

BEYOND
FASTER
HORSES

CHAPTER THREE

Assembly lines were the real innovation

The engine was the marvel, but it was the assembly line that changed the world. Clinical AI needs its assembly line.

AUTHOR

Tahseen Omar

RESEARCHER

Dorothea Dwomoh

FEATURING

Aman Bhasin
Saul Howard
Gautam M. Shah
Rob Duffy

READING TIME

9 min read



SEPTEMBER 2026

[ANTERIOR.COM/BEYOND](https://anterior.com/beyond)



In the fall of 1806, two brothers pushed a flat-bottomed boat upstream on the Saône in eastern France using nothing but a series of controlled explosions. Claude and Nicéphore Niépce had spent years perfecting their machine, which worked by igniting a fine dust of clubmoss spores roughly a dozen times a minute. They called it the *Pyréolophore*.¹ It was the first internal combustion engine.^{1*}

It was also a far cry from the automobile. So, in 1863, Étienne Lenoir mounted his gas engine on a cart and drove from Paris to Joinville-le-Pont and back—eighteen kilometers in about three hours, barely faster than walking.² Two decades later, Gottlieb Daimler and Wilhelm Maybach solved the power-to-weight problem with a small, fast-running engine, then bolted it into a horse carriage. The retrofit demonstrated what was possible, but carriage wheels, springs, and steering had not been designed around an engine's vibration.³ No one sold their horses.

Karl Benz's genius in 1886, on the other hand, was the engineering *around* the engine: a purpose-built tubular-steel chassis, a differential, evaporative cooling, electric

ignition—assembled from an industrial base of precision tubes, chains, and wire-spoked wheels that the bicycle trade had only just made available.⁴ In August 1888, his wife Bertha took the *Motorwagen* (without telling him) and drove their two sons more than a hundred kilometers from Mannheim to Pforzheim, buying fuel at a pharmacy, clearing a blocked fuel line with her hatpin, and having a cobbler nail leather onto the failing brake blocks, inventing the brake lining somewhere along the way. It was the first long-distance drive in an automobile.⁵

But for all its ingenuity, the car still didn't change the world until after 1913 when, on the factory floor at Highland Park, Michigan, Ford's engineers tied a rope to a Model T chassis and dragged it past lines of parts bins with a windlass, assemblers walking alongside, someone holding a stopwatch. They'd just invented the first moving assembly line. After months of iteration and engineering, the labor in a chassis fell from twelve and a half hours to ninety-three minutes, and the price of a Model T fell from \$825 to as little as \$260 by 1925, even as quality improved.⁶

1 Nicéphore Niépce's House Museum, "The *Pyréolophore*: A New Engine Principle"; Musée Nicéphore Niépce, "Biographie de Nicéphore Niépce"; Archives de l'INPI, dossier 1BA395, brevet d'invention of Joseph-Nicéphore and Claude Niépce, "*Pyréolophore, machine dont le principe moteur est l'air dilaté par le feu, propre à mettre en mouvement toutes sortes de mécaniques*" (filed November 12, 1806; patent certificate issued April 3, 1807).

1* The *Pyréolophore*'s claim is contested. It was not the first internal combustion engine to be patented: Robert Street had patented the concept in England in 1794 (*The Repertory of Arts and Manufactures* 1 [1794]: 154–56), and François Isaac de Rivaz, a Swiss inventor, received a French patent for his hydrogen-fueled engine, about two months before the Niépces received theirs, and described mounting the engine on a cart (brevet no. 781, January 30, 1807; *Description des machines et procédés spécifiés dans les brevets d'invention* 9 [1824]: 182).

2 Musée des Arts et Métiers, *Le moteur Lenoir*, Les Carnets (Paris: Musée des Arts et Métiers, n.d.), 5 (the eighteen-kilometer round trip from Lenoir's Paris workshop to Joinville-le-Pont and back in 1863, completed in three hours).

