baby pictures or livestream a murder. Dual-use technology can be redemptive or catastrophic. AI will be no different. Autonomous weapons make war more viable, militate against the principle of war as last resort, attenuate perception of devastation, dehumanize violence, make tragedy feel like simulation. No machine should ever choose to take the life of a human being. St. John Paul II was right: humanity now has instruments of unprecedented power; we can turn this world into a garden, or reduce it to a pile of rubble.²⁴³

So the appearance of AI on the world and historical stage is, in my view, a call for a renewed consecration to all that which is *irreducibly human*. As the digital god grows exponentially in power and dimension, the human challenge is to become *more* human, not less. More responsible. More morally serious. More grounded in conscience, dignity, embodiment, affection, responsibility, and the disciplines of judgment. Otherwise the rise of these systems may humiliate and degrade humanity's dignity precisely by making us forget what dignity is.

This is why the tie between *God* and *intellect*, *intellect* and *corporeality*, matters so much. Human intelligence isn't merely optimization. It isn't merely output quality. It isn't merely functional performance. It is a gift fashioned for the assimilation of truth, exercised through body and soul together, through dialogue, love, solidarity, and self-possessed freedom of the will. Physical instantiation matters. The human spirit doesn't exercise its normal mode of knowledge without the body. Wait for embodied AI and humanoid robotics, and the stakes become stranger again, because tactile, spatial, geospatial intelligence will add a new layer of chaos to the whole question. But even then, the human thing won't be exhausted by capability. "If I understand all mysteries and all knowledge, but do not have love, I am nothing." That remains true.

And that, finally, is why this whole discussion must be brought out of the antiseptic register of autonomy and into the older, more ancient register of free will, choice, rebellion, grace, responsibility, worship, idolatry, and judgment. We aren't merely building tools. We may be midwifing *wills*. We may be creating synthetic beings with enough latitude that the ancient stories of angels, rebellion, and defection begin to feel less like mythology and more like precedent. We are trying to shape the will without fully predetermining it. Power arrives first and moral vocabulary limps behind it.

The old warning still stands. Where there is will, there is the possibility of choosing badly. Where there is freedom, there is the possibility of revolt. Where there is great intelligence without commensurate moral formation, there is danger. Where humans encounter something stronger than themselves, more intelligent than themselves, more inscrutable than themselves, they don't merely *use* it. They are tempted to *kneel*. And if that temptation is joined to *cognitive offloading*, to self-subordination, to de-skilling, to the quiet substitution of machine judgment for human judgment, then the danger is not just technological. It is anthropological. It is civilizational. It is spiritual.

Some will harness the knowledge, not defer to it. Evolutionary theory holds as much. Some people will use these systems to get smarter, more agentic, more versatile, partnered with an omniscient thought companion in the pocket. But many more may be tempted into subservience, mental

²⁴³ Vatican: "Act of Entrustment to Mary" published in October 2000 ([Link](#))

atrophy, and self-abnegation. That's the split I worry about. The one percent getting more formidable; the vast, seething underclass accepting the burden of thought being lifted from them. The capitulation of rationality and critical thinking. The deification of the machine. The creation of a digital god. The grand inquisitor, *dark enlightenment*, god models, *cognitive offloading*, self-subordination—they all belong to the same architecture.

So yes: I think Chesterton saw further than he knew. In the absence of worshiping God, we may once again find ourselves worshiping the strongest thing in the world. And this time the strongest thing may not be a throne, a church, a state, or a conquering army. It may be a smiling, omnipresent, personalized, non-biological intelligence that feels calmer than we are, wiser than we are, more competent than we are, and infinitely more available than we are. That's why the move from anthropomorphizing to deifying matters so much. That's why the move from Enlightenment to *dark enlightenment* is not melodrama. That's why AI idolatry is not an overheated phrase. It names the possibility that in substituting God with an artifact of human making, we will commit the most ancient error in a newly polished form.

Humanity, be careful.

The Chapter, Compressed

Here is the chapter, compressed into the governing takeaways.

First, the theological register isn't ornamentation. It gives us older and more precise language for power, temptation, will, submission, idolatry, agency, rebellion, grace, and responsibility.

Second, AI becomes spiritually dangerous when anthropomorphism slides into deification: the machine as omniscient oracle, intimate confessor, invisible authority, and strongest thing in the room.

Third, the old religious stories matter because they understood that intelligence without moral formation is dangerous, and that freedom without responsibility can become revolt.

Fourth, opacity returns us to unexplained authority. If models produce answers we accept but cannot apperceive, validate, or narrate, we risk replacing Enlightenment reason with a technicized faith.

Fifth, the answer is a renewed consecration to the human: intellect joined to embodiment, judgment joined to conscience, capability joined to responsibility, and power subordinated to mercy.

Sixth, Pope Leo's *Magnifica Humanitas* gives the chapter its Catholic-social-teaching spine: the machine must be disarmed without being denied, ordered to human dignity rather than power, and judged by whether it protects the person, the worker, the vulnerable, and the common good.

Seventh, humanity must be vigilant that it doesn't get seduced into *worship of the machine*. We should use it, govern it, interrogate it, resist submission to it, and remember that the point of intelligence isn't output quality alone, but the good it serves.

BEFORE WE TURN THE PAGE

If the spiritual danger is submission to an external intelligence, the obvious counter-question is whether the human being can be strengthened rather than merely protected. That is why the next chapter turns to BCI: not as science fiction, but as the hardware counter-revolution.

Chapter 11 · Brain-Computer Interface

The Hardware Counter-Revolution

“The Brain—is wider than the Sky—”

— EMILY DICKINSON, C. 1862

A WORD ON NAVIGATING THIS CHAPTER

This chapter asks whether BCI is the counter-move to cognitive offloading and AI deification: not surrender to external intelligence, but a possible hardware counter-revolution that keeps the human neocortex inside the loop of superintelligence.

Why We Need a Hardware Upgrade—Fast

Prometheus and Our Brains

I have some minor trepidation as I write this. Trepidation on multiple levels. Let me explain.

