Impacting the Economics of Care

Why governance, policy, and architectural alignment must precede the next wave of healthcare AI investment

About 13point1 Advisors

13point1 Advisors is a boutique management and strategy consulting firm specializing in payment integrity and claims processing. We partner with health plans and solution providers to align governance, policy, technology, and analytics so that technology investments compound capability rather than fragmentation. Our perspective is informed by decades of operating experience inside the nation's largest health plans and solution providers.

Abstract

U.S. healthcare administration consumes nearly twice the share of total spending that peer systems require. The prevailing industry response—digitization, AI, and predictive analytics—produces mixed results because technology amplifies whatever foundation it sits on. This paper argues that impacting the economics of care requires multiple solutions operating on common governance, coherent policy, and aligned architecture. Using the 13point1 Reference Framework as its structural spine, the paper traces how each foundational layer either compounds value or compounds fragmentation. The implications differ for health plan executives, solution providers, and the broader ecosystem, but the core conclusion is consistent: the governance question precedes the consolidation question, and the returns to alignment are measurable, durable, and already visible among operators who have made the investment.

Published May 2026
© 2026 13point1 Advisors. All rights reserved.
No reproduction, redistribution, or derivative use without written permission.

I. The Question Worth Asking

The United States spends more on healthcare administration than any nation on earth, and the gap is not close. In a landmark study, Himmelstein, Campbell, and Woolhandler found that administrative costs consume approximately 34 percent of total U.S. health expenditures, roughly double the 17 percent that Canada spends under its single-payer system. If the United States had streamlined its administrative spending to Canadian levels in 2017 alone, the savings would have exceeded $600 billion.[1]

This figure is staggering, yet it tells us something more nuanced than the standard waste narrative suggests. A substantial share of that administrative cost reflects the structural complexity of a multi-payer system that serves over 330 million people through commercial insurance, Medicare, Medicaid, and hybrid programs like Medicare Advantage. A synthesis of micro-costing evidence by Jiwani et al. estimated that billing and insurance-related activities alone accounted for $471 billion in 2012, approximately 80 percent of the excess administrative cost relative to simplified financing systems.[2] The administrative architecture carries real infrastructure, real contractual relationships, and decades of compounded investment.

A separate analysis published in JAMA by Sahni, Carrus, and Cutler estimated that administrative simplification could save more than $265 billion annually—a figure that underscores the scale of waste attributable not to clinical complexity but to the friction between policy intent and operational execution. The gap between the Himmelstein and Sahni estimates reflects differences in methodology and scope, but the directional conclusion is consistent: the administrative cost burden is structural, not incidental, and the returns to simplification are enormous.[1, 29]

> ***This is not a system that is broken. It is a system that is opaque.***

The distinction matters profoundly. Broken implies replacement. Opaque implies evolution: bringing transparency and integration to what already exists. Every major reform attempt of the past three decades, from the Affordable Care Act (ACA)'s value-based payment provisions to the No Surprises Act and Centers for Medicare and Medicaid Services (CMS) price transparency mandates, has chosen evolution over revolution. The system is not going to reset. It is going to evolve.

The federal government's own actions confirm this trajectory. In February 2026, CMS published a comprehensive Request for Information (RFI) known as Comprehensive Regulations To Uncover Suspicious Healthcare, or "CRUSH," soliciting stakeholder feedback on regulatory changes to strengthen fraud prevention across Medicare, Medicaid, Children's Health Insurance Program (CHIP), and the Marketplace. The RFI is not proposing to replace the system. It is proposing to make it visible. CMS's Fraud Defense Operations Center, formed in 2025, used real-time cross-functional targeting to produce $1.8 billion in taxpayer savings, over $100 million from suspect laboratories alone.[3] That is what happens when governance catches up to operations.

CMS's own improper payment data underscores the urgency. In fiscal year 2025, Medicare Fee-for-Service reported an improper payment rate of 6.55 percent, representing $28.83 billion. Medicare Part C added $23.67 billion at a 6.09 percent rate, and Medicaid contributed $37.39 billion at 6.12 percent. Combined, improper payments across these three programs alone approach $90 billion annually.[4] The American Hospital Association reported in 2024 that hospitals spend nearly $20 billion annually appealing denials, more than half of which involve claims that should have been paid at the time of initial submission.[5]

The data tells a consistent story: billions flowing through correction and recovery cycles that exist because the system cannot measure its own accuracy in real time. That visibility gap—not any single policy failure or technology deficit—is what makes the economics of care so resistant to improvement.

The industry's instinct, understandable given market incentives, is to reach for technology. AI-powered claims review. Predictive analytics for fraud detection. Machine learning for coding accuracy. Yet as Brynjolfsson's foundational research on the IT productivity paradox has demonstrated, technology investments produce outsized returns only when matched with organizational complements: governance structures, process alignment, and workforce readiness. Without these complements, technology amplifies whatever it sits on top of, including misalignment.[6]

This paper argues that the industry is chasing analytics and AI while the foundation underneath remains misaligned. That is building on sand. The question worth asking is what must be true in our organization before a portfolio of technology investments compounds.

No single solution spans the full payer operational landscape—utilization management, pre-service decisioning, payment integrity, post-pay recovery—and none needs to. What separates payers that compound value from those that compound fragmentation is whether their portfolio of tools shares a common governance, coherent policy, and aligned architecture. That foundation is what impacting the economics of care actually requires.

A note on sources: This paper draws on published research, regulatory filings, and practitioner perspectives from across the healthcare ecosystem—payers, solution providers, and regulatory bodies—to ground the argument in operational reality. Where specific organizations are cited, it is to illustrate a published viewpoint or operational methodology relevant to the analysis, not to endorse a product or service. Some organizations cited herein maintain commercial relationships with the authors or their clients; inclusion is determined by analytical relevance, not commercial affiliation.