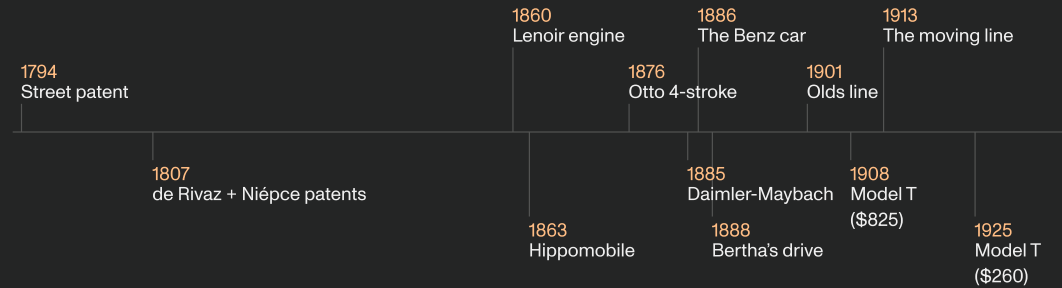
3 Mercedes-Benz Group, "Beginnings of the Automobile: The Predecessor Companies (1886–1920)"; German Patent and Trade Mark Office, "Wilhelm Maybach," Milestones; Mercedes-Benz Store, "Daimler Motor Carriage (1886)."

4 Mercedes-Benz Group, "Benz Patent Motor Car: The First Automobile (1885–1886)." Also Deutsches Museum, "Der *Motorwagen* von Carl Benz," Meisterwerke (page archived November 13, 2015).

5 Mercedes-Benz Group, "Bertha Benz"; German Patent and Trade Mark Office, "Bertha Benz," Ingenious Women.

6 Stewart Brand, "Three Maintenance Philosophies Fought for Control of the Auto Industry," in *Maintenance: Of Everything*, chap. 2, "Vehicles (and Weapons)," eighth draft, Books in Progress (Works in Progress, August 2024). Also Henry Ford with Samuel Crowther, *My Life and Work* (Doubleday, Page, 1922), chap. 5; The Henry Ford, "1909 Ford Model T Touring Car Drawn to Scale, November 27, 1972," object ID 84.1.1660.P.D.964; The Henry Ford, "1925 Ford Advertisement, 'Supreme in Service as in Saving,'" object ID 64.167.19.453.

The long road from engine to car 1794 to 1925

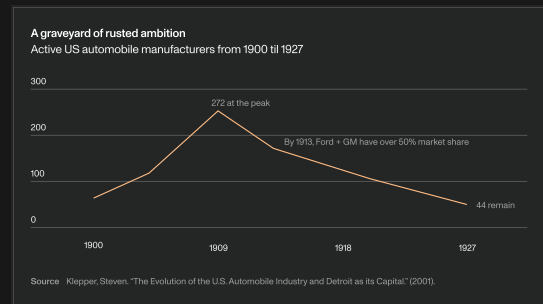


Sources French Patent and Trademark Office (INPI), Britannica, Mercedes Benz Group, Works in Progress

Ultimately, the automobile companies that survived were the ones that invested in building assembly lines. During the industry's first decades, more than seven hundred American companies entered automobile manufacturing.⁷ By 1913, Ford and General Motors already controlled more than half the market⁸ and, by 1927, only forty-four manufacturers remained.⁹ The firms that fell behind did not lose on the quality of their car, but on the ability to produce quality at scale.

The system that builds the system

Consider the combustion engine as the large language model (LLM); the car as the AI agent; and the assembly line as the platform that designs, builds, and continually improves reliable agents at scale.



The models are already capable of clinical reasoning, and agents of completing discrete but complex workflows, even in high-stakes and regulated settings like healthcare. Today, the frontier opportunity is in scaling the value of AI, safely and reliably.

For health plans, the focus must shift away from demos and pilots (engines and cars) and toward infrastructure that scales (assembly lines). Aman Bhasin, the Chief Information Officer of Capital

⁷ Steven Klepper, "The Evolution of the U.S. Automobile Industry and Detroit as Its Capital" (unpublished paper, Carnegie Mellon University, November 2001).

⁸ Richard S. Tedlow, "The Struggle for Dominance in the Automobile Market: The Early Years of Ford and General Motors," *Business and Economic History*, 2nd ser., 17 (1988): 49–62.

⁹ Robert R. Ebert, "Number of U.S. Automakers Falls to Forty-Four," EBSCO Research Starters, 2021.