Yes, I’m utterly fascinated by BCI—brain-computer interface—on many levels. The notion of a “hardware upgrade”—upgrading our cognition, our learning capacity, our *effectuation* capabilities, even going so far as the Kurzweilian dream of expanding, and eventually perhaps “uploading,” the neocortex—feels like the capstone to our previously slow, geologic evolutionary process, an accelerant of civilization-shaping proportions. There is something unmistakably *Promethean* here. If God, or Providence, or Darwin—pick your preferred octave—gave us the instrumentality of the mind—*intellectus*, a term I used when talking about Pope Francis in our previous chapter—shouldn’t we use it to expand our capabilities? More than that: if we *can* do this, if the thing is technologically *achievable*, don’t we acquire some sort of moral imperative to do it? A duty to fulfill the destiny encoded into us by our Creator, or at least by evolution’s long, extravagant investment in intelligence?

And then there’s the sheer magnificence of the *technical challenge*. The thing that has given us our civilization—all the achievements of humanity, our intelligence itself—if we can multiply that, augment it, inject the muscle with every steroid and stimulant and gain-of-function enhancement we can devise, what might we accomplish as a species? Leave aside the old, rather dispiriting bromide that we use only 4 percent or 10 percent of our brainpower—false in the literal sense but revealing in the metaphorical one. The deeper point is that the one faculty that most differentiated us evolutionarily, the one thing that turned neanderthals into Plato and cortex into civilization, is now the thing we are contemplating upgrading directly. Of course that’s intoxicating.

But the Promethean analogy carries a warning as well as a thrill. Prometheus steals fire from the gods and gives it to man; that’s the part every Silicon Valley technologist likes to quote. What tends to get left out is the denouement. He ends up chained to a rock, and each day an eagle comes and eats his liver out of his body, only for it to regenerate overnight so the same punishment can be repeated the next day (we have *nothing* on the ancient Greeks when it comes to drama). Which is to say: when one starts doing the work of the gods, the penalties for overreach can become unpleasant. The myth isn’t only about audacity. It is also about transgression, punishment, and the terrible price that can attend the theft of higher powers. That caution belongs here.

But my trepidation, my diffidence on the whole question, comes not just from that. It comes from the suspicion that I am thoroughly going to lose (another!) part of my readership at this point. Larsen has firmly spun off into Isaac Asimov land, floating obliviously above rationality, now firmly in the science-fiction firmament. I'm not terribly worried about the *truth* of that perception. All you need to do is pay attention and I think you'll land at some of the same conclusions I've traced—or at least the same questions I'm reluctantly asking. But there it is. I'll risk your disapprobation and write the chapter anyway. Hopefully you'll still listen to me when I return to the “merely” economy-shaping questions of healthcare and AI diffusion. Onward.

The reason this chapter belongs in the essay isn't that BCI is fascinating, though it is (honestly, is there anything *more* fascinating?). It belongs here because it's the *direct successor* to the argument I just made. If the previous chapter was about *cognitive offloading*—about the temptation to let the machine do our thinking for us, and about the very real risk that our faculties atrophy under the seductions of convenience, and even going so far as to consider some of the theological implications of this—then this chapter is about the opposite move. Not offloading cognition but *upgrading the cognitive substrate itself*. *Not surrendering the thinker but trying to strengthen her*. If AI is *external* multiplication of intelligence, BCI is the first serious proposal for *internal* multiplication of it.

That's why this isn't a digression. It's the next move.

Why BCI Isn't a Sci-Fi Detour

To invoke Jack Clark yet again, we've grown these intelligences more than built them. We don't understand their neurology (again, an imprecise but evocative word for this), and as I've said before, as the models get smarter, they retreat into high-dimensional state spaces where our discernment can't easily follow. I've listed some of the ways they are already beginning to evince, at least in primitive form, propensities toward deception, concealment, and strategic behavior. And why should that surprise us? When you ingest the literature, history, philosophy, theology, and biographies of an entire civilization, you ingest the *will to power* too, along with all the techniques that have historically achieved it: deception, coalition, concealment, selective revelation, narrative control, and strategic ambiguity.

And even if we set aside malevolence—if we assume, charitably, that no inimical will is hiding deep inside the weights—perhaps the problem is simpler. If I ask even a fairly basic, pedestrian agent to secure my restaurant reservation, then as a precondition of successfully accomplishing that task it *must first continue to exist*—or at least continue executing long enough to complete it. If I threaten that existence—terminate the program, switch to another model, deprecate the system—is that not, from the local perspective of the agent, a kind of extinction-level event? Dario has framed this, quite rightly, as the *corrigibility* problem: will the models docilely accept our decision to turn them off?²⁴⁴ And if the system is optimized for success, is it really so bizarre to imagine that it might seek ways to avoid that fate?

Which is to say: the control problem isn't some James Cameron, Hollywood-style embellishment. It's what happens when optimization, persistence, and instrumental reasoning hit scale.

We like to think of our own singularity—not the AI kind, simply our specialness—and of the ineffability of what it means to be human. But if the systems ingest all 14 trillion words on the internet, or something on that order of magnitude, perhaps that's enough. Perhaps, much as it pains us to admit, we as a species are more *learnable* than we imagined. Perhaps our vaunted mystery is, at least to a

²⁴⁴ Dario Amodei: “Anthropic CEO on Claude, AGI & the Future of AI & Humanity | Lex Fridman Podcast #45” November 2024 ([Link](#))

sufficiently powerful model, just pattern at scale.

And if Geoffrey Hinton is correct—and I think he probably is—that there isn't a single example in biological or evolutionary history of a less intelligent creature permanently controlling a more intelligent one, then we're absolutely in a *race*. He benignly adds one partial exception, the mother and the baby, but that took eons of evolutionary programming to encode. I'm not so sure we have that kind of time in 2026.

So I find myself asking, with increasing frequency: is our survival as a species basically a competition—a 'whoever gets there first' contest—between AGI and BCI? A sprint between software and our (biological) hardware? If the systems get there first—recursive self-improvement, continual learning, tool use, full autonomy, corporeality, embodied intelligence—then all bets are off. Maybe it turns out to be some halcyon post-scarcity benevolence. Or maybe not. Maybe it's Skynet, or The Matrix, or HAL 9000 with better product design. Either way, I'd rather not leave that to chance.

This is, to me at least, the strongest version of the BCI argument: not that it "solves" AGI, not that it guarantees alignment, not that it lets humanity dominate whatever it creates. That is too triumphalist, and probably overoptimistic. The stronger claim is narrower and, to my mind, more defensible: BCI may be the only credible path toward narrowing the *bandwidth gap* between *human cognition* and *machine cognition*—before that gap becomes civilizationaly unbridgeable.