II. The Actionable vs. the Optimal

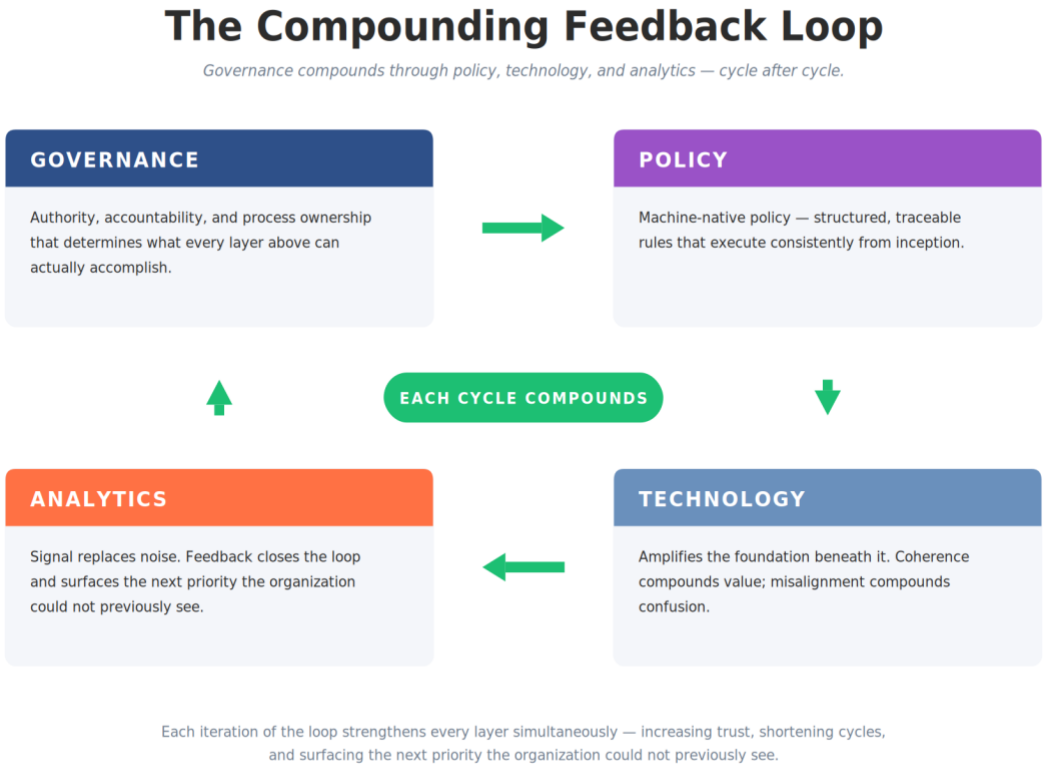
Every healthcare executive has heard the pitch: a comprehensive, multi-year transformation that will fundamentally reimagine operations. The vision is compelling. The timeline is not. In an industry where CMS issues hundreds of transmittals and change requests annually, where the regulatory landscape shifts with each administration, and where medical cost trends can swing several percentage points in a single year, waiting for the optimal solution is itself a cost, and it compounds.

The actionable move is different: bring speed and transparency to what exists today. Not a multi-year roadmap to a perfect state, but an immediate commitment to shorter feedback cycles and clearer visibility into current operations. This is not a modest aspiration. It is the most leveraged intervention available.

The Compounding Loop

Peter Senge's research on organizational learning identifies reinforcing feedback loops as the fundamental mechanism through which organizations improve. Chris Argyris's distinction between single-loop and double-loop learning sharpens the point: single-loop learning fixes the error; double-loop learning fixes the system that produced it. Most payment integrity operations today are trapped in single-loop learning—recovering overpayments without interrogating the governance and policy structures that produced them. The central mechanism of this paper is an engine for double-loop learning: a compounding feedback loop that operates through four interconnected stages. Speed and transparency bring clarity to what exists today. Consider a health plan that deploys real-time claim editing with transparent policy logic. A provider submits a claim for evaluation and management services billed alongside a procedure. Instead of a silent denial followed by a 90-day appeals cycle, the system surfaces the specific edit, the policy criterion behind it, and the clinical evidence supporting that criterion—at the point of submission. That single act of transparency shortens the feedback cycle from months to moments. Shorter feedback cycles build trust. When providers can see why a decision was made and verify it against published criteria, the interaction shifts from adversarial dispute to collaborative correction. Trust, once established, restores visibility: providers share clinical context earlier, payers surface decision logic more openly, and the bilateral information gap that drives the economics of care begins to close. Visibility, in turn, exposes what is actually broken—and those broken parts can be attacked one at a time. The claim editing example above may reveal that the underlying policy criterion is outdated, that the governance process for updating it has no owner, or that the technology interpreting it lacks access to the clinical record. Each of these is a discrete, addressable problem. This is not a one-time fix. It is a reinforcing cycle where each iteration strengthens the next, and the system learns faster with every pass.

A critical nuance: this loop does not require perfection at any layer before it begins operating, but it does require that each layer exist. Governance does not need to be complete before policy can improve. Technology deployment does not need to wait for policy to be fully aligned—but it must operate within an established governance boundary, with feedback mechanisms that surface misalignment as the system learns. The loop starts wherever the organization is today, and each iteration through it strengthens the layer below. The strongest version of this argument is not governance-first-then-technology. It is governance-informed deployment—where each deployment cycle generates governance intelligence that was previously invisible. Organizations building governance through operation, rather than before it, are the ones compounding fastest.



The theoretical foundation introduced above bears elaboration. Senge's reinforcing loops are not merely a metaphor; they describe a specific dynamic in which small initial improvements create conditions for larger subsequent ones.[7] Argyris's double-loop framework, applied here, means that the compounding loop does not merely correct individual claims faster—it systematically interrogates the assumptions embedded in governance, policy, and technology with each cycle.[8] Most payment integrity operations today are trapped in single-loop learning, identifying and recovering overpayments without interrogating the governance and policy structures that produced them. The compounding loop this paper describes is an engine for double-loop learning: each iteration does not merely fix a claim; it surfaces a systemic insight that prevents future claims from needing to be fixed.

The trust dimension is equally well-documented. Mayer, Davis, and Schoorman's integrative model of organizational trust identifies three antecedents: ability, benevolence, and integrity.[9] In the context of payment integrity, transparency directly builds perceptions of integrity and ability. Amy Edmondson's research on psychological safety confirms the operational mechanism: in organizations where transparency is low and trust is fragile, errors are hidden, feedback is delayed, and the system's ability to learn is throttled.[10] The compounding loop begins with transparency precisely because transparency creates the conditions under which errors can surface without triggering defensive routines.

The lean healthcare tradition provides the operational proof. Virginia Mason Medical Center adopted the Virginia Mason Production System and, through Rapid Process Improvement Weeks, achieved measurable reductions in inventory, floor space, lead time, and staffing

