Surfing the AI Tidal Wave

A Discussion Paper for Health Systems

Prepared by:

John Sambrook

Common Sense Systems, Inc.

john@common-sense.com

https://common-sense.com

December 2025

Version 1.2

Request for Feedback

This is a discussion paper, not a finished analysis. The author seeks comments, criticism, and suggestions for improvement. If you see flaws in the logic, obstacles not considered, or ways to strengthen the argument, please reach out. Your input will help refine these ideas.

Change Log

Version	Date	Changes
1.2	December 2025	Strengthened credibility and addressed blind spots. Added "Ambient Clinical Documentation" as industry-wide example (Section 5.1) to complement proprietary tools. Added "The Integration Tax" subsection (Section 4.2) acknowledging that data liberation, not just analysis, is a constraint. Added "The Trust Dimension" subsection (Section 6.3) addressing data governance, the training vs. inference distinction, and human verification as safeguards; frames trust as a binding constraint and uses the "watertight vessel" metaphor for governance frameworks. Nuanced the Microsoft coding analogy (Section 3) to distinguish deterministic software from probabilistic clinical care. Added visual description of the shifting constraint (Section 4). Reframed Section 5 intro to present tools as "applications of this logic" rather than the only solutions.
1.1	December 2025	Significant reframe of core argument. Previous version emphasized navigating disruption; this version argues that the fundamental constraint on knowledge-work organizations (scarcity of high-quality cognitive capacity) is being relaxed. Added logical foundation: organizational capacity is bounded by (thinking quality) $\times$ (thinking throughput). AI addresses throughput; systems thinking addresses quality. The combination is close to optimal. Revised Executive Summary, Section 4, Section 6, and Conclusion to reflect opportunistic rather than defensive framing. Added AI cost decline data with citations (280x reduction, 10x/year). Added RaDonda Vaught case reference. Added AI Voice Agents as third concrete example.
1.0	December 2025	Initial discussion paper

Contents

Executive Summary

Health systems are knowledge-work organizations. This may not be obvious when we picture nurses at bedsides and surgeons in operating rooms. But the clinical moment is the visible tip of an iceberg. Beneath it lies an enormous mass of information gathering, synthesis, coordination, documentation, and decision-making. Administrative functions now consume 34–40% of healthcare spending.[1] Physicians spend nearly twice as much time on documentation as they do face-to-face with patients.[2]

A knowledge-work organization's capacity is bounded by its ability to think clearly, multiplied by the volume of clear thinking it can produce. For decades, this constraint has limited what healthcare organizations could achieve. High-quality cognitive capacity (the ability to gather information, synthesize it, identify what matters, and act on it) has been scarce and expensive.

That constraint is now being relaxed. Since late 2022, the cost of AI inference has fallen by more than 280 times. The rate of decline, roughly 10x per year, is faster than the fall in compute costs during the PC revolution or bandwidth costs during the dotcom boom. Tasks that were previously impractical, not merely expensive but genuinely beyond reach, are now achievable at negligible cost.

This paper argues that health systems have an opportunity to achieve outcomes that were previously impossible. The combination of AI (which dramatically increases thinking throughput) and systems thinking (which ensures thinking quality) attacks both terms in the equation that bounds organizational capacity. This is close to optimal.

The opportunity is not automatic. AI without rigorous thinking scales confusion. Reactive adoption, driven by vendor timelines and competitive pressure, produces the disappointing results that fuel legitimate skepticism. But proactive adoption, grounded in systems thinking and aligned with mission, can reduce administrative burden, improve care coordination, free clinicians for clinical work, and achieve levels of organizational effectiveness that scarcity of cognitive capacity previously made impossible.

This paper presents the case for proactive adoption, describes what thoughtful implementation looks like, and offers concrete examples of AI-enabled tools developed using systems thinking methodology. It is written as an invitation to dialogue, not a sales pitch. The author seeks criticism, correction, and collaboration.

[1] Himmelstein DU, Campbell T, Woolhandler S. Health Care Administrative Costs in the United States and Canada, 2017. *Annals of Internal Medicine*. 2020;172(2):134–142. https://pubmed.ncbi.nlm.nih.gov/31905376/

[2] Sinsky C, Colligan L, Li L, et al. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Annals of Internal Medicine*. 2016;165(11):753–760. https://pubmed.ncbi.nlm.nih.gov/27595430/

1 Introduction

Help me check my thinking.

I've been working on two related problems: complex patient discharge, and tooling for reducing friction in healthcare organizations. This work has led me to a perspective I'd like to test with others who care about the future of our health systems.

While this discussion paper references my local health system, EvergreenHealth in Kirkland, Washington, as a concrete example, the logic isn't specific to them. I'd invite you to consider whether it applies to yours.