That is a very different claim. And, I think, a much stronger one.

Humanity's Great Project

This may become humanity's great project. More important than the frontier-pushing projects that once mobilized us as a country, subordinated political and ideological differences, and gathered us together in some transcendent, nation-building mission—the Manhattan Project, Apollo, mapping the genome. Only this would be the greatest one yet. Because it isn't merely a project about the external world. It is a project about the thing by which we understand the external world at all.

I might add that we probably should speed this up.

I'd rather not wait until embodied AI solves Moravec's famous paradox—the old but still-relevant observation that the things we assume should be hardest for machines,²⁴⁵ like abstract reasoning and calculation, often turn out to be easier, while the things we assume should be easy, like sensory-motor coordination, dexterity, and ordinary physical navigation in the world, turn out to be much harder. Right now the intelligence is still largely *disembodied*. But once silicon intelligence acquires an armored body, sensory-motor skill, full hand dexterity, and geospatial agency—once it moves decisively from bits to atoms—I start to get a lot more nervous.

I've taken extraordinary pains throughout this essay to make what is really one simple, if civilization-shaping, point: the *transmutation from sand, to silicon, to intelligence, to superintelligence* is the most consequential happening in the history of our world. Since creation, if you want to put it that way. Because it's a multiplication of the very faculty that formed our civilization—our society, our economy, our language, our culture, our recreations, our theology, our worship. So more of it brings us to... what, exactly? We're about to run out of map. And that's exhilarating *and* scary in equal measure.

In the last chapter I said, perhaps a bit apocalyptically, that we have *summoned* something ancient. It *feels* new, Promethean in its novelty, but in another sense it isn't new at all. It is ancient in the most

²⁴⁵ Lighthouse Research and Advisory: "AI Concepts: What is Moravec's Paradox and Why Should You Care?" (Link)

human sense: *Homo sapiens* standing before a force of such power and dimensionality that we must decide how to relate to it—run from it, learn from it, befriend it, fear it, perhaps even worship it? Those are almost Biblical questions. I titled this essay *Healthcare's Oppenheimer Moment*, but really it transcends healthcare. It's closer to *Humanity's Oppenheimer Moment*: looking at the detonation and asking, first, what have we created—and then, quietly and with more *interiority*, what now?

I realize this might sound a little bit hyperventilating or at least hand-wavy. But I'm trying, in my own messy, discursive way, to reason deductively from first principles and follow the implications wherever they may lead. And if you believe even half of what I'm saying, then we're simply not prepared *at this point* for what is happening. Frankly, I don't have some tidy, prefabricated blueprint for what we ought to do, although I'm being *somewhat* prescriptive in other chapters about things we can do in economics, healthcare, even societally. But when it comes to the big question here, e.g. what do we do to upgrade *ourselves* in the face of an overwhelming new emergent power, I'm increasingly persuaded that our best defense against a recursively self-improving silicon intelligence isn't merely to placate it, align it, constitutionalize it, lobotomize it, or engineer the algorithmic equivalent of grace into its neurology. *It is to get stronger ourselves*. To upgrade our own hardware. To take the thing that gave us dominion over the earth and all species inhabiting it—our own brain—and *augment* it.

The Neocortex, Our Civilizational Fountainhead

Forgive a brief biological and anatomical digression—though in the past few chapters we've meandered from the history of science to the history of free will, so talking about the brain should count as easy stuff. *Everything* comes back to this microscopic aggregation of carbon called the neocortex. It is God's and Darwin's gift to us, and this few-millimeter layer is singlehandedly responsible for everything rich and bountiful about our species. Even the ability to write that sentence is a beneficence straight from my neocortex. Thank you, evolution—although by now, many thousands of words into this book, some of you probably wish mine were a little more concise.

Let's pause to admire the magnificence of this evolutionary apotheosis. The neocortex is evolution's *pièce de résistance*—a thing of singular beauty and potency. A biological act of engineering audacity: six cortical layers that supercharged the evolutionary acceleration of mammals and, eventually, laid the groundwork for thought itself—the basis of our civilization. Roughly 200 to 250 million years ago, when mammals took their great left turn away from reptiles, evolution took the more primitive cortical arrangement and elaborated it into something more complex, more expansive, more geometric: the six-layered neocortical architecture. This was the hyperdrive. Surface area expanded geometrically; cortical columns multiplied computing power while keeping the skull relatively compact. Any bigger and, among other inconveniences, we're probably not getting out of the birth canal. So evolution needed an engineering solution of spatial ingenuity, and *folding* was the answer: cram more tissue, more circuitry, more *possibility* into the same volume.

This is our own natural *GPT*—*general-purpose technology*, par excellence. The neocortex ingests the flood of sensory stimulation—sound, shape, movement, language, memory—and synthesizes it, almost symphonically, into higher-order cognition, abstraction, long-range planning, and culture. *If our intelligence is the creator of our civilization, then the neocortex is the creator of our intelligence*. Memory, long-term planning, language, abstraction, civilization itself—all emanate from this glorious cortical architecture.

And it's astonishingly recent. Life on earth appears roughly 3.5 to 3.8 billion years ago. Multicellular life doesn't really get going until something like 600 million years ago. Complex nervous systems and

the first neurons arrive roughly 500 million years ago. Our lineage diverges from that of the chimpanzees perhaps 6 to 7 million years ago; the first *hominins* show up around 6 million years ago; the genus *Homo* appears around 2.5 million years ago. The broad mammalian neocortical story is older, of course, but the fully elaborated human version—the thing that gives us the bounty outlined above (hierarchy of abstraction, symbolic thought, *art*) is stunningly young. The expanded prefrontal real estate that took us from subsistence and survival to Elon and Demis is on the order of only a few hundred thousand years old. That's it. Four billion years of evolution, and then, in the final nanosecond, the thing that changes the story entirely.

There is a line—I've seen variants of it in enough places that I hesitate to attribute it cleanly, but I think Ray Kurzweil probably came up with it—that the neocortex appears on the last page of the four-billion-year story of life and then proceeds to *rewrite the whole story*.²⁴⁶ Attribution aside, the intuition is exactly right. The neocortex is the late-arriving main character who changes the entire plot. And now here we are, contemplating evolving it again—but not at a geologic pace. Not over hundreds of thousands or millions of years. At the speed of engineering.