Blue Cross, is leading that shift inside a health plan today.¹⁰



Aman Bhasin is the Chief Information Officer of Capital Blue Cross.
Photograph: Cody Cutter

“I think about it in the sense of a car. A track where I can go fast because I have predictability... those are the guardrails and the chassis we want to build. Instead of being reckless out on the street, driving a hundred miles an hour where there are a lot of unknowns.”

Aman Bhasin
Chief Information Officer, Capital Blue Cross

When investing in an AI system, health plan leaders must consider scale across six qualities: accuracy and reliability; observability; privacy and security; interoperability; customizability; and cost.

Accuracy and reliability for safety

LLMs are, by design, non-deterministic; reliability is not an intrinsic property—it has to be engineered into the system. After several years building secure systems at scale on Apple's iCloud team, Saul Howard, VP of Engineering at Anterior, now spends most of his time thinking about this challenge.

“You have this LLM—the non-deterministic core—and to make use of that within your system, it's absolutely imperative that every other piece of the system is as deterministic, as predictable, as schema-driven, as auditable as possible,” says Howard. “How do you build the power station around that core that can harness the power of the unpredictability and make it perform within a predictable system?”¹¹

“How do you build the power station around that core that can harness the power of the unpredictability and make it perform within a predictable system?”

Saul Howard VP of Engineering, Anterior

It's also important not to treat software SLAs as a safety standard. When Firestone's Decatur plant went through one troubled production period,

¹⁰ Aman Bhasin, shoot interview, 2026.

¹¹ Saul Howard, shoot interview, 2026.

Scaling AI in healthcare

6 qualities health plan leaders should require in their AI systems



ACCURACY & RELIABILITY FOR SAFETY

Engineer determinism around the non-determinism, with reliability measured against consequences of failure, not traditional software SLAs.



OBSERVABILITY FOR BETTER CARE

Make each decision traceable to its inputs, evidence, and reasoning. Give teams the visibility to investigate failures, audit outcomes, and use what they learn to improve care.



PRIVACY & SECURITY BY DESIGN

Enforce protection through architecture: control ingress/egress, isolate customer environments, and separate control planes and data/compute planes.



INTEROPERABILITY AS A PRECONDITION

Integrate with the systems and standards healthcare already uses. Embrace shared data models to limit fragmented tools and workflows between organizations.



CUSTOMIZABILITY ACROSS THE STACK

Maintain a common quality bar across users while letting each plan encode its own policies, thresholds, workflows and boundaries for human judgment.



COST REDUCTION THROUGH ENGINEERING & OPERATIONS

Make workflows legible, then optimize around them: route intelligence efficiently and focus human judgment where it creates most value.

the same tire became roughly fifteen times more likely to fail. NHTSA logged reports of 271 deaths associated with the tires under investigation.¹² And one degraded Takata propellant formulation forced the recall of some 67 million airbags.¹³ Today, automotive quality is measured in defective parts per million, not percentages, with the best suppliers held to single digits.¹⁴ A fleet of 10,000 cars of which 9,999 are safe is “99.99 percent reliable,” but scaled to a population of a few million, it puts several hundred lives at risk.

“For an engineer, a 99.99% might be fine. That’s a great error rate...four nines,” Howard says. “But for a doctor, four nines means that somebody had a bad outcome in their healthcare journey.” The infrastructure that determines the care humans receive deserves the same engineering seriousness.

Observability for better care

The automotive industry learned to trace a single defective part to every car that carried it. As the Takata recalls widened, any driver in America could look up their VIN and know within seconds whether their vehicle was affected.¹⁵ Clinical AI can hold itself to something similar, with each micro-decision auditable. Every time an agent acts, it leaves an observable trail behind it: every fact extracted, where on the page it came from, each criterion applied, any and all actions taken, and why. The agent’s output is not a “chat message” but an assertion: an independently checkable clinical or operational fact, written with its provenance to an append-only log that can be replayed exactly, years later.

¹² Alan B. Krueger and Alexandre Mas, “Strikes, Scabs and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires,” NBER Working Paper 9524 (National Bureau of Economic Research, March 2003); David R. Francis, “Labor Dispute Caused Poor Quality Products,” *NBER Digest*, August 2003.

¹³ National Highway Traffic Safety Administration, “FAQs: Takata Desiccated Inflators and Volkswagen Recalls,” Takata Recall Spotlight.