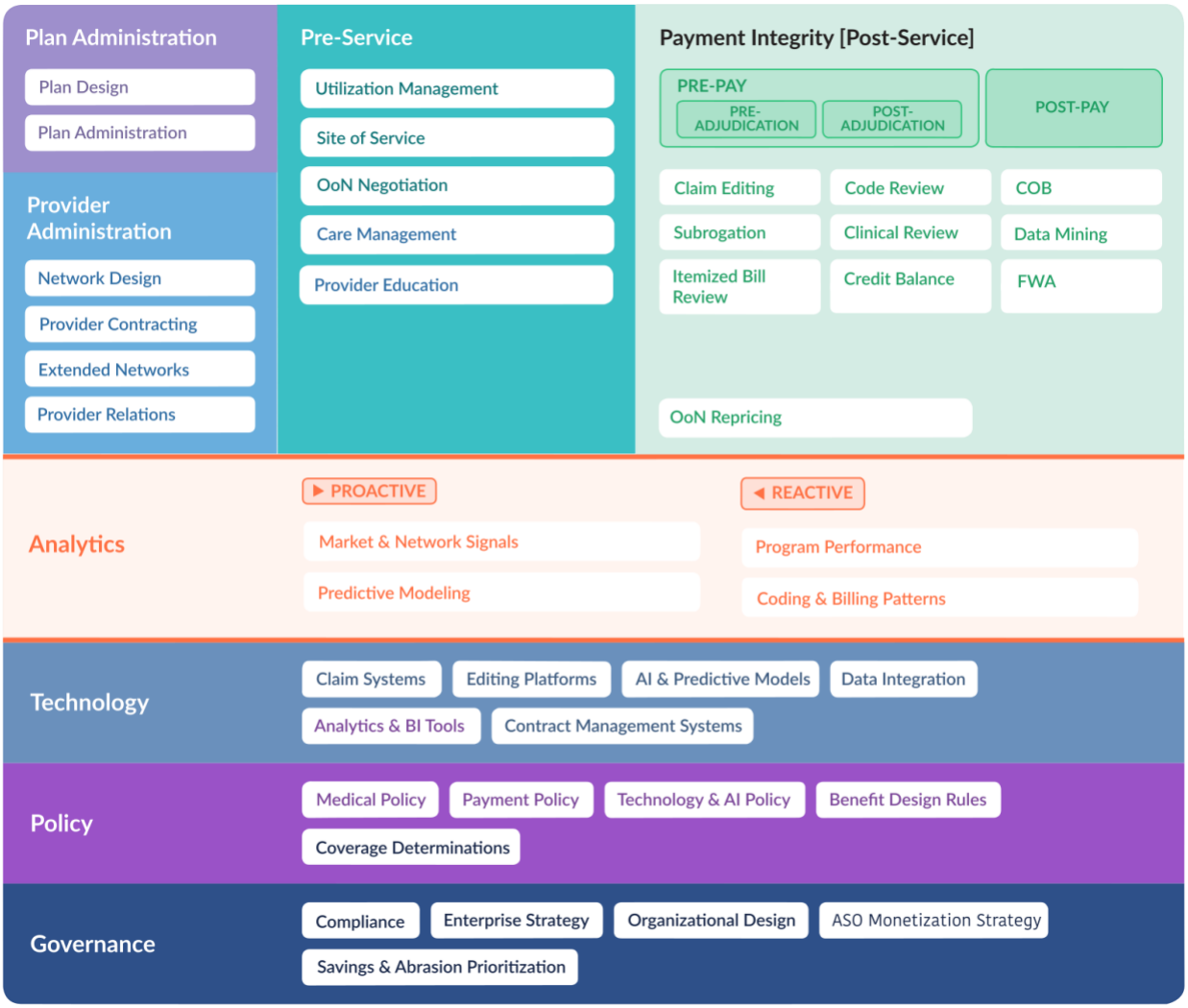
requirements. ThedaCare followed a similar trajectory, producing millions in annual savings through event-based improvement cycles. These organizations succeeded because they built feedback systems that allowed the organization to see itself clearly and act on what it saw.[30]

The organizations that move fastest are the ones that can see most clearly, and the ones that see most clearly are the ones that invested in transparency first.

III. Reading the Roadmap Bottom-Up

The 13point1 Advisors Reference Framework presents the foundational layers of cost containment in the sequence they must be addressed: bottom-up, not top-down. The framework's layered architecture runs from governance through policy, technology, and analytics to the operational landscape. Each layer depends on the one below it. Misalignment at any layer propagates upward, compounding through every function it touches.

This section walks each layer in sequence, with the supporting evidence that establishes why the order matters.



A. Governance — The Foundation No One Wants to Prioritize



The governance layer encompasses compliance, enterprise strategy, organizational design, administrative services only (ASO) monetization strategy, and the prioritization framework for balancing savings against provider abrasion. These are not administrative overhead. They are the operating rules that determine what every layer above can actually accomplish.

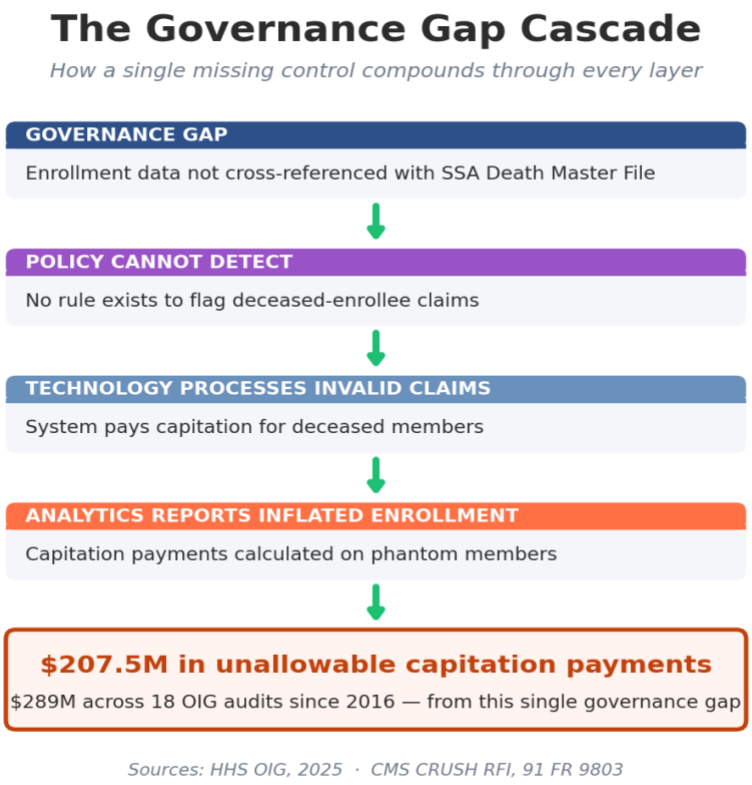
Cross-industry evidence consistently demonstrates that governance maturity, not technology investment, is the primary predictor of operational performance. The ISACA COBIT framework, the prevailing standard for IT governance worldwide, structures its assessment around capability levels ranging from incomplete to optimizing.[11] Research applying COBIT to enterprise transformation consistently finds that organizations at lower governance maturity levels realize significantly less return on technology investment than peers at higher levels, even when technology spending is equivalent. ISACA's 2025 guidance on AI governance reinforces the point: governance frameworks emphasizing continuous improvement are essential for evolving AI systems that must adapt to changing business environments.[12]

The healthcare-specific evidence is equally stark. A recent HHS Office of Inspector General (OIG) audit found that Medicaid agencies made an estimated $207.5 million in unallowable capitation payments to managed care organizations on behalf of deceased enrollees, not because of malicious intent, but because governance systems did not cross-reference enrollment records with the Social Security Administration's Death Master File.[13] Since 2016, OIG has conducted 18 such audits identifying roughly $289 million in improper payments from this single governance gap alone.

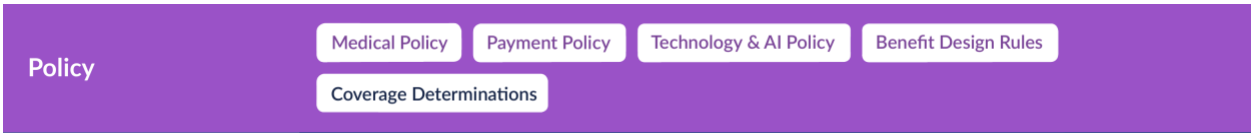
The Government Accountability Office's independent audit of the federal government's FY 2024 consolidated financial statements identified two continuing material weaknesses: the government's inability to determine the full extent to which improper payments occur and to reasonably assure that appropriate actions are taken to reduce them. These are governance failures at the highest level of the system, and they cascade downward through every program, plan, and payment.[31]

CMS's own regulatory actions confirm the diagnosis. The CRUSH RFI explicitly solicits feedback on strengthening program integrity oversight, including provider enrollment screening, payment suspension authority, and data analytics within a governance context.[3] It identifies the Medicare Advantage preclusion list gap: providers revoked from Traditional Medicare shifting their billing to Medicare Advantage (MA) plans because governance systems do not cross-reference enrollment status. CMS is not merely acknowledging governance failure. It is soliciting solutions.