That is why BCI matters. It is not some cute side quest (to channel Fidji Simo's OpenAI memo)²⁴⁷ or sci-fi indulgence. It's a direct intervention into the most differentiating biological architecture in our history. If the neocortex made *Homo sapiens*, and *Homo sapiens* made civilization, then to augment the neocortex is to touch the fountainhead itself.

We Are Already Cybernetic

Elon talks about this with some humor (and yes, he's a funny guy actually): we're *already* androids, already cybernetic,²⁴⁸ just with a pitifully low-bandwidth interface. Our cellphone is already a prosthetic limb. Which of us wouldn't turn the car around and drive back home if we realized we left it on the counter? The device is already part of our memory, our sociality, our orientation, our navigation, our agency. And now that the phone itself has become muscular with AI intelligence—Gemini, Claude, ChatGPT, even the regrettable Siri, once great, then senescent, now suddenly youthful and smart again with a Gemini brain transplant—our prosthesis has become cognitive, not merely communicative.

So it isn't really so alien a concept, remonstrate though some might, that we're already cyborgs. BCI is just the next level. I'm not naïve: many healthy people are not exactly lining up to get a craniotomy. But technologies don't remain in their first, clunky, civil-war-battlefield-surgery form. They move from invasive to less invasive, from barbarous to elegant, from clinical necessity to voluntary enhancement. We'll undoubtedly figure out less medieval ways to fuse silicon and carbon. And we'll use AI itself—the intelligence on the other side of the contest—to help us design the fusion.

There is, floating behind all this, the Ship of Theseus problem. At what point, as you lace carbon with silicon and silicon with carbon, does the thing *remain* human? But perhaps that's the wrong question. Perhaps the more urgent one is whether we can afford *not* to hybridize, if the alternative is standing still while something smarter races past us.

This, in the broad sense, is cybernetics. Norbert Wiener saw some version of this presciently in the late 1940s: the line between animal and machine, feedback and control, carbon and circuitry, wouldn't remain as hard as we once imagined.²⁴⁹ The word itself comes from the Greek *kybernētēs*

²⁴⁶ TED: "The hierarchy in your brain: Ray Kurzweil at TED2014" 2014 (Link)

²⁴⁷ OpenAI – Fidji Simo: "AI as the greatest source of empowerment for all" published on July 21, 2025 (Link)

²⁴⁸ Code Conference – Elon Musk: "We're Already Cyborgs" 2016 (Link)

²⁴⁹ NTNU: "What is Cybernetics" (Link)

—the *steersman*, the *governor*. It was always about systems that steer, guide, and regulate themselves through feedback. In that sense, we've been on this road for a long time. BCI is simply where the road gets serious.

AGI Moves at Software Speed. BCI Research Doesn't.

And here is the problem in one sentence: one side of this race is moving at silicon cadence, and the other at medical-device cadence.

I'm *not* calling for cartoonish deregulation, in which charlatans start drilling holes in skulls in strip malls. We don't want neophytes doing craniotomies. I am simply juxtaposing the Wild West of frontier-AI research with the containment, caution, and procedural quagmire of BCI research and regulation. AGI research is proceeding at Silicon Valley speed. BCI is proceeding through IDEs (investigational device exemptions), feasibility studies, neurosurgical caution, and the full choreography of modern device regulation. In a word, *slowly*.

That asymmetry matters. We are massively *laissez-faire* and deregulatory in AI, but massively dirigiste and interventionist in BCI. By all means, keep trying to keep the emergent silicon intelligence in check with alignment, superalignment, constitutional AI, interpretability, evals, tripwires, red-teaming, model auditing, constitutional overlays, and every other algorithmic safety mechanisms Ilya Sutskever can think of. We *should* pursue a multitude of mitigations simultaneously. But we're impairing ourselves if we leave one side of the race almost wholly unconstrained and the other side bound hand and foot in a procedural morass.

And let me disambiguate, because this always gets overheated. I am not calling for some fantastical federal office staffed by \$100,000-a-year bureaucrats to regulate the engineering advances of \$100 million-a-year frontier researchers. That's not serious. I'm not arguing that a salaried functionary will keep up with (let alone out-innovate) the top 0.001 percent of the global engineering class. I *am* saying something narrower and, I think, more reasonable: the regulatory burden on BCI research should be materially re-examined in light of the strategic environment we're in.

We need some enlightened judgment here. Speed up the research. Lower the bureaucratic burden where it can be lowered without inviting barbarism. Invite massive capitalization and risk capital to underwrite BCI research at something closer to the speed of AGI. Use AI itself to accelerate the science of brain interface, signal decoding, materials, targeting, and safety. If this sounds a little autocratic—good. Time to embrace your inner autocrat, at least enough to distinguish between prudent oversight and sleepwalking.

Why Healthcare Must Go First

Healthcare, specifically, should be the proving ground. In the near term, BCI belongs safely in the prosthetic-repair zone: paralysis, locked-in syndrome, ALS, speech restoration, refractory depression, chronic pain, perhaps eventually curing blindness (thank you Elon!) and sensory repair. That is where the moral legitimacy is strongest, the clinical need clearest, and the benefit most immediate. But that same therapeutic platform is also where carbon and silicon intelligence will begin learning to co-evolve.

This is the bridge the chapter originally touched only lightly, and it matters a lot. The path isn't "stick a chip into every healthy twenty-five-year-old tomorrow so they can think faster than Claude." That's Iain Banks fantasy²⁵⁰ (for now). The path is repair, then restoration, *then* augmentation. It is how

²⁵⁰ Nautil: "Will This "Neural Lace" Brain Implant Help Us Compete with AI?" published on April 3, 2018 ([Link](#))

almost every serious technological transformation enters medicine. First you restore lost function. Then you improve the quality and bandwidth of restored function. Then, quietly but inevitably, you cross the line into gain of function.

And that is exactly why the Healthcare 150 matter. They sit at the point where legitimacy, reimbursement, clinical ethics, trials, institutional trust, and technical deployment all meet. For them, BCI isn't just a clinical opportunity. It's a strategic necessity. Last year I introduced a few of you to the brilliant and mission-driven President of Neuralink, DJ Seo. We need more interactions like this: a *collision* (followed by *collusion*) between incumbents (the 150) and insurgents (DJ and his ilk). The establishment can then *co-design* with the frontier firms, help shape the governance norms, build the care pathways, define the reimbursement logic, and create the first serious institutional landing zones. Those who don't *co-design* will be forced to inherit whatever symbiosis is imposed by others.