Here are two observations that, taken together, suggest we're at an important decision point:

First: Health systems are fundamentally knowledge-work organizations. This may not be obvious. We naturally picture healthcare as hands-on clinical care delivered at the bedside. But consider what surrounds every clinical encounter: scheduling, insurance verification, prior authorization, documentation, coding, care coordination, discharge planning, compliance, quality reporting. The clinical moment is the visible tip; the iceberg beneath is information, coordination, and decision-making. Data suggest that administrative functions now consume 34–40% of healthcare spending,[3] and that physicians spend nearly twice as much time on documentation as they do face-to-face with patients.[4]

Second: The cost of performing knowledge work has begun falling rapidly, and will continue to fall. Since late 2022, AI capabilities have advanced to the point where tasks requiring judgment, synthesis, and language can increasingly be performed by machines at a fraction of previous cost. This shift is already displacing skilled knowledge workers, and it will not bypass healthcare.

The question I'm sitting with: *How do we ensure this capability serves our communities, rather than simply disrupting them?*

[3] Himmelstein et al. (2020), op. cit.
[4] Sinsky et al. (2016), op. cit.

2 Healthcare Is Knowledge Work

When we think of healthcare, we picture the nurse at the bedside, the surgeon in the operating room, the phlebotomist drawing blood. Physical acts performed by skilled humans.

But step back and look at what surrounds every clinical encounter.

Before the patient arrives: Scheduling, insurance verification, prior authorization, medical history synthesis, protocol selection, resource allocation.

During the encounter: Diagnosis, which is pattern recognition across symptoms, history, and test results. Treatment planning. Medication reconciliation. Documentation. Coding.

After the encounter: Care coordination, discharge planning, referral management, follow-up scheduling, quality reporting, billing, appeals when claims are denied.

Continuously: Policy interpretation, compliance monitoring, credentialing, supply chain decisions, capacity planning.

The hands-on clinical moment is real and essential. But it is the visible tip of an iceberg. The mass beneath the waterline is knowledge work: information gathering, synthesis, coordination, documentation, decision-making.

The data bear this out. Administrative functions now consume 34–40% of total healthcare spending.[5] Physicians in ambulatory settings spend nearly twice as much time on desk work as they do in direct clinical contact with patients.[6] U.S. insurers and providers together spent $812 billion on administration in 2017, amounting to $2,497 per capita, roughly double the rate of comparable countries.[7]

This is not a critique. Complex organizations require coordination. The question is whether that coordination must consume the resources it currently does, or whether new capabilities might change the equation.

[5] Himmelstein et al. (2020), op. cit.
[6] Sinsky et al. (2016), op. cit.
[7] Himmelstein et al. (2020), op. cit.

3 The Canary

What happens when the cost of knowledge work falls? We don't have to speculate. We can look at what's already happening in the Seattle Eastside economy.

The region built its prosperity on knowledge work, specifically software development. Microsoft, Amazon, and hundreds of smaller technology companies created an ecosystem where skilled cognitive labor commanded premium compensation. For decades, the trajectory pointed upward: more jobs, higher salaries, growing demand for people who could think systematically and solve complex problems.

That trajectory has bent.

Tech employment in Washington fell 6% from mid-2022 to early 2025, even as the national economy added jobs.[8] Amazon and Microsoft together have laid off more than 46,000 employees since 2023.[9] Job postings for developer positions on Indeed dropped from over 200 in February 2022 to just 69 by July 2024.[10]

The mechanism is visible. Microsoft now generates 30% of its code automatically.[11] Tasks that once required human judgment, synthesis, and careful implementation are increasingly performed by AI systems at a fraction of the cost. This isn't a temporary correction or a post-pandemic adjustment. It's a structural change in what human cognitive labor is worth.

A caveat is warranted here. Software is largely deterministic: code either runs or it doesn't, and automated testing can verify correctness at scale. Clinical care is probabilistic and high-stakes, involving judgment calls where context and nuance matter profoundly. The technology sector is a directional indicator of cognitive price deflation, not a direct operational map for clinical workflows. The pattern is instructive; the application requires translation.

The economics are stark. The cost of AI inference, the "running cost" of machine intelligence, has collapsed faster than any comparable technology shift. Between November 2022 and October 2024, the cost to run AI at GPT-3.5's level of performance dropped by more than 280 times, from $20 to $0.07 per million tokens.[12] Andreessen Horowitz describes this as "LLMflation": for an AI of equivalent performance, cost is decreasing by roughly 10x every year, faster than the decline in compute cost during the PC revolution or bandwidth during the dotcom boom.[13] This trajectory shows no sign of slowing.