¹⁴ Johnson Controls Automotive Experience, *Global Supplier Standards Manual*, chap. 4, “Quality” (July 2016); “The Monthly Metric: Defect Rate,” *Inside Supply Management* (Institute for Supply Management), November 26, 2018.

¹⁵ National Highway Traffic Safety Administration, “Check for Recalls: Vehicle, Car Seat, Tire, Equipment,” recall lookup by VIN.

Half of this is compliance; regulators already require payers to give specific reasons for adverse decisions,¹⁶ and an emerging design literature argues for tamper-evident audit logs of AI recommendations.¹⁷ But the other, more important half is improving care; a decision that can be decomposed is one that can be investigated. “If we are making a denial on a request for coverage, we need to understand why we did it. What was the actual data?” says Bhasin. “When you think about a regulator, when you think about a member, we ought to be able to look back and know exactly why the decision was made.”

Privacy and security by design

In 1854, at the Crystal Palace exhibition in New York, Elisha Otis stood on a hoisted elevator platform and ordered the rope cut. The platform dropped a few inches and stopped; a spring-set brake locked into the notched rails the moment the rope lost tension, and Otis called down to the crowd, “All safe, gentlemen, all safe.” The elevator’s safety did not depend on an operator’s vigilance or an inspector’s checklist; it was designed into the machine.¹⁸

The most sensitive data humans have deserves the same safety by construction, or “zero trust architecture.” For example, an append-only decision ledger that records the shape of the work (which facts, produced by which steps, in what order), rather than the patient data itself, so that engineers and clinicians can debug a workflow without unnecessary access to PHI; strict ingress and egress rules on protected data;

customer environments carefully segregated; separating the system’s control plane from its data and compute plane.

“Security is very different in AI,” Bhasin says. “Instead of having a guard at the gate, at the entry and the exit point...it has to be at every gate, at every step. We have to have security enabled because there’s a lot more ingress and egress points with AI.”

Interoperability as a precondition

No health plan runs on a single system, and the automotive industry had to solve a similar challenge. From the 1910s, the Society of Automobile Engineers—“Automotive” from 1916—collected specifications into shared standards so that a supplier’s part could work in any maker’s machine.¹⁹ Healthcare has its own emerging standard: FHIR (Fast Healthcare Interoperability Resources), a data model all healthcare organizations, not just health plans, can operate in.²⁰ The system also has to sit alongside the existing platforms—headless where needed, integrable everywhere, taking data from where it lives and writing results back without disrupting or fragmenting it—and help lead the industry’s convergence rather than inventing a private ontology.

As Gautam Shah, who led product at Caelon before becoming Chief Product and Strategy Officer of NextGen Healthcare, notes, the challenge of building a system with these qualities looks different at scale. “Every solution is a point solution unless

¹⁶ Centers for Medicare & Medicaid Services, “CMS Interoperability and Prior Authorization Final Rule (CMS-0057-F),” fact sheet, January 17, 2024.

¹⁷ Fidelis Fidelis Alu and Sunkanmi Oluwadare, “An Auditable and Source-Verified Framework for Clinical AI Decision Support: Integrating Retrieval-Augmented Generation with Data Provenance,” *Frontiers in Artificial Intelligence* 9 (2026): 1737532.

¹⁸ Otis Worldwide, “Our History: A Story of Innovation and Progress”; Mark Crawford, “Elisha Graves Otis,” ASME, June 8, 2012.

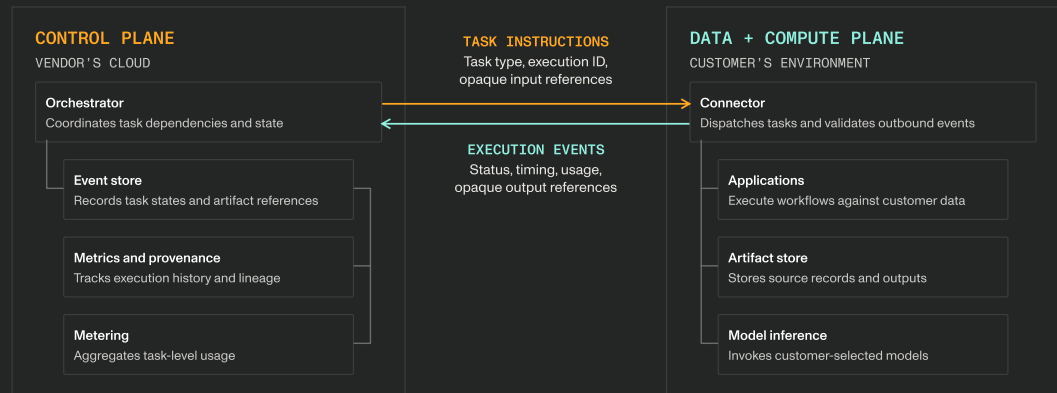
¹⁹ SAE International, “The Early Years: 1905–19,” History.