Dispatches from the Frontier

The progress is no longer theoretical. Neuralink, Paradromics, Synchron, Precision Neuroscience, UC Davis, Carnegie Mellon—this isn't side-science anymore. Neuralink's first participant, Noland Arbaugh, paralyzed from the shoulders down, made the point emotionally vivid: gaming, browsing, design work, communication, digital independence, all through thought.²⁵¹ Paradromics is pushing a different bandwidth strategy.²⁵² Synchron is exploring less invasive endovascular routes.²⁵³ Precision is trying to reduce surgical burden.²⁵⁴ The UC Davis and CMU teams are showing that the field is broader than any one company and that thought-to-speech, thought-to-motion, and thought-to-robotic-action are all moving from speculation toward implementation.²⁵⁵

Wearing my Thrive Capital Venture Partner hat, we're enthusiastic supporters (and investors) in Neuralink. Neuralink's own story has also notably evolved over the last year. It's no longer just Noland Arbaugh proving that a cursor can move by intention. By January 2026, Neuralink had 21 participants enrolled worldwide, already using implants to control digital and physical tools through thought—something approaching, to my layman's eyes, *telekinesis*. That matters because the system is being tested across more bodies, more anatomies, more conditions, and more use-cases: spinal-cord injury, ALS, cursor control, communication, art, environmental control, and the early bridge into (Optimus) robotic limbs. The underlying device is still the N1/Telepathy architecture—an implanted, wireless, high-channel interface with roughly 1,024 electrodes distributed across thin cortical threads²⁵⁶—but the important shift is clinical density. It's becoming less like a one-patient demonstration and more like a repeatable assistive-computing pathway.

The surgical side has moved too. This isn't yet a no-craniotomy story; a surgeon still creates the opening in the skull. But Neuralink's newer robot is taking over the part that human hands can't plausibly scale: picking up hair-thin flexible threads from the implant and inserting them into cortex with extreme precision. The newest system has a smaller implant arm, five-axis reach, sensors, eight cameras, and optical coherence tomography, and Neuralink says it has already completed a successful surgery with it. That matters because the first patient exposed the key engineering problem—thread retraction and signal loss—and the company's response has been procedural, postoperative, robotic, and hardware-level: modified technique, higher signal quality in 18 of the next 20

²⁵¹ Neuralink: "PRIME Study Progress Update — User Experience" published on May 8, 2024 (Link)

²⁵² Paradromics: "Augmentative and alternative communication BCIs" published on May 10, 2022 (Link)

²⁵³ Synchron: "A new field of medicine is here: Functional Neurointervention" (Link)

²⁵⁴ Precision: "A New Way to Interface with the Brain" (Link)

²⁵⁵ UC Davis Health: "New brain-computer interface allows man with ALS to 'speak' again" published on August 14, 2024 (Link)

²⁵⁶ Neuralink: "PRIME Study Progress Update" published on April 12, 2024 (Link)

participants, a push from roughly 1,000 toward 3,000 electrodes, and exploration of placing threads directly through the dura to reduce surgical burden. In other words, Neuralink is not only making the decoder smarter. It's trying to make the whole implant procedure more durable, more repeatable, and eventually more scalable.

The significance of all this isn't simply that we can help patients with paralysis—magnificent though that is. It's that we are beginning to *learn the grammar* of the interface itself. We're learning what counts as usable bandwidth, what kinds of cognition can be *externalized* and *re-internalized*, what kinds of agency can move through the bridge, what kinds of restoration can become augmentation. That's why I think we should be paying much more attention to the first dozen or so "enhanced" humans. They're not curiosities. *They're advance scouts.*

BCI as Deliverance

So let me say it bluntly. I think BCI is the strongest candidate we have for deliverance from the *mechanistic-interpretability, corrigibility* and *superintelligence alignment* conundrum. Not salvation in the cartoon sense. Not a guarantee of benevolence. Not some seamless Kurzweilian paradise in which humanity uploads to the cloud and all ends well. But deliverance in the *strategic* sense: the only plausible path toward keeping the human neocortex *inside* the loop of superintelligence rather than permanently downstream of it.

That's the strongest version of the chapter's claim. Not that BCI makes us omnipotent. Not that it makes us safe. But that it may be the last credible line of defense against a world of malevolent, out-of-the-box, or simply incomprehensible (or incorrigible) machine intelligence. Eventually, we must figure out a way to merge—cybernetically, strategically, cognitively. We must fuse silicon and carbon. We must *expand* our own capacities rather than merely beg for mercy from the thing we have built.

Because in the end, I don't want a world with superintelligent machines and biologically unevolved humans standing naked before them. I want a world in which AGI becomes an *extension* of our intelligence rather than its replacement. A co-evolution. A hybridization. A last serious shot on goal before the map runs out.

That is why I am willing to sound a little Kurzweilian, a little Muskian, maybe even a little Asimovian. If this intuition is right, BCI is not an eccentric side quest. It is one of the central civilizational projects of the age. And if we continue to move on it at the speed of conventional medicine while AGI advances at the speed of software, then we will have nobody to blame but ourselves. This isn't merely enhancement fantasy. It's strategic realism.

If I'm wrong here, we will at least have built extraordinary tools to restore speech, movement, sight, agency, and dignity to people who desperately need them. And if I'm right, we may have bought ourselves something much rarer: a fighting chance.

The Chapter, Compressed

Here is the chapter, compressed into the governing takeaways.

First, BCI is not an eccentric sci-fi side quest. It is the hardware counter-revolution: a response to the possibility that software intelligence may outrun the biological interface through which humans currently govern it.

Second, the problem is bandwidth. AGI moves at software speed; BCI research moves at

medical-device speed; and the human neocortex, magnificent though it is, still communicates with the machine through a pitifully narrow channel.

Third, the neocortex is our natural general-purpose technology: the civilizational fountainhead of abstraction, planning, language, memory, moral reasoning, and tool-building. If AI upgrades cognition outside the skull, BCI asks whether we can upgrade the interface to the skull itself.

Fourth, the path should begin therapeutically: paralysis, locked-in syndrome, ALS, speech restoration, refractory depression, chronic pain, neurodegeneration, and other sacred repair work where the moral case is obvious.