Misaligned governance means every downstream decision inherits that misalignment. The cost is invisible until it compounds into systemic dysfunction. Governance is not a compliance exercise. It is the operating system—the foundational layer of decision-making authority, accountability, and process ownership on which policy, technology, and analytics depend. Everything above it runs on the rules set here.



B. Policy — Translating the Aspirational State into Operational Rules



If governance sets the destination, policy draws the route. The policy layer encompasses medical policy, payment policy, technology and AI policy, benefit design rules, and coverage determinations. Aligning governance and policy to the aspirational state is the first step, before technology enters the conversation.

The velocity of policy change in healthcare is extraordinary. CMS issues hundreds of transmittals, change requests, and manual updates annually, each carrying implications for how claims are adjudicated, how edits fire, and how payment rules are applied across lines of business. When governance structures are static while policy is dynamic, organizations build an ever-widening gap between their operating rules and the rules that actually govern their business.

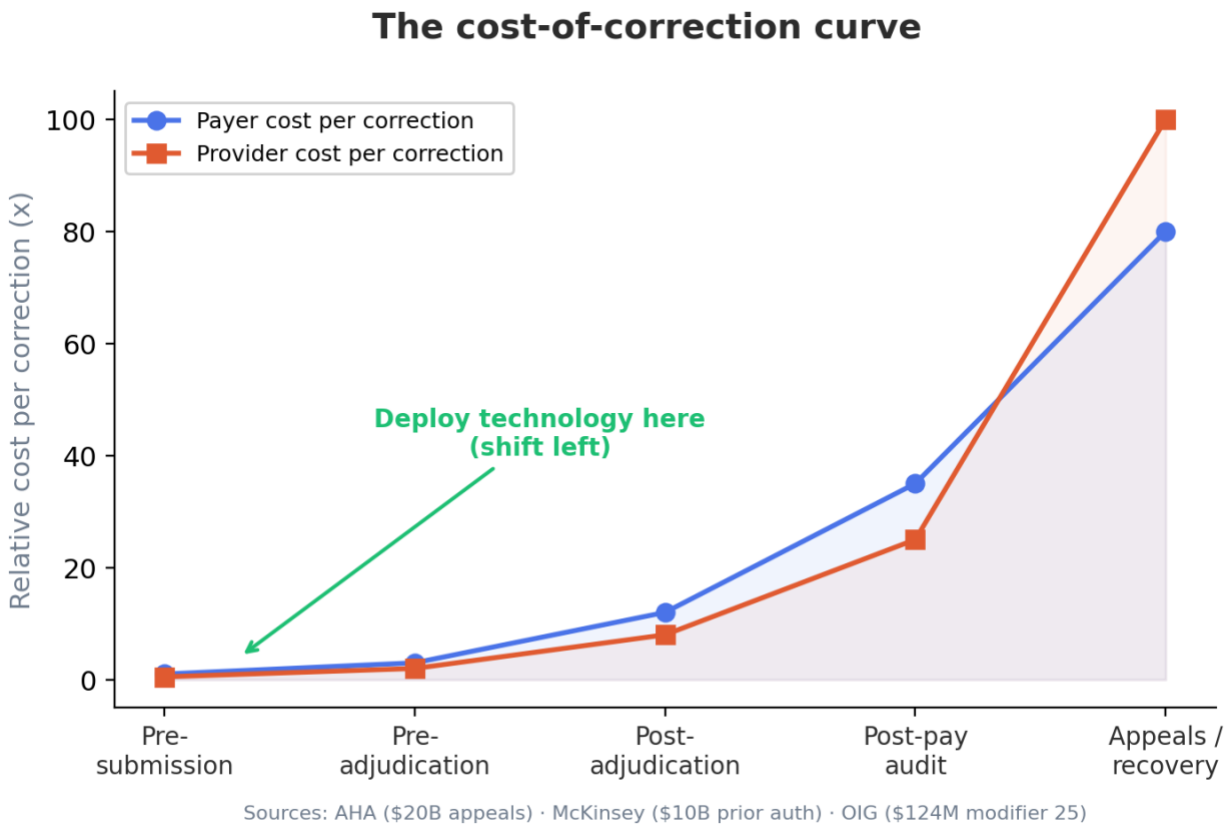
The OIG audit portfolio quantifies the cost. A 2025 audit of evaluation and management services billed with modifier 25 on the same day as intravitreal injections found that documentation for 22 of 24 sampled claims did not support the modifier's use, yet CMS paid $124 million nationwide because internal controls were not adequate to detect the pattern.[14] This is not fraud. It is policy misapplication at scale, the predictable consequence of policy rules outpacing governance alignment.

The architectural end state is **machine-native policy**: policy authored from inception for computation, where the human-readable document and the machine-readable criteria are synchronized outputs of a single structured source of truth. A change to clinical criteria propagates simultaneously to both the policy document a medical director reads and the claim edit that enforces reimbursement, eliminating the disconnect between clinical intent and payment execution that produces the misapplication documented above. The architecture is operational today within targeted use cases, but it is not yet, at industry scale, how the bulk of payer policy is authored. The vast majority of payer policy still exists as documents written for human reviewers—PDFs, contract clauses, clinical guidelines—and the corpus is too large, too embedded in vendor and provider relationships, and too operationally consequential to be rewritten wholesale on any schedule that meets payers' current cost-pressure timelines. Solving for that corpus is its own architectural decision, and one this section returns to.

The CRUSH RFI surfaces a second failure mode: policy absence rather than policy misapplication. Medicare Part B spent $8.4 billion on clinical lab tests in 2024. Genetic tests represented only 5 percent of tests but 43 percent of spending at $3.6 billion. OIG, DOJ, and CMS have issued multiple fraud alerts and enforcement actions, yet the fraud persists because policy mechanisms have not kept pace with billing patterns. The RFI asks about the MolDX program, noting that other payers and MA organizations require MolDX registration as a governance control for molecular diagnostics. CMS is asking why it does not.[3] This is a known fraud vector with known controls available, unimplemented because governance has not mandated the policy.

Hospitals bear the bilateral cost. McKinsey has estimated that hospitals spent $10 billion annually dealing with insurer prior authorizations.[15] A 2023 Premier study found hospitals spend just under $20 billion annually appealing denials, more than half wasted on claims that should have been paid initially. Claims denials increased 20.2 percent in 2023, and payer processing time increased 19.7 percent in the same year.[32] Policy without aligned governance produces rules that optimize for the wrong outcomes. The sequence matters.

The industry's path forward is to pursue payment integrity solutions where applied criteria are transparent and traceable by design. This is a governance decision, not a vendor selection: the ground rule for participation in a healthcare payment transaction must be that every claim disposition traces to specific criteria, definitions, and supporting literature. "Black box" edits—where the policy language behind a decision is unavailable—and "gotcha" edits that surprise anyone familiar with the underlying evidence must give way to structured traceability from edit, to criterion, to clinical guideline or evidence review, to the rationale for the coverage position. When traceability is a structural property of the policy layer rather than a reporting feature layered on after the fact, the bilateral dispute shifts from "was this edit legitimate" to "does this evidence support this position"—a fundamentally different and less expensive conversation for both parties.