²⁰ Health Level Seven International, *HL7 FHIR Release 5*, version 5.0.0 (March 2023); CMS-0057-F requires FHIR-based APIs of Medicare Advantage organizations, Medicaid and CHIP fee-for-service programs and managed-care plans, and qualified health plan issuers on the federally facilitated exchanges (see note 16).

Separate the control plane from the data + compute plane

Control plane in the vendor's cloud; data and compute plane in the customer's environment.

ILLUSTRATIVE REFERENCE ARCHITECTURE



it's built to be integrated. You're never going to be the be-all, end-all system, so you have to be built to sit well inside a workflow—whether or not you're the thing someone sees in the UI—connecting into multiple health systems, multiple EMRs, multiple downstream systems.”²¹

Customizability across the stack

Ford offered the Model T in any color “so long as it is black.” When buyers wanted variety, General Motors offered “a car for every purse and purpose” and took the market. Ford had to shut his factories for months in 1927 to retool for a different car.²²

Every plan runs the same workflows differently: different criteria interpretations, different thresholds, different rules about what a machine may decide and what a human must. A scalable system holds

quality to the same bar but layers each plan's proprietary rules and methods on top: versioned, swappable, and configurable. A plan's accumulated judgment and processes stay its own intellectual property, walled off from unauthorized parties. And the plan steers the system with tools through which its own teams configure workflows, adjust thresholds, and improve outcomes without over-reliance on an external engineering queue.

“I think about it as a blend of what Henry Ford did, but then what Toyota did in lean process to give you quality—but quality with customization and options. And that together, I think, is what the AI chassis of healthcare today looks like.”

Aman Bhasin

Chief Information Officer, Capital Blue Cross

²¹ Gautam Shah, shoot interview, 2026.

²² Henry Ford with Samuel Crowther, *My Life and Work* (Doubleday, Page, 1922), chap. 4; The Henry Ford, “Advertisement for General Motors, ‘A Car for Every Purse and Purpose,’ 1925,” object ID 64.167.657.405; *Encyclopaedia Britannica*, s.v. “Henry Ford,” by Carol W. Gelderman, sec. “Later Years.”

“We do have an analogy that we can compare with what Ford did with the Model T in terms of having repeatable process. That’s our governance, that’s our rapid data availability,” Bhasin says. “But at the same time, as a CIO, I also feel like we have customers who are asking us for more. They want specific customized solutions. I think about it as a blend of what Henry Ford did, but then what Toyota did in lean process to give you quality—but quality with customization and options. And that together, I think, is what the AI chassis of healthcare today looks like.”

Cost reduction through engineering and operations

Low-cost intelligence is a core motivation for the AI transformation, and reducing costs at scale is both an engineering and an operational challenge.

In engineering, for example: routing each task to the smallest model that can do it verifiably rather than the largest one available; caching what has already been reasoned; designing the architecture so the expensive calls happen once, not per workflow; choosing hosted or dedicated capacity where appropriate.

And in operations: sampling completed work for statistical confidence; putting human judgment at the points of highest leverage; automating high-confidence, low-risk workflows and maintaining close manual involvement in low-confidence, high-risk ones; and building the tooling for small QA teams to supervise large volumes of work.

But to manage costs, the unit of production—the workflow—has to be made legible. Most clinical

operations today live in scattered systems and tacit knowledge: SOPs and policies as PDFs, procedures as training decks, judgment in the minds of clinicians who have completed the same procedure several thousand times. The preparatory work is in mapping where the data lives and how it flows; decomposing workflows into discrete tasks, each with defined inputs, outputs, and a standard for what “done” means; and digitizing policies into machine-executable form.