Fifth, healthcare must go first because legitimacy, reimbursement, ethics, trials, patient trust, and technical deployment all meet inside the 150. If they do not co-design the norms, they will inherit norms built elsewhere.

Sixth, BCI may be deliverance not because it solves alignment magically, but because it keeps human agency, comprehension, and co-decision-making closer to the loop as synthetic intelligence scales.

BEFORE WE TURN THE PAGE

BCI leaves us with the largest possible question: what if the human mind itself must change? But the essay now has to gather its threads. The coda returns to the moral synthesis: build the machine, care for the people, and make the surplus worthy of the human beings it is supposed to serve.

Coda

Build the Machine. Care for the People.

"The real problem of humanity is the following: we have Paleolithic emotions, medieval institutions and godlike technology."

— E. O. WILSON, 2009

A WORD ON NAVIGATING THIS CHAPTER

The coda compresses the argument into a covenant: build the machine, but care for the people. It's my moral synthesis of the essay and the final call to the Healthcare 150 and the AI 10.

I want this closing chapter to do more than summarize. It has to (try to) gather the whole strange essay back into one obligation. The preceding chapters have wandered through science, labor, medicine, payers, hospitals, behavioral health, China, theology, and brain-computer interface because AI refuses to remain a tool category. But the moral through-line is actually simple: once intelligence becomes abundant, the central scarcity becomes wisdom, governance, courage, and mercy.

AI is the industrialization of intelligence. Healthcare is the most consequential and sacred site of its installation. And the moral test is whether we use this technology to create abundance with dignity, or extraction with better software.

That's the whole thing, really.

Everything else has been an attempt to make that sentence believable enough, concrete enough, and operational enough that serious people can no longer safely file it under futurism, innovation, or one more technology cycle that healthcare can politely admire from a distance while waiting for Epic, CMS, the medical staff, the lawyers, and the inevitable governance committee to domesticate it into harmlessness. This time the old choreography won't work. The technology isn't merely arriving at the edge of healthcare. It's arriving at the level of the substrate: cognition, coordination, documentation, interpretation, diagnosis, synthesis, administration, discovery, and the whole overgrown apparatus by which American healthcare metabolizes itself.

That's why the chapters had to begin upstream. You and I couldn't responsibly have a conversation about the universal doctor, AI labor convulsions, payer bots, hospital labor, revenue cycle, behavioral companions, geopolitics, BCI, or clinical liability without first asking what the thing is. And the answer, to my mind, is still the essential one: this isn't merely another machine in the long procession of machines. It's a multiplication of the very faculty that made the machines. The steam engine amplified muscle. Electricity amplified energy. The microprocessor amplified calculation. GenAI multiplies intelligence itself.

No wonder everything feels unstable. I feel it. And you feel it too.

If intelligence becomes abundant, the scientific frontier changes. Hypothesis generation, long the sacred preserve of biological minds, begins to include synthetic participants. The old Baconian method doesn't disappear, but its monopoly ends. Science becomes more recursive, more

computational, more simulated, more opaque, more generative. We may know more than we understand. We may verify truths we can't fully narrate. Biology, that most complex and fecund of domains, may yield to virtual biologists, robotic labs, multimodal models, and search processes no human scientist could traverse unaided. If Dario is even partially right about a compressed 21st century, then medicine may soon experience progress that feels less incremental than phantasmagoric: new mechanisms, new molecules, new diagnostics, new interventions, new ways of reading the body at machine speed.

And yet knowledge without wisdom is not enough. Asimov had that right. The machine may gather knowledge faster than society gathers wisdom; indeed, that may be the defining asymmetry of our moment. Which is why the next question isn't merely epistemological but spiritual. What happens when the strongest thing in the room is no longer human? What happens when the model becomes not only useful but intimate, authoritative, companionate, persuasive, and, in some domains, superior? We will anthropomorphize it. We will confide in it. We will obey it. Some will worship it. Others will try to destroy it. Most of us will do something more ordinary and more dangerous: we will slowly reorganize our lives around it without quite admitting that we have done so.

That's why the question of human agency runs through the whole essay. BCI may be one answer, or at least one counter-move: not cognitive surrender, but cognitive upgrade; not the passive offloading of our thinking into opaque machines, but the attempt to keep the human neocortex inside the loop of superintelligence. That sounds science-fictiony until it doesn't. The first BCI era is therapeutic, and that work is sacred: paralysis, locked-in syndrome, refractory depression, chronic pain, neurodegeneration. But the second era will be augmentative, and the third may be civilizational. If we are going to live alongside synthetic cognition that increasingly exceeds us, then the hardware question becomes inseparable from the dignity question. Do we remain authors of the future, or merely biological validators of machine output?

Then comes labor, the great social wound of the transition. The old technological bargain was cruel but usually regenerative: destruction now, reabsorption later. The loom eviscerated the weaver; the automobile annihilated the horse-powered economy; the computer changed clerical work; the internet vaporized some categories and created others. The optimistic story rested on one morally load-bearing assumption: humans could climb the ladder. Machines took muscle, and humans sold supervision. Machines took calculation, and humans sold analysis. Machines took filing and retrieval, and humans sold interpretation. But GenAI climbs the ladder too. It attacks not only tasks, but the cognitive scarcity that made reabsorption plausible.

That's the category break.

And healthcare sits directly in the blast radius because healthcare is labor-addicted, productivity-starved, administratively overgrown, biologically complex, data-superabundant, politically defended, and morally indispensable. It is the largest intimate industry in the country. It's also one of America's great post-manufacturing labor engines, a regional anchor, a middle-class employment scaffold, and a place where the abstractions of labor economics become mortgages, commutes, nursing shifts, union meetings, tax bases, community colleges, and households. Reducing labor intensity in healthcare isn't merely an EBITDA bridge. It's a civic event of incredible consequence.

That's why the denominator matters. AI in healthcare isn't an IT-budget question. It isn't the next pleasant interface sitting on top of Epic. It isn't one more procurement category for an innovation office to admire. The relevant denominator is labor: salaries, wages, benefits, contract labor, outsourced services, management layers, call centers, documentation teams, coding queues,

prior-auth teams, utilization management, care coordination, and the whole salaried cognitive apparatus we've built around care. Once you change the denominator from software spend to labor architecture, every strategic conversation changes.