C. Technology — Built on Top, Not Before



Technology accelerates whatever it sits on top of. If governance and policy are misaligned, technology accelerates the misalignment.

Erik Brynjolfsson's research provides the foundational evidence.[6] His analysis of firm-level data found that the organizational factors that unlock the value of IT are costly and time-consuming, and that for every dollar of IT investment there are several dollars of organizational investments that, when combined, generate the large rise in measured firm productivity. Firms that coupled IT investment with decentralized, process-aligned work practices were approximately 5 percent more productive than firms with equivalent IT spending but traditional structures.

McKinsey's 2025 payer transformation analysis quantifies the magnitude when organizational readiness is present.[16] For every $10 billion of payer revenue, AI solutions could save $150 million to $300 million in administrative costs and $380 million to $970 million in medical costs, but only as part of their operational framework, not as standalone technology deployments. In aggregate, payers embracing AI within their operational framework could reduce administrative expenses by 13 to 25 percent and medical costs by 5 to 11 percent. The qualifier is the thesis: technology is an amplifier, not a foundation.

The traditional policy development process is fundamentally segmented: a clinical team writes a document, and a separate IT team translates that text into code and configurations. The disconnect between clinical intent and implementation is what makes maintenance time-consuming and error-prone. The first architectural response addresses the policy that already exists. Across the existing corpus—provider contracts, clinical guidelines, coverage criteria, non-standard formats, the vast majority of payer policy on any 6 to 12 month action horizon—full rewriting is not economically justifiable on the timelines payers face. Governance must therefore mandate translation as a core capability of every participant in the ecosystem: any selected technology must be capable of converting unstructured policy inputs into structured, machine-readable form, paired with human review where clinical or contractual judgment is required. The mandate is technology-agnostic: it specifies a capability, not a vendor, which is what makes it durable governance rather than vendor advocacy. The same architectural property—translation of unstructured inputs into computable rules, paired with expert review where judgment is decisive—is what defines a consolidated post-service payment integrity platform when it is built correctly, as Section IV examines.

The second architectural response addresses policy not yet written. For greenfield domains and narrow scopes, governance should mandate that all new policy be authored as machine-readable structured data from inception, so the human-readable document and the computable criteria are synchronized outputs of a single source of truth from the moment the policy is written. This is the cleaner architecture wherever it can be applied—it eliminates the translation step entirely and removes the disconnect between clinical intent and implementation at the source. The mandate does not require rewriting the existing corpus; it requires that the corpus stop growing through unstructured authorship.

This architectural principle is operational today in production at targeted scale, and through more than one path. Concert has published its authorship-at-source methodology and reports that health plans applying it have reduced prior authorization volume and complexity by up to 50 percent while lowering medical costs, across a footprint serving more than 30 million members.[22] A parallel approach addresses the policy corpus already in place: AI-driven policy interpretation at the point of review, in which unstructured policy stays in its native form and is interpreted by AI in real time at the moment of adjudication. The two paths carry different governance profiles — authorship-at-source delivers traceability at the source, while interpretation-at-review delivers reach across the corpus that already exists — and most payers will need both. What they share is the defining structural property: the policy layer ceases to be documentary and becomes operational.

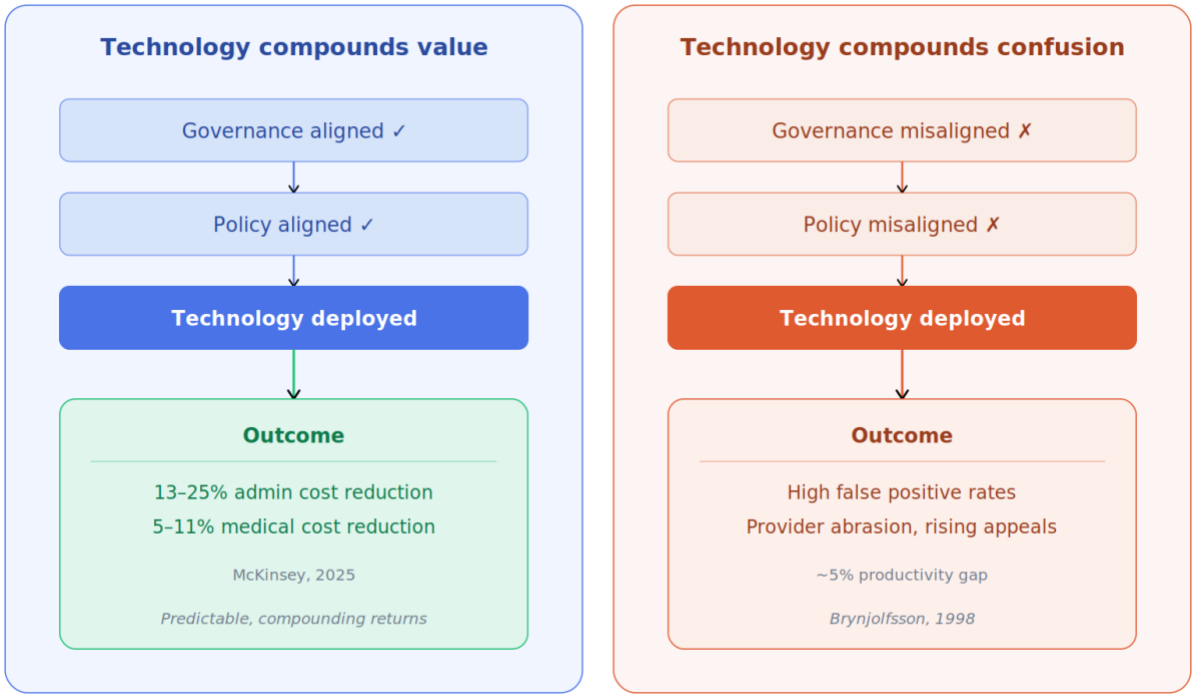
The CRUSH RFI validates this framing from the regulatory side.[3] Its questions about AI in coding oversight assume AI operates inside a governance framework: How should AI recommendations be displayed to human reviewers? What compliance risks should be mitigated? How do we prevent hallucinations? How do we incorporate coder feedback? CMS is not asking whether to deploy AI. It is asking how to govern AI deployment. The urgency is compounded by the reality that AI in 2026 does not operate within a single layer. AI agents are surfacing governance gaps, enforcing policy in real time, integrating data across systems, and generating forward-looking analytics simultaneously. When the foundation is aligned, this cross-layer capability multiplies value. When the foundation is misaligned, it multiplies misalignment across every layer at once.