The future is already here

Since the early 1920s, deaths per mile driven in America have fallen by more than ninety percent (from roughly 18 deaths per hundred million miles to about 1.3) while miles driven have multiplied beyond anything the 1920s could have imagined.²³

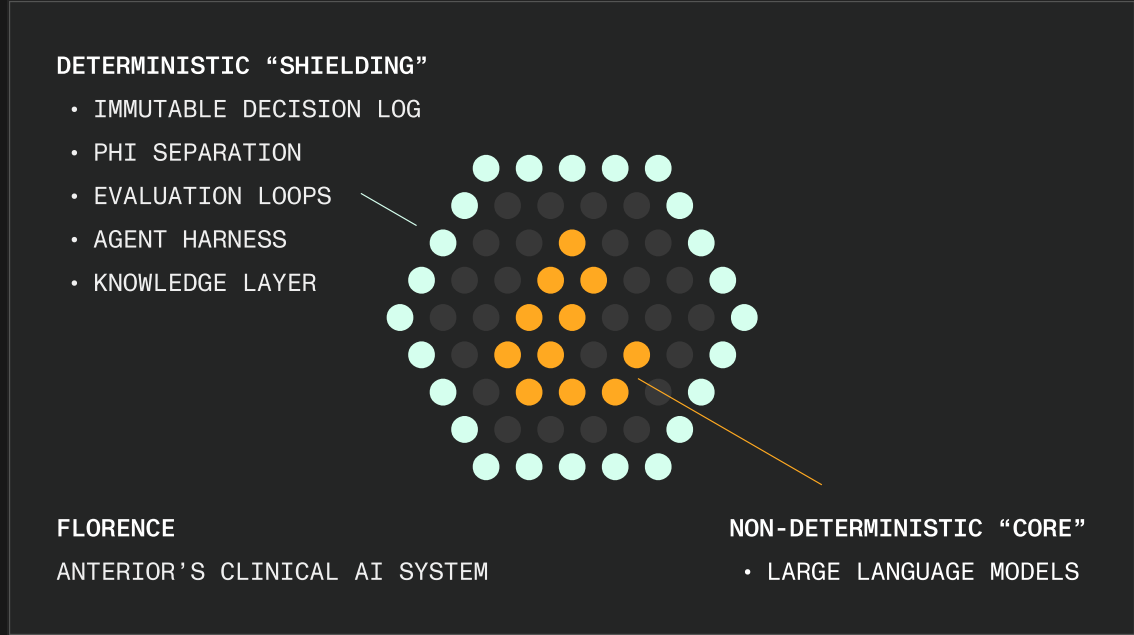
Clinical AI is earlier on the same curve, but the inflection is already visible. On some of the hardest clinical workflows that exist, from oncology determinations to spinal reviews, decision quality scores (precision, recall, and F1 scores) already run between 95 and 100 percent.²⁴

“When you’ve reached that threshold level of accuracy...you gain the confidence that the system is producing the output that you need, so that you can then scale it across to millions or tens of millions of claims, or tens of millions of insured lives.”

Rob Duffy Chief Technology Officer, HealthEdge

²³ National Safety Council, “Historical Car Crash Deaths and Rates,” *Injury Facts*, 1913–2024 series.

²⁴ Anterior internal evaluation data, 2026.



"You get to a point where the humans are just saying yes, that output is accurate," Duffy says of scaling a deployment. "And when you've reached that threshold level of accuracy...you gain the confidence that the system is producing the output that you need, so that you can then scale it across to millions or tens of millions of claims, or tens of millions of insured lives."

The combustion engine was a marvel.
But it waited eighty years for the car,

and the car waited another generation for the assembly line that built it a million times without breaking. The same opportunity is in front of the biggest industries today, of which healthcare, nearly a fifth of the American economy, is perhaps the most consequential.²⁵ The engine is here, and the car is already on the road. Whether we wait another generation to scale its value, or do it today, is up to us.

²⁵ Centers for Medicare & Medicaid Services, Office of the Actuary, "National Health Expenditures 2024 Highlights," January 14, 2026.