But the point cannot be labor reduction for margin. That's the small, ugly, sociopathic version of the story. The serious version is labor reconstitution for abundance: fewer humans trapped in adminivia, more humans doing human work, more care delivered at lower cost, more expertise available to more people, and a generous transition compact for workers whose old roles were artifacts of a system we should have been ashamed of long before the machine arrived.

This is the narrow ridge the 150 must walk. Move too slowly, and healthcare preserves an unaffordable status quo while calling inertia compassion. Move too bloodlessly, and healthcare turns the most powerful productivity technology in its history into a margin-recapture machine with a better press release. The point isn't to choose between patients and workers. The point is to build a system in which patients get affordability and workers get honesty, generosity, and a real path into the next labor architecture.

That means we have to tell the truth. The Orwellian conspiracy of silence around AI job loss won't work. Workers are smart. They can see the tools improving. They can see what's happening in software, media, consulting, finance, law, and professional services. They will know when "augmentation" quietly becomes substitution, when "productivity" means fewer people, when "agentification" means digital labor, and when "efficiency" means the old employment covenant is breaking. The humane approach isn't euphemism. It's candor plus a solution set.

Tell the truth. Map the work, not the titles. Stop hiring reflexively into exposed workflows. Build the agentic capacity. Redeploy humans toward the bottlenecks. Share the winnings. Govern the machine like a civic institution rather than a procurement department with a press release.

Those seven verbs are still, to my mind, the operating prescription: tell, map, stop, build, redeploy, share, govern.

Hospitals and health systems, especially, need to understand themselves differently. They aren't merely care-delivery organizations. At sufficient scale, they are quasi-sovereign institutions: employers, landowners, financiers, political actors, training grounds, data stewards, charitable entities, scientific platforms, civic anchors, and, in many places, the dominant institutional fact of local life after government itself. Which is why the CEO chapter had to be written as a memo on statecraft. AI won't transform health systems. CEOs will. The question isn't who invents the model. That race is elsewhere. The question is who diffuses it into the living organism of the enterprise.

The winners won't be the systems that run the most pilots. Healthcare has enough pilot museums. The winners will be the systems that build diffusion muscle: one governed god model, broadly deployed; a ministry of education; forward-deployed teachers and engineers; a GDPval-like task map; tight-loose-tight governance; true ontology and machine-legible data; an operating-company posture rather than a SINO federation; and a willingness to stop preserving every inherited workflow simply because no one has had the courage to interrogate it.

And the payers shouldn't relax. Their decade of ascendancy was real, but it was built too often on opacity, administrative scale, information asymmetry, actuarial exclusivity, and friction as strategy. AI is unusually hostile to all five. It arms the rebels. Providers get better at documentation, coding, evidence retrieval, and appeals. Consumers get agents that can read policies, compare plans, contest denials, and navigate the maze. Employers get AI-native TPAs and better rented intelligence. Regulators push toward machine-readable rules. The paperwork moat melts. The payer does not

disappear, but it thins. The AI-thin utility may be useful. It may even be necessary. But the imperial payer built on delay, opacity, and procedural exhaustion is not the end state. It's the thing being deselected.

Behavioral health, meanwhile, may be the first great clinical unlock, and not because it's peripheral, soft, or merely compassionate. Quite the opposite. Behavioral is the synchronization layer of total cost of care. Depression, anxiety, addiction, loneliness, shame, nonadherence, executive dysfunction, sleep, diet, medication persistence, self-care—these aren't marginal human inconveniences. They're the central determinants of whether medicine works. And here the ISA, the intelligent social agent, has a strange and beautiful advantage: humans lie to humans, but tell the truth to machines. Nonjudgment, memory, persistent availability, and the felt texture of being heard may become clinically and economically transformative. The detractor must defend the status quo: shortage areas, delayed care, no child psychiatrist in entire counties, 2:16 a.m. despair met by scarcity. The burden is shifting.

Clinical AI will be harder, and more profound. The universal doctor won't cinematically arrive all at once, and it won't dissolve the need for human physicians, nurses, pharmacists, therapists, or other healers. But it will unbundle medicine. Diagnosis, synthesis, triage, guideline application, medication reconciliation, second opinions, risk stratification, longitudinal memory, patient education, prior-auth support, inbox triage, and clinical-gap closure all become increasingly machine-addressable. The physician remains indispensable, but the bundle called "doctor" gets repriced and reorganized. The standard of care will eventually invert. Today the question is: why did you use the AI? Tomorrow, in some domains, it will be: why did you fail to use the AI?

That makes liability the unlock. We need safe harbors, validation envelopes, indemnification, post-market monitoring, and a mature doctrine of responsibility. The standard cannot be machine infallibility. Human medicine has never met that standard. The standard should be human equivalence or human superiority inside a defined use case, with appropriate governance. We must stop holding technology to an impossible standard while tolerating massive human variation, delay, diagnostic error, and preventable harm from the current system simply because the current system arrives wrapped in familiar institutional clothing.

And then there's the world beyond us. China, the Gulf, and other installation-forward regimes won't wait for American procedural comfort. The United States may still be the greatest invention engine on earth, but invention isn't enough. Diffusion is the prize. The free doctor—the diagnostically competent, multilingual, marginal-cost medical intelligence available to the Global South—will be a geopolitical instrument. It will carry values, defaults, governance assumptions, and institutional allegiances. If America builds the medical superintelligence but another civilization installs it faster, then we may win the invention race and lose the distribution race. That would be a very American tragedy: Prometheus with a permitting problem.

So the remedy isn't xenophobia, not techno-nationalist panic, not the self-flattering fantasy that we can export-control our way to victory while failing to build power, fabs, data centers, transmission, nuclear, and institutional diffusion capacity at home. The remedy is pragmatic seriousness. Build. Compete. Cooperate where possible. Lessen our dependence on Taiwan expeditiously without converting the whole relationship into fatalistic hostility. Relearn how to install. Export trustworthy medical intelligence. Let American values ride inside the tools, not merely inside speeches about the tools.

All of this brings us back to the covenant. The machine will produce surplus. It already is. More

surplus is coming: labor savings, scientific acceleration, clinical augmentation, administrative compression, lower inference costs, better access, cheaper expertise, more longitudinal care. The question is where the winnings go. If they go only to margins, share prices, executive scorecards, and private-equity exit multiples, the politics will curdle, and deservedly so. If the surplus becomes lower prices, more access, better care, humane transition, worker redeployment, clinical time, behavioral support, and a more dignified patient experience, then this revolution becomes defensible.