> ***Technology built on a coherent foundation compounds value. Technology built on a misaligned foundation compounds confusion.***

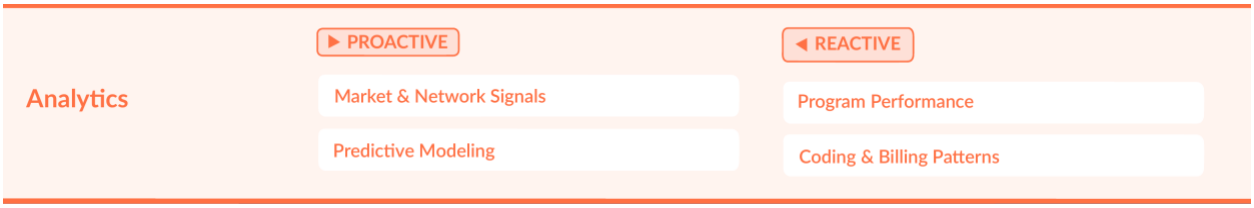
This is not a future-state argument. McKinsey's own data demonstrates the divergence: for every dollar invested in digital and AI transformation, payers with aligned governance and process foundations realize 13 to 25 percent administrative cost reductions and 5 to 11 percent medical cost reductions, while peers with equivalent technology spending but weaker foundations realize a fraction of that return. The competitive gap is observable and widening. Organizations that have not begun the governance work are not merely behind—they are compounding the gap.

The argument that governance is what distinguishes AI success from AI failure has been articulated by operators deploying AI-native payment integrity platforms at scale. Machinify, whose AI operating system processes more than serves 85 plus customers including 18 of the top 20 health plans, supporting 270 million member lives across commercial, Medicare, and Medicaid, and has delivered more than $6 billion in annual cost avoidance and recoveries to client health plans, has published the position directly: model quality is rarely the decisive variable.[21] Responsible AI principles only produce durable outcomes when enforced at the decision level—through systems where governance is built into how decisions are made, not bolted on after deployment. Where governance fragments across tools and teams, policies apply inconsistently, models drift as guidelines change, and documentation falls short under audit. The technology stack inherits whatever governance condition sits beneath it—and at scale, that inheritance is the single largest determinant of whether AI delivers compounding value or compounding cost.

If the amplifier principle holds—that technology compounds whatever foundation it sits on—then the design requirements for health plan AI infrastructure become specific. Four architectural properties distinguish systems that compound value from systems that compound fragmentation. Continuity means the system carries clinical and operational context forward across encounters and workflows; a prior authorization decision is not an isolated transaction but part of an ongoing narrative the system understands and builds on. Traceability requires that every AI-generated decision can be explained, reviewed, and audited—not merely for compliance, but to sustain the operational trust that clinicians and auditors require before relying on automated outputs. Proportionality matches the level of AI autonomy to the risk and complexity of each decision: routine tasks are automated, complex clinical judgments retain human oversight, and high-risk decisions trigger escalation. Upstream orientation redirects intervention earlier in the care lifecycle—identifying care gaps before they are missed, intervening before a denial occurs—reducing the need for downstream correction rather than accelerating it. Autonomize AI's Ganesh Padmanabhan, writing on what he terms "living intelligence," has proposed these four properties as the design requirements for healthcare AI systems that learn and adapt rather than merely execute.[17] The connection to the compounding loop is direct: continuity enables the feedback cycle to accumulate intelligence across iterations; traceability builds the trust that sustains the loop; proportionality ensures governance boundaries are respected as automation expands; and upstream orientation is the mechanism through which visibility converts to prevention.



D. Analytics — The Layer Everyone Wants to Start With



The analytics layer operates along two dimensions that are structurally significant. Proactive analytics, including market and network signals and predictive modeling, requires aligned governance and coherent technology to generate real signal. Reactive analytics, including program performance measurement and coding pattern analysis, is achievable with less foundational maturity but delivers diminishing returns without the context that governance and policy provide.

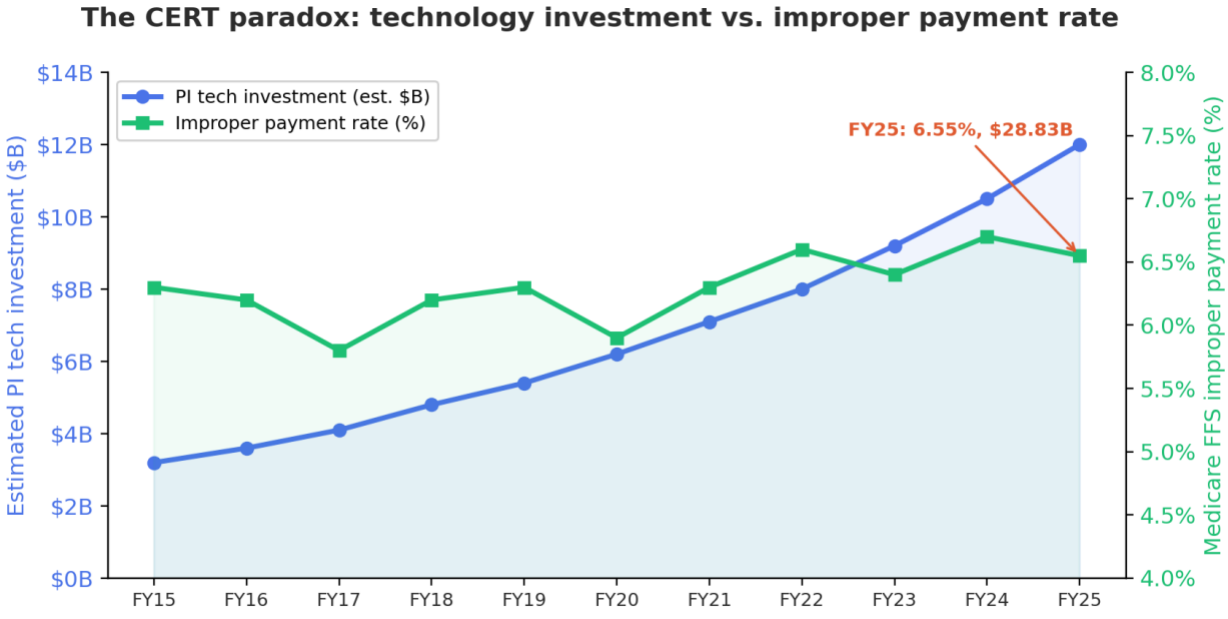
Thomas C. Redman's foundational research on data quality economics establishes the baseline: most organizations lose between 15 and 25 percent of operating revenue due to poor data quality.[18] In healthcare, where claims data passes through multiple systems, coding conventions, and policy filters before reaching an analytics engine, the compounding effect of upstream data quality issues is severe.

The CMS Comprehensive Error Rate Testing program provides the most telling evidence.[19] Despite continuous growth in technology investment across the healthcare industry, the Medicare Fee For Service (FFS) improper payment rate of 6.55 percent in FY 2025 has oscillated within a narrow band that has persisted for a decade. If technology alone drove accuracy, the trendline would slope downward in proportion to investment. It does not. The variable is not the analytics. It is the foundation underneath.

The most striking evidence comes from Obermeyer et al.'s landmark study, which demonstrated how governance assumptions embedded in data propagate through analytics into real-world harm.[20] A widely used commercial algorithm exhibited significant racial bias because it used health costs as a proxy for health needs, a governance decision. The authors estimated that reformulating the algorithm would increase the percentage of Black patients identified for extra care by more than half. This is not an analytics failure. It is a governance assumption, amplified by technology, producing systematic inaccuracy at population scale.

The CRUSH RFI acknowledges this dependency directly.[3] Its state-level Medicaid questions ask about using federal databases to move further away from a pay-and-chase model and towards pre-pay review. CMS is publicly acknowledging the reactive model's limitations and soliciting alternatives grounded in governance, not technology alone.

Analytics is the canary. If analytics is producing noise, high false positive rates, inconsistent savings projections, unpredictable model performance after policy changes, the problem is not the analytics. It is the foundation underneath. When governance, policy, and technology are aligned, analytics shifts from retrospective reporting to forward-looking intelligence.



If technology alone drove accuracy, the green line would slope downward. It does not.

IV. The Operational Landscape This Enables

The operational layers of the Reference Framework—plan administration, provider administration, pre-service, and payment integrity post-service—each carry their own specialization patterns. The central thesis holds at the operational level: no solution addresses every layer for every modality, and none needs to. What makes the landscape coherent is not tool count, but whether the tools operate on the foundation the preceding sections have described.

The top of the Reference Framework describes the operational functions that execute on the foundation: plan administration, provider administration, pre-service, and payment integrity. When the layers below are aligned, these functions transform from a collection of siloed activities into an integrated system where decisions at one point inform actions at every other.

Plan Administration & Provider Administration

These functions set the structural context: network design, provider contracting, extended networks, provider relations, and plan design. When governance is aligned, plan administration decisions are made with full visibility into their downstream payment integrity implications. When governance is misaligned, plan design creates complexity that the claims operation must absorb, and provider contracting creates ambiguity that the editing engine must interpret.

Provider abrasion, the measurable degradation of the payer-provider relationship caused by administrative burden, is the observable cost of operational misalignment. The American Hospital Association (AHA)'s 2024 Costs of Caring report quantified the scale: commercial health insurers' practices, including automatic claims denials and onerous prior authorization requirements, generated billions in administrative costs that divert dollars away from patient care. When the foundation is aligned, provider administration becomes a relationship management function. When it is not, it becomes a dispute resolution function.

Pre-Service

Utilization management, site of service, out-of-network negotiation, and care management represent the proactive levers, the intervention points where aligned governance and policy translate into better decisions before a claim is generated. The CMS Transparency in Coverage rule, effective July 2022, required non-grandfathered plans to post machine-readable files disclosing negotiated rates and out-of-network allowed amounts. CMS's December 2025 proposed rule identified three persistent barriers: inaccessibility, data ambiguity, and misalignment with the Hospital Price Transparency rule. Hospital compliance with federal price transparency regulations increased from 70.4 percent in 2021 to 87.7 percent in 2022 following increased financial penalties.[37] This regulatory trajectory toward structured, accessible, interoperable transparency is the ecosystem moving whether individual participants choose it or not.

The operational vision is concrete: a patient sitting with their provider should be able to ask whether a service will be covered, how much they will owe, and receive an accurate answer instantly. In published production deployments of machine-native policy for laboratory diagnostics, API and machine-readable criteria deliver coverage determinations in milliseconds, powering electronic prior authorization and pre-payment claim editing in real time.[22] This capability has enabled health plans to reduce prior authorization volume and complexity by 50 percent for genetic testing while simultaneously lowering medical costs, because transparent policy enforced at the point of service replaces the manual review process that generates the delays and denials documented elsewhere in this paper. The principle extends regardless of clinical domain: when policy is machine-native, pre-service determinations become automatable, and the administrative burden that the AHA and McKinsey have quantified begins to compress. The CMS interoperability rule (CMS-0057-F) requiring production-ready FHIR-based prior authorization by 2027 is the regulatory confirmation that this trajectory is not optional.[38]

The operational evidence for what aligned foundations produce in pre-service is now measurable. When AI-native systems operate on coherent governance and policy, prior authorization transforms from a manual bottleneck into a scalable, compliant operation. Published case studies document the pattern: a Fortune 100 health plan reduced prior authorization processing time by 55 percent—from 35 minutes to under 15—while cutting decision turnaround time by half and recouping more than 36,000 clinical hours per month for higher-value clinical work. Compliance improved 16 percent as real-time guideline adjudication eliminated rework. At Altais Health Solutions, turnaround time fell 63 percent and manual errors dropped 54 percent. In appeals and grievances—the downstream consequence of pre-service failures—a national Medicare Advantage plan reduced case preparation time by 80 percent, with 98 percent of AI-generated clinical summaries accepted by reviewing clinicians without modification. Autonomize AI, which deployed these implementations, reports a 3–5x aggregate return on investment across its production portfolio.[23] These outcomes are consistent with the paper's central argument: the returns are not attributable to the technology alone. They emerge because governance, policy, and technology were aligned before the system began operating—and each cycle through the loop compounds the accuracy of the next.

Payment Integrity (Post-Service)

Payment integrity spans the pre-pay/post-pay continuum with specific functions: claim editing, code review, coordination of benefits, subrogation, clinical review, data mining, itemized bill review, credit balance recovery, and fraud, waste, and abuse detection.

The payment integrity market illustrates the consequences of operating without foundational alignment. Health plans typically engage seven to thirteen vendors for payment integrity functions, each contracted in isolation, each optimizing for its own savings metric. The result is overlapping audits, conflicting edit logic, and a total program cost that is invisible to any single function. The CRUSH RFI surfaces a related fragmentation: Durable medical equipment, prosthetics, orthotics, and supplies (DMEPOS) payments exceed $7 billion annually in Original Medicare alone, and OIG and MA organizations have identified non-participating DMEPOS suppliers fraudulently billing MA plans for millions because enrollment systems are not integrated. CMS is asking whether MA organizations would prefer suppliers to be accredited and enrolled similar to Original Medicare. This is an integration question, not a technology question.

When the foundation layers are aligned, payment integrity becomes the intelligence layer, not merely recovering money that should not have been paid, but feeding insights upstream into governance, policy, and pre-service functions that prevent the error from recurring. This is the structural difference between payment integrity as a cost center and payment integrity as a strategic capability.

When foundation alignment is achieved, post-service payment integrity consolidates. Rather than coordinate a portfolio of seven to thirteen point solutions, the payer operates a single intelligence layer — what Machinify has publicly positioned as an operating system for payers — that evaluates and prices claims at machine speed and routes exceptions to expert review. Published outcomes at a top-ten payer illustrate the operational shift: the system automates 40 percent of recoveries, cuts claims review time by more than half, doubles reviewer capacity, eliminates false positives at scale, and measurably reduces payer dependence on fragmented

vendor ecosystems.[24] The architecture is not a replacement for human expertise but a pairing with it. Machinify CEO David Pierre has described the operating model as "algorithms, machine learning, and then our subject matter experts — attorneys, analysts, clinical specialists" engaging together on claims that require specialized judgment.[39] Ryan Little, president of Right Payer Solutions, frames the discipline more plainly: the right approach is to "automate drudgery, but don't replace expertise."[40] This is the actionable state today — the interim architecture that moves payment integrity forward without waiting for perfect conditions. Human judgment remains decisive where it matters; routine volume moves at machine speed within codified policy guardrails; and each cycle of paired AI-plus-expert review generates governance intelligence that expands what can be safely automated next. The architectural implication is that the seven-to-thirteen-vendor pattern documented earlier in this paper is not the natural state of payment integrity—it is the artifact of a decade of point-solution acquisition on top of misaligned foundations. When the foundations are aligned, payment integrity stops being a portfolio of disconnected recovery mechanisms and becomes a single intelligence layer feeding insights back into policy and governance. The shift is from vendor consolidation as a procurement problem to vendor consolidation as an architectural outcome.

V. The Compounding Effect

The evidence across all four layers converges on a single thesis. When governance is aligned, policy becomes coherent rather than contradictory, technology amplifies that coherence rather than inheriting misalignment, and analytics finally produces signal instead of noise. None of these layers delivers value in isolation—and no single tool within any layer delivers the full answer either. The economics of care improve when a portfolio of solutions operates on a governance, policy, and architectural foundation coherent enough to compound their outputs rather than fragment them. Each iteration through the compounding loop strengthens every layer simultaneously increasing trust, shortening cycles, and surfacing the next priority the organization could not previously see.

Dierickx and Cool's foundational research on capability accumulation provides the theoretical anchor: organizational capabilities compound because each increment builds on the last, and the time-compressed nature of this accumulation creates sustainable competitive advantage that cannot be replicated through spending alone.[25] Teece, Pisano, and Shuen's dynamic capabilities framework maps directly to this loop: the capacity to sense opportunities, seize them through aligned resources, and transform organizational structures is precisely what the governance-to-analytics sequence enables.[36]

James March's exploration-exploitation framework completes the picture. Organizations that can exploit their existing infrastructure faster outperform in the near term while building the foundation for long-term exploration. But March warned that adaptive processes, by refining exploitation more rapidly than exploration, are likely to become effective in the short run but self-destructive in the long run.[33] The compounding loop this paper describes addresses March's paradox: by continuously surfacing the next priority through transparency and trust, it ensures that exploitation does not calcify into complacency. The system learns faster because it can see itself.

> ***The compounding is not in the savings. It is in the capability.***

VI. Implications

The argument of this paper carries specific implications for three audiences. Each should see a clear next step, not a prescription, but a logical consequence of the analysis.

For Payers

The first move is not a technology investment. It is a governance audit. Understanding where governance and policy are misaligned reveals why downstream investments have underperformed. The National Association of Corporate Directors has long advocated governance assessment as a prerequisite to strategic investment,[34] and the principle translates directly: before committing capital to AI-powered payment integrity, a payer should be able to articulate its governance maturity level, identify the specific misalignments between governance and policy, and quantify the downstream cost of those misalignments in terms of false positive rates, appeals volume, and total program cost.

Deloitte's Center for Health Solutions has published payer operating model maturity assessments that validate governance as the starting point.[26] McKinsey's analysis reinforces the point: payers that have already undertaken administrative transformations still have significant opportunities to optimize, but only if the optimization is built on aligned governance and policy.[16] The governance audit is the beginning of the compounding loop. Critically, this does not mean delaying technology deployment until governance is perfected. Governance-informed deployment—where each AI implementation generates governance intelligence that was previously invisible—is the operational model that compounds fastest. The audit reveals where to start. The deployment reveals what the audit missed. The loop accelerates from there.

Governance also establishes the terms of participation for outside parties in the solution. Vendors, solution providers, and data partners operate inside a governance boundary the payer defines—not the other way around. Requirements for transparency, traceability, data stewardship, and policy alignment become pre-conditions for engagement rather than negotiated concessions. Without these requirements specified as governance outputs, every new vendor relationship reintroduces the fragmentation the organization is trying to resolve.

For Solution Providers

Selling into misaligned governance creates fragile relationships and unpredictable outcomes. The same payment integrity technology deployed at two different health plans can produce dramatically different results, not because the technology varies, but because the governance foundation varies. Solution providers who recognize this reality and help payers address foundational alignment build durable partnerships with predictable performance. Those who sell technology into governance vacuums will continue to experience the cycle of pilot, underperformance, and replacement that characterizes the current vendor landscape.

As Erica Dhawan argues in her work on connectional intelligence, sustainable collaboration requires building trust through visible valuing, careful communication, and confident

collaboration.[27] Solution providers who invest in helping their clients achieve governance maturity are not giving away competitive advantage. They are building it.

Solution providers operating in this environment are publishing their own guidance on governance-informed deployment. Ryan Little, President of Right Payer Solutions at Machinify, has articulated a disciplined sequence: start small, prove the case, then scale. His framing—"test before you tell, measure before you mandate, remove friction before adding ambition"—maps directly to the foundation-before-amplifier principle this paper has traced throughout.[28] Little echoes a well-established research finding: approximately 70 percent of large-scale transformations fail — not from technology limitations, but from insufficient leadership conviction, organizational engagement, execution discipline, and follow-through.[35] The implication for solution providers is that the most durable deployments are built through disciplined piloting with measurable outcomes—and the vendors who respect this discipline tend to become long-term partners rather than replaced point solutions.

For the Ecosystem

Transparency is not a threat to any participant acting in good faith. It is the prerequisite for sustainable economics. The regulatory trajectory confirms this. The No Surprises Act, the CMS Hospital Price Transparency rule, and the Transparency in Coverage rule are moving the ecosystem toward structured, accessible, interoperable pricing data whether individual participants choose it or not. Hospital compliance with price transparency increased from 70.4 to 87.7 percent after increased penalties.[37] CMS's December 2025 proposed rule would further strengthen requirements. The CRUSH RFI, with its comment period having closed on March 30, 2026, adds enforcement mechanisms for provider enrollment, payment suspension, AI governance, and cross-program integrity.[3] This is active rulemaking, not directional signaling.

The operators building infrastructure compatible with this sequence are already publishing their experience. Concert has operationalized machine-native policy at scale. Autonomize has described the architectural properties required for AI to learn rather than merely execute. Machinify has demonstrated what integrated AI infrastructure produces when post-service intelligence feeds upstream into governance and policy. None of these organizations covers the full landscape alone—and none needs to. What they share is a recognition that governance is the precondition, not a downstream compliance exercise. The next paper in this series, *Beyond Pre-pay and Post-pay: Building Flexible Architecture for First-Time Accuracy*, examines what that flexible architecture looks like in practice—how payers design infrastructure that accommodates the modalities they must address today while remaining adaptable to those they will need to address tomorrow.

The ecosystem is already deciding this question. The organizations that align their governance, policy, and technology to the transparency trajectory will compound capability while their peers compound friction. Reinforcing loops do not wait for consensus to begin operating.

References

1. Himmelstein, D.U., Campbell, T. & Woolhandler, S. "Health Care Administrative Costs in the United States and Canada, 2017." Annals of Internal Medicine, 172(2), 2020. https://doi.org/10.7326/M19-2818
2. Jiwani, A. et al. "Billing and Insurance-Related Administrative Costs in United States Health Care." BMC Health Services Research, 14(556), 2014. https://doi.org/10.1186/s12913-014-0556-7
3. CMS. "Request for Information (RFI) Related to Comprehensive Regulations To Uncover Suspicious Healthcare (CRUSH)." 91 FR 9803 (CMS-6098-NC). February 27, 2026. https://www.federalregister.gov/documents/2026/02/27/2026-03985
4. CMS. "FY 2025 Improper Payments Fact Sheet." January 2026. https://www.cms.gov/data-research/monitoring-programs/improper-payment-measurement
5. AHA (American Hospital Association). "2024 Costs of Caring." April 2025. https://www.aha.org/costsofcaring
6. Brynjolfsson, E. "Beyond the Productivity Paradox." Communications of the ACM, 