The impact falls unevenly. Developers over 45 represent less than 6% of the industry workforce.[14]

[8] Washington Policy Center. "Could the Latest Round of Layoffs be Big Tech's 'Will the Last Person Leaving Seattle Turn Out the Lights' Moment?" November 2025. https://www.washingtonpolicy.org/publications/detail/could-the-latest-round-of-layoffs-be-the-big-techs-will-the-last-person-leaving-seattle-turn-out-the-lights-moment

[9] Amazon cut 27,000 corporate jobs in 2023; Microsoft has cut over 16,000 since 2023. See GeekWire and Layoffs.fyi tracking data.

[10] Axios Seattle. "Unemployment slightly up, tech jobs down in Seattle area." August 2024. https://www.axios.com/local/seattle/2024/08/13/unemployment-tech-jobs-washington

[11] Microsoft CEO Satya Nadella, speaking at Meta's LlamaCon, April 2025. Reported in *The Register*, https://www.theregister.com/2025/04/30/microsoft_meta_autocoding/

[12] Stanford AI Index Report 2025; Epoch AI data. See also Cerulean, "The Decreasing Cost of Intelligence," https://www.joincerulean.com/blog/the-decreasing-cost-of-intelligence

[13] Andreessen Horowitz, "Welcome to LLMflation," November 2024. https://a16z.com/llmflation-llm-inference-cost/

[14] Stack Overflow Developer Survey and industry data. See also: "Shocking Age Discrimination in Tech," Employment & Business News, August 2025. https://www.whatjobs.com/news/shocking-age-discrimination-in-tech-

Those who lose positions at older ages rarely recover their previous compensation. Research suggests they earn, on average, about half as much in subsequent roles.[15]

I am one data point in this pattern. My career was built on embedded software for medical devices: ultrasound systems, defibrillators, safety-critical systems where mistakes can harm patients. Thirty years of experience. The kind of work that once commanded $250,000 annually. In the past year, my consulting revenue has been approximately $10,000.

I mention this not to elicit sympathy but to establish standing. When I say that the cost of knowledge work is falling and that this will affect healthcare, I'm not theorizing from a distance. I'm reporting from the impact zone.

The pattern that has already reshaped Seattle's technology sector will not bypass healthcare. The administrative burden that consumes 34–40% of healthcare spending, the documentation, coordination, coding, and compliance work, involves exactly the kind of cognitive tasks that AI systems are learning to perform. The question is not whether this shift will occur, but how organizations will navigate it.

Emad Mostaque, founder of Stability AI, frames the timeline starkly. In his recent book *The Last Economy*, he argues that we have roughly 1,000 days to prepare for a fundamental restructuring of how knowledge work creates value.[16] That clock started when ChatGPT launched in November 2022. We are now approximately 750 days in.

Whether his specific timeline proves accurate, the direction is clear. And the Seattle Eastside offers a preview of what happens when organizations wait for market pressure rather than moving proactively.

how-61-of-developers-over-45-face-career-extinction/

[15] See CNBC, "Laid-off techies face 'sense of impending doom,'" March 2024, reporting job offers requiring 50%+ pay cuts. https://www.cnbc.com/2024/03/15/laid-off-techies-struggle-to-find-jobs-with-cuts-at-highest-since-2001.html

[16] Mostaque E. *The Last Economy: A Guide to the Age of Intelligent Economics*. 2025. Available at Amazon and other retailers. https://www.amazon.com/Last-Economy-Guide-Intelligent-Economics-ebook/dp/B0FNDMWRZT

4 The Opportunity

Healthcare has always operated under an implicit assumption: high-quality cognitive capacity is scarce and expensive. Every system, every workflow, every staffing model reflects this assumption. The 34–40% administrative burden exists partly because thinking is expensive, so organizations build bureaucratic structures to routinize decisions and reduce the need for judgment.

What if that assumption becomes obsolete?

A knowledge-work organization's ceiling is set by (thinking quality) × (thinking throughput). For decades, throughput was the binding constraint. There were only so many hours, only so many skilled people, only so much cognitive capacity to go around. Organizations rationed it carefully, reserving human judgment for decisions that seemed to warrant it and routinizing everything else.

The Shifting Constraint
Imagine a balance scale. On one side sits *Throughput*: how much cognitive work an organization can perform. On the other side sits *Quality*: how rigorous, accurate, and well-directed that thinking is.
For decades, Throughput was the heavy side, the bottleneck. Organizations could not think about enough problems, fast enough, to address everything that mattered. Quality was often sacrificed to get through the volume.
AI has placed a thumb on the Throughput side of the scale. The bottleneck is shifting. The scarce resource is no longer raw cognitive capacity but the quality and direction of thinking: knowing which problems to solve, how to frame them correctly, and how to avoid scaling confusion alongside insight.
This is why systems thinking matters. It addresses the constraint that is becoming binding as the old constraint relaxes.

The throughput constraint is now being relaxed. The same capabilities that are displacing software developers (AI systems that can read, synthesize, identify patterns, and generate coherent output) can address healthcare's cognitive load. Between November 2022 and October 2024, the cost of AI inference fell by more than 280 times. This is not a marginal improvement. It is a structural change in what cognitive work costs.

4.1 The Integration Tax

A note of realism is necessary here. While AI inference costs have dropped 280 times, the cost of moving data securely out of electronic health records has not. Healthcare data lives in silos, protected by layers of access controls, governed by regulations, and locked into vendor ecosystems that do not always prioritize interoperability.

The constraint is shifting from *analyzing* data to *liberating* data. An AI system that could synthesize a patient's complete history in seconds is of limited value if extracting that history from the EHR takes weeks of IT negotiation and compliance review.

This is not an argument against the opportunity. It is an argument for systems thinking. The integration challenge is real, but it is a problem of policy, governance, and architecture, not a problem of cognitive capacity. Organizations that approach AI adoption with discipline can design data flows that respect security requirements while enabling the analysis that creates value. Organizations that layer AI tools onto existing silos without examining data architecture will discover

that the bottleneck has merely moved, not disappeared.

The point stands: high-quality thinking about *which* data to liberate, *how* to govern its movement, and *where* to apply AI analysis is now the scarce resource. The tools to do the analysis, once the data is accessible, are increasingly cheap.

Consider what this means practically. AI systems can now: read and synthesize large volumes of text, identify patterns across documents, generate coherent summaries, flag inconsistencies, draft communications, answer questions about complex material. These are precisely the cognitive tasks that consume so much of healthcare's resources.

Prior authorization requests. Documentation review. Care coordination across multiple providers. Policy interpretation. Coding and compliance verification. Discharge planning logistics. Each of these involves gathering information, synthesizing it, identifying what matters, and communicating clearly. Work that humans currently do, often under time pressure, often with incomplete information, often while juggling competing demands.

AI doesn't get tired. It doesn't experience moral injury from being asked to do the impossible. It can review a hundred pages of policy documents and surface the three paragraphs that create an operational conflict. It can monitor a patient's chart continuously and flag when discharge planning should begin. It can draft the prior authorization appeal while the case manager focuses on the family conversation that requires human judgment and compassion.

The opportunity is not to replace clinicians. It is to remove the friction that prevents clinicians from doing clinical work.

Recall the data from earlier: physicians in ambulatory settings spend nearly twice as much time on documentation as they do with patients. Administrative functions consume 34–40% of healthcare spending. These numbers represent human effort devoted to information management rather than care delivery. If even a fraction of that effort could be redirected, or simply eliminated because the underlying task was automated, the benefit would cascade across the system.

Better use of clinician time. Reduced burnout from impossible workloads. Faster resolution of discharge barriers. Fewer patients waiting in acute care beds for paperwork to clear. Lower costs without reduced quality. Potentially, *higher* quality because the humans in the system have capacity to exercise judgment rather than process transactions.

This is not utopian speculation. It follows directly from what the technology can already do, applied to problems we already understand.

4.2 A Note on Skepticism

A note of honesty is warranted here. There is significant skepticism about AI in healthcare, and much of it is earned.

What we are seeing in many organizations is reactive adoption: AI tools purchased because competitors are purchasing them, implemented in fragments, configured to vendor defaults, layered onto existing processes without examining whether those processes make sense. The results are predictable. Point solutions that don't integrate. New friction introduced alongside old friction reduced. Staff asked to learn yet another system. Promised efficiencies that fail to materialize because the underlying workflow was never redesigned.

This is not a failure of AI. It is a lack of systems thinking.

AI is a capability. Like any capability, it produces value only when deployed within a system that understands its own constraints, bottlenecks, and goals. An organization that doesn't know where its real problems lie will not solve them by adding technology. It will automate the wrong things, or automate the right things in ways that create new problems elsewhere.

The skeptics are observing this pattern and concluding, reasonably, that AI in healthcare is overhyped. What they may be missing is that the failures they observe are failures of *adoption*, not failures of *capability*. The technology can do what it claims. The question is whether organizations can deploy it thoughtfully.

This is why I believe systems thinking is not optional. It is the difference between AI that helps and AI that generates expensive new problems. Theory of Constraints methodology, applied before and during AI adoption, can identify where the real leverage points are, what assumptions need examination, and how to sequence changes so they reinforce rather than undermine each other.

5 Concrete Examples

The shift described above is not theoretical. Organizations are already applying AI to healthcare's cognitive burden. Let me describe four examples: one that is gaining broad industry adoption, and three that I have been developing. The point is not to advocate for specific tools but to demonstrate that the opportunity is practical and the pattern is reproducible.

5.1 Ambient Clinical Documentation

The most visible example of AI reducing cognitive burden in healthcare is ambient documentation: AI systems that listen to patient-clinician conversations and automatically generate clinical notes.

Products like Nuance DAX, Abridge, and similar tools are now deployed across thousands of healthcare organizations. The value proposition is straightforward. Physicians spend nearly half their time on documentation. Ambient AI can draft notes in real-time, reducing that burden to review and approval rather than composition from scratch.

Early adopters report meaningful reductions in after-hours documentation (the "pajama time" that contributes to burnout) and increased time available for patient interaction. The technology is imperfect; clinicians must review and correct the output. But the shift from *writing* to *editing* represents a genuine reduction in cognitive load.

This is not my work. I mention it because it demonstrates that the pattern described in this paper is already operating at scale. The constraint (documentation burden) was identified. AI capability was applied. The result is measurable improvement in how clinicians spend their time.

The examples that follow are applications of the same logic to different constraints.

5.2 Policy Conflict Detection

Organizations accumulate policies over time, written by different people for different purposes. These documents often contain what I call "embedded conflicts": language that establishes mutually exclusive requirements, creating situations where frontline staff cannot comply with one mandate without violating another.

The accommodation is not passive acceptance. Staff develop workarounds, informally learned and passed from experienced workers to new ones, that allow work to proceed but technically violate policy. When a workaround leads to a bad outcome, the individual is blamed for the violation rather than the system that made violation necessary.[17]

These conflicts are difficult to detect through normal review. No single document contains the problem. It emerges from interactions between documents, or from requirements that seem reasonable in isolation but prove impossible in combination.

I have developed tooling, grounded in Theory of Constraints methodology, that systematically examines policy documents for these structural conflicts. The tool doesn't just identify problems. It proposes resolutions that address root causes rather than symptoms. In demonstration analysis

[17] The 2017 case of nurse RaDonda Vaught at Vanderbilt University Medical Center illustrates this pattern. Systemic conditions, including normalized override cultures, software transition challenges, and workload pressures, converged to create a fatal medication error. Vaught was criminally prosecuted; the system that made the workaround necessary was never on trial. For detailed analysis, see Sambrook J. "The Hazard We Don't See: How Policy Conflicts Harm the People Organizations Are Meant to Protect." Common Sense Systems, Inc. 2025.

of a single end-of-life care policy, the methodology surfaced seven major conflicts, each generating operational friction, legal exposure, and staff moral distress.

The broader application: any organization with a substantial policy corpus likely harbors embedded conflicts that create daily impossible situations for staff. Finding and resolving these conflicts eliminates self-imposed constraints. It removes waste that the organization is inflicting on itself.

5.3 The Post-Acute Care Plan (PACP)

Complex hospital discharges cost health systems billions annually. These are patients who remain hospitalized despite being medically ready, because discharge arrangements aren't in place. At some Washington hospitals, patients awaiting guardianship services occupy acute care beds for an average of 124 days.[18]

The standard approach is reactive: discover what's missing during hospitalization, then scramble to resolve it while the patient waits at $2,500 or more per day.

PACP inverts this. It captures discharge-relevant information proactively, during the pre-hospitalization phase, while patients are still living in the community. Legal documents. Insurance details. Care preferences. Support networks. Home environment. When hospitalization occurs, discharge planning becomes execution of an existing plan rather than crisis management.

This is systems thinking applied to a specific constraint. The bottleneck isn't clinical care. It's information that should have been gathered earlier. PACP addresses the constraint directly.

5.4 AI Voice Agents

A third area of development involves AI-powered voice agents: conversational interfaces that can explain complex information, answer questions, and guide users through unfamiliar territory.

One application is the knowledge worker assistant. The premise is simple: everyone gets a free executive assistant. Voice agents can handle scheduling, answer routine questions, take messages, route inquiries, and manage the administrative overhead that consumes professional time. For healthcare organizations, this could mean staff spending less time on phone trees and more time on work that requires human judgment.

Another application is educational. Complex topics that affect communities, such as ballot measures, bond issues, or policy changes, often go unexplained because the explanation itself is expensive. A voice agent can patiently walk a caller through what a hospital district bond measure means, how it affects their taxes, what services it funds, and why it matters. The same capability could explain insurance options, care decisions, or discharge instructions in ways that adapt to each caller's questions and pace.

These are not hypothetical. Working prototypes exist. The underlying technology has matured to the point where natural conversation with an AI is no longer jarring. What remains is applying this capability to problems worth solving.

All four of these examples share a common structure: AI capability combined with systems thinking, applied to problems that currently consume resources without producing proportionate value.

[18] Data from author's analysis of complex discharge cases in Washington State. Documentation available upon request.

Ambient documentation is gaining industry-wide traction. The other three are applications of the same logic to constraints that remain largely unaddressed. The point is not any specific tool but the pattern. When you understand your constraints and apply appropriate technology, you get leverage. These examples exist to show that the pattern works, not to claim that any particular implementation is the only path forward.

6 The Choice

The constraint that has bound healthcare organizations since their inception, scarcity of high-quality cognitive capacity, is being relaxed. This creates an opportunity, but opportunities must be seized.

An organization that grasps this early could stop rationing cognitive effort and start applying it abundantly to problems that were previously "not worth the time." It could redirect expensive human cognition to the places where it is genuinely irreplaceable: presence, judgment, relationship, the genuinely novel. It could solve coordination problems that were previously intractable because the cost of thinking through all the interdependencies was prohibitive.

But organizations do not transform themselves automatically. The default path is reactive adoption: waiting for market pressure to force the issue, adopting tools when competitors adopt them or when payers require them, implementing quickly under pressure without time to examine how new capabilities interact with existing workflows. This path produces the disappointing outcomes the skeptics rightly criticize.

The question is not whether AI arrives in healthcare. It is whether your organization captures the opportunity or merely absorbs the disruption.

Option one: Reactive adoption. Wait for external pressure. Let vendors define what the tools do. Implement quickly, without time for systems thinking. The technology serves whoever designed it, which may or may not align with your mission or community's needs. This is the path of least resistance. It is also the path that leaves value on the table and produces the disappointing results that make AI look like hype.

Option two: Proactive adoption. Begin now, while there is time to be thoughtful. Understand your actual constraints and bottlenecks. Pilot AI capabilities in controlled settings where learning is possible. Develop internal expertise so you can evaluate vendor claims and configure tools appropriately. Engage frontline staff so adoption addresses their real problems.

This path requires investment before the pressure becomes acute. It requires leadership willing to allocate attention to something that is not yet an emergency. It requires the discipline to apply systems thinking before layering in new capabilities.

The reward is an organization operating at a level of effectiveness that scarcity of cognitive capacity previously made impossible. Not merely surviving the transition, but flourishing because of it.

6.1 The Human Dimension

There is a human dimension to this choice as well.

The displacement occurring in Seattle's technology sector is painful for those experiencing it. Skilled people who did nothing wrong find themselves struggling to contribute, watching capabilities they spent careers developing become commoditized. That pain is real and should not be dismissed.

Health systems have an opportunity to navigate this transition more thoughtfully than the technology sector has. Healthcare involves work that genuinely requires human presence, judgment, and compassion. The goal is not to eliminate human contribution but to redirect it toward what humans do best. AI handles the information management. Humans handle the relationships, the judgment calls, the moments when a patient needs someone who sees them as a person.

Done well, this transition could make healthcare work more sustainable, not less. Clinicians freed from documentation burden. Case managers focused on complex family situations rather than paperwork. Administrative staff redeployed to higher-value work rather than eliminated.

Done poorly, it will produce the same dislocation the technology sector is experiencing, with the added harm of degraded care for the communities health systems exist to serve.

6.2 The Fiscal Dimension

A software developer earning $250,000 pays substantially more in taxes than the same person earning $10,000. Multiply this across thousands of displaced knowledge workers in King County, and the arithmetic becomes stark. The tax base that supports public services, including public hospital districts, is eroding from one direction while obligations continue to grow from the other.

Public hospital districts operate within this environment. Their funding depends on a community whose economic foundation is shifting. The displacement of knowledge workers is not an abstract phenomenon happening to other people. It is a change in the fiscal ground on which public institutions stand.

6.3 The Trust Dimension

There is a third dimension to this choice, perhaps the most critical: trust.

The economic argument for AI, that inference costs have dropped 280 times, inevitably raises a safety counter-argument. If "cheap thinking" relies on sending sensitive patient data to external cloud providers, is the efficiency worth the risk?

In healthcare, trust is a binding constraint. If clinicians or patients suspect that their data is being harvested to train public models, or that privacy is being traded for speed, the system will reject the technology regardless of its utility.

A systems thinking approach treats data sovereignty not as a compliance checkbox but as a foundational design requirement.

Sovereignty over speed. The collapsing cost of inference applies to private, enterprise-grade instances just as it does to public chatbots. We must distinguish between *training* (teaching the AI) and *inference* (asking the AI to think). Health systems can and must demand architectures where data is used solely for the immediate task and then discarded, never retained to train the model for others.

Verification over blind faith. AI hallucinations, plausible but incorrect outputs, are a real risk. This is why the human in the loop remains essential. The AI's role is to synthesize and propose; the human's role is to verify and decide. This structure leverages the AI's speed while retaining the human's accountability.

We cannot "surf the tidal wave" if the board, the legal team, and the community are terrified of the water. Proactive adoption requires constructing a vessel, a governance framework, that keeps the data watertight while harnessing the power of the current.

The choice is real. The window for making it thoughtfully is finite.

7 How to Proceed

If the argument in this paper has merit, what should a health system actually do?

I want to suggest an approach, independent of whether I am involved in it.

Start with systems thinking, not technology. Before purchasing AI tools or launching pilots, understand where the real constraints lie. What are the bottlenecks that limit your organization's performance? Where do staff encounter impossible situations created by conflicting requirements? What information gaps cause downstream problems? AI applied to the wrong problems will produce disappointing results. AI applied to the actual constraints can produce substantial gains.

Work on a real problem, bounded but meaningful. Pick something concrete. A policy domain where conflicts create operational friction. A workflow where information gaps cause delays. A coordination challenge that consumes disproportionate staff time. The goal is to learn how AI can help with *your* problems, not to implement AI in the abstract.

Involve the people who do the work. Frontline staff know where the friction is. They experience the impossible situations daily. Any AI adoption that does not incorporate their knowledge will likely automate the wrong things or create new burdens. Their involvement also builds ownership. A tool that staff helped develop is a tool staff will actually use.

Build internal capability, not dependency. The goal is not to hire consultants indefinitely. It is to develop your organization's capacity to use AI effectively. This means your people learning to work with AI as a thinking partner. It means developing judgment about what AI does well and where human oversight remains essential. The engagement model that serves you best is one where external help decreases over time as internal capability grows.

Apply systems thinking throughout. AI adoption is itself a change that interacts with existing systems. Resistance, unintended consequences, and second-order effects are predictable. Theory of Constraints methodology, or similar frameworks, can help sequence changes so they reinforce rather than undermine each other. The point is not to adopt a specific methodology but to bring disciplined thinking to a process that otherwise tends toward chaos.

This approach is faster than adopting tools on an ad-hoc basis, though it may feel slower at first. Ad-hoc purchases create the illusion of progress. The problems they introduce, such as tools that don't integrate, workflows that conflict, staff who resist systems imposed without their input, only become visible months or years later. By then, the organization has compounded the problem with additional ad-hoc decisions, and untangling the mess is genuinely slow and expensive. Starting with systems thinking avoids that trap.

8 Conclusion

This paper began with a request: help me check my thinking.

I have argued that health systems are knowledge-work organizations, and that the capacity of knowledge-work organizations is bounded by (thinking quality) × (thinking throughput). For decades, the throughput term was the binding constraint. High-quality cognitive capacity was scarce and expensive, and every system reflected that scarcity.

That constraint is now being relaxed. AI has made cognitive throughput abundant and cheap. The cost has fallen 280 times in two years, and continues to fall at roughly 10x per year. Tasks that were previously impractical, not merely expensive but genuinely beyond reach, are now achievable.

I have argued that systems thinking addresses the other term in the equation. AI without rigor scales confusion. But AI combined with disciplined thinking about constraints, bottlenecks, and unintended consequences produces leverage. The combination attacks both terms. This is close to optimal.

I have offered four examples: ambient documentation (already gaining industry-wide adoption), tooling that surfaces embedded conflicts in policy documents (a task that was genuinely impossible before, regardless of budget), a framework for proactive discharge planning that addresses barriers before hospitalization occurs, and voice agents that can explain complex information or serve as assistants for knowledge workers. These are not the only possibilities. They are instances of a broader principle: the cognitive capacity constraint has moved, and organizations that recognize this can achieve outcomes that were previously impossible.

I may be wrong about some of this. The timeline may be longer than I suggest. The opportunity may be harder to capture than I believe. If you see errors in the logic or obstacles I have not considered, I want to know.

What I am confident about is this: the ceiling has lifted. Organizations that continue operating as if cognitive capacity were scarce will leave enormous value on the table. Organizations that recognize the shift and combine AI throughput with rigorous systems thinking can achieve levels of effectiveness that scarcity previously made impossible.

The opportunity is real. Execution determines who captures it.

I welcome the opportunity to continue this conversation.

John Sambrook
Common Sense Systems, Inc.
john@common-sense.com
common-sense.com

References and Sources

Healthcare Administrative Costs

Himmelstein DU, Campbell T, Woolhandler S. Health Care Administrative Costs in the United States and Canada, 2017. *Annals of Internal Medicine*. 2020;172(2):134–142. doi:10.7326/M19-2818
https://pubmed.ncbi.nlm.nih.gov/31905376/

Sinsky C, Colligan L, Li L, et al. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Annals of Internal Medicine*. 2016;165(11):753–760. doi:10.7326/M16-0961
https://pubmed.ncbi.nlm.nih.gov/27595430/

American Hospital Association. "Costs of Caring." 2025.
https://www.aha.org/costsofcaring

Sahni NR, et al. "Administrative Simplification: How to Save a Quarter-Trillion Dollars in US Healthcare." McKinsey & Company. October 2021.

Health Affairs. "The Role Of Administrative Waste In Excess US Health Spending." Health Affairs Health Policy Brief. October 2022.

Seattle Area Tech Employment

Washington Policy Center. "Could the Latest Round of Layoffs be Big Tech's 'Will the Last Person Leaving Seattle Turn Out the Lights' Moment?" November 2025.
https://www.washingtonpolicy.org/publications/detail/could-the-latest-round-of-layoffs-be-the-big-techs-will-the-last-person-leaving-seattle-turn-out-the-lights-moment

Axios Seattle. "Unemployment slightly up, tech jobs down in Seattle area." August 2024.
https://www.axios.com/local/seattle/2024/08/13/unemployment-tech-jobs-washington

Seattle Times. "Seattle job market is past its 'golden age' for tech workers." November 2025.
https://www.seattletimes.com/

Layoffs.fyi. Tech Layoff Tracker.
https://layoffs.fyi/

GeekWire. "Amazon's total Washington state employment declines for the first time, new numbers show." April 2024.
https://www.geekwire.com/2024/amazons-total-washington-state-employment-declines-for-the-first-time-new-numbers-show/

Age Discrimination in Technology

Employment & Business News. "Shocking Age Discrimination in Tech: How 61% of Developers Over 45 Face Career Extinction." August 2025.
https://www.whatjobs.com/news/shocking-age-discrimination-in-tech-how-61-of-developers-over-45-face-career-extinction/

IEEE Spectrum. "Does Age Discrimination in Tech Start at 40?" August 2022.
https://spectrum.ieee.org/does-age-discrimination-in-tech-start-at-40

CNBC. "Laid-off techies face 'sense of impending doom' with job cuts at highest since dot-com crash." March 2024.
https://www.cnbc.com/2024/03/15/laid-off-techies-struggle-to-find-jobs-with-cuts-at-highest-since-2001.html

diginomica. "The tech industry has an ageism problem - here's why that matters to us all." November 2024.
https://diginomica.com/tech-industry-has-ageism-problem-heres-why-matters-us-all

AI and Economic Transition

Mostaque E. *The Last Economy: A Guide to the Age of Intelligent Economics*. 2025.
https://www.amazon.com/Last-Economy-Guide-Intelligent-Economics-ebook/dp/B0FNDMWRZT

The Register. "30 percent of some Microsoft code now written by AI." May 2025.
https://www.theregister.com/2025/04/30/microsoft_meta_autocoding/

R&D World. "Microsoft CEO says AI now writes up to 30% of company code." May 2025.
https://www.rdworldonline.com/microsoft-ceo-says-ai-now-writes-up-to-30-of-company-code/

Indeed Hiring Lab. "The US Tech Hiring Freeze Continues." July 2025.
https://www.hiringlab.org/2025/07/30/the-us-tech-hiring-freeze-continues/

AI Cost Trends

Andreessen Horowitz. "Welcome to LLMflation - LLM inference cost is going down fast." November 2024.
https://a16z.com/llmflation-llm-inference-cost/

Epoch AI. "LLM inference prices have fallen rapidly but unequally across tasks." March 2025.
https://epoch.ai/data-insights/llm-inference-price-trends

Stanford HAI. *Artificial Intelligence Index Report 2025*. Stanford University.
https://aiindex.stanford.edu/report/

Cerulean. "The Decreasing Cost of Intelligence." 2025.
https://www.joincerulean.com/blog/the-decreasing-cost-of-intelligence

Wing Venture Capital. "The plummeting cost of intelligence." 2024.
https://www.wing.vc/content/plummeting-cost-ai-intelligence

Theory of Constraints

Goldratt EM, Cox J. *The Goal: A Process of Ongoing Improvement*. North River Press. 1984.

Goldratt EM. *Theory of Constraints*. North River Press. 1990.

Dettmer HW. *The Logical Thinking Process: A Systems Approach to Complex Problem Solving*. ASQ Quality Press. 2007.

Related Papers by the Author

Sambrook J. "Theory of Constraints Analysis for Healthcare Policy Conflict Identification." Concept Paper. December 2025. Available from the author.

Sambrook J. "The Hazard We Don't See: How Policy Conflicts Harm the People Organizations Are Meant to Protect." Concept Paper. December 2025. Available from the author.

Sambrook J. "The Post-Acute Care Plan (PACP): A Proactive Approach to Discharge Readiness." Concept Paper. November 2025. Available from the author.

Sambrook J. "PACP and Legal Barriers to Discharge: From POA Gaps to Guardianship." Companion Paper. December 2025. Available from the author.

Common Sense Systems, Inc.
https://common-sense.com
john@common-sense.com