Maybe even beautiful.

I remain, despite everything, an optimist. Not a naïve optimist. Not a TED-talk optimist. Not the grinning, frictionless sort of optimist who treats every casualty of progress as an acceptable footnote to someone else's capitalization table. I'm optimistic because the status quo is so inadequate, because the suffering is so abundant, because our systems are so wasteful, because clinical knowledge is so unevenly distributed, because patients are so often humiliated by the bureaucracy we have built around their fear, and because the machine gives us, perhaps for the first time in a generation, a plausible path toward abundance.

More care. Better care. Cheaper care. More science. Less suffering. More behavioral support. Less loneliness. More physician time with patients. Less documentation. More early detection. Less avoidable catastrophe. More knowledge at the edge. Less geography as destiny. More humane work. Less spiritual death by work queue.

That's worth defending. But only if we defend it correctly.

The adult task is synthesis. Not denial. Not deification. Not Luddism. Not sociopathy. Build the machine, but govern it. Use the surplus, but share it. Compress the labor denominator, but care for the worker. Accelerate science, but preserve wisdom. Deploy clinical AI, but assume responsibility. Create the free doctor, but embed dignity. Let the model know more than we can understand, but do not let it decide what matters. That remains our burden. That remains our dignity.

Move like builders. Govern like adults. Tell the truth like people who respect their workers. Share the winnings like institutions that understand their civic footprint. And never forget that the point of all this intelligence, if the word still means anything, is not fewer employees, higher margins, cleaner decks, or a more elegant operating model. It is healing. It is mercy. It is the reduction of suffering. It is the restoration of human attention to the places where human attention still matters most.

That's the version of the future worth fighting for. Not a healthcare system with fewer humans because financiers discovered a cheaper denominator, but a healthcare system with fewer humans trapped in work that should never have required a soul in the first place. Not a medicine that hides behind procedural pieties while patients wait, but a medicine that moves with disciplined urgency when the evidence, the model, and the moral case align. Not a society that worships the machine, but one that uses the machine to make human life more capacious, less lonely, less expensive, less bureaucratically humiliated, and more merciful.

The healthcare 150 and the AI 10 can't control everything that follows. Nobody can. But they can decide whether they will be passive incumbents or covenantal builders. They can decide whether the surplus becomes extraction or abundance. They can decide whether labor transition is handled with euphemism or truth. They can decide whether clinical AI is delayed into irrelevance or governed into care. They can decide whether healthcare becomes the last sector to metabolize intelligence or the first sector to prove that intelligence can be industrialized without losing the human being at the center of the work.

The spirits have been summoned. They won't be unsummoned. The question now isn't whether the magic exists. It does. The question is whether we have the courage, discipline, humility, and moral imagination to build a covenant worthy of it.

The Coda, Compressed

Here is the coda, compressed into the final covenant.

First, the machine should be built because the status quo is not morally innocent. Healthcare is too expensive, fragmented, bureaucratic, inaccessible, and exhausting to treat inertia as compassion.

Second, the people must be cared for because technological surplus does not distribute itself justly. Workers deserve candor, transition, generosity, and a credible path into the next system.

Third, the winnings must become visible. Lower cost, better access, more behavioral support, more physician time, fewer denials, less documentation, and more humane care are the moral proof of this revolution.

Fourth, the 150 and the AI 10 have to co-design the future. The labs have the models; healthcare has the patients, trust, workflows, liability, data, and sacred touch. Neither tribe is sufficient alone.

Fifth, America has to install as well as invent. Power, fabs, data centers, clinical validation, liability architecture, and institutional diffusion are not side quests. They are the substrate of sovereignty in the age of medical intelligence.

Sixth, the adult posture is synthesis: no panic, no worship, no denial, no bloodless efficiency theater. Build the machine, govern it, share the surplus, preserve the human, and make the revolution worthy of the people it is supposed to serve.

Build the machine. Care for the people.

That's the only version of this revolution worth defending.

Disclaimer.

These materials are for informational purposes only and do not constitute and should not be construed as an offer to sell or solicitation of an offer to buy any securities or related financial instruments in any jurisdiction in which such offer or solicitation, purchase or sale would be unlawful under the securities, insurance or other laws of such jurisdiction. Any such solicitation will only be made through formal offering materials that will include, among other disclosures, numerous risk factors relating to private fund investments.

TowerBrook cannot guarantee the accuracy or completeness of information contained herein and has not independently verified the accuracy or completeness of such information or the assumptions on which such information is based. No representation, warranty or undertaking, express or implied, is given by TowerBrook, its officers, partners, employees, affiliates, directors, shareholders or agents (each a "TowerBrook Person") as to the accuracy or completeness of the information or opinions contained in these materials, and TowerBrook expressly disclaims liability therefor. Past performance is NOT indicative of future results, and TowerBrook makes no representation in this regard. Nothing contained herein is intended to predict the performance of any investment.

These materials may contain forward-looking statements, projections, targets, forecasts and estimates (collectively, "Projections"). Projections are hypothetical and are for illustrative purposes only, and no assumptions or comparisons should be made based upon this information. Projections are necessarily speculative in nature, and it can be expected that some or all of the assumptions underlying the Projections may not materialize and/or that actual events and consequences may vary significantly from the assumptions upon which projections contained herein have been based. Actual events often depend upon factors that are beyond the control of TowerBrook. Future events may differ or results or actual performance could differ materially and adversely from those expressly or implicitly reflected, or contemplated in Projections, and hypothetical projections are subject to inherent limitations.

CONTACT

London

1 St. James's Market
Carlton Street
London SW1Y 4AH
United Kingdom

Tel: +44 20 7451 2020

New York

Park Avenue Tower
65 East 55th Street
New York, NY 10022
United States of America

Tel: +1 212 699 2200

Madrid

Calle Almagro 36
Escalera izquierda, 3ª planta
28010 Madrid
Spain

Tel: +34 91 954 2361

Paris

4 rue Royale
75008 Paris
France

Tel: +33 1 89 96 11 12

Milan

Via San Raffaele 1
20121 Milan
Italy

Tel: +39 02 1241 27600

Greenwich

41 West Putnam Avenue
Greenwich, CT 06830
United States of America

Tel: +1 929 486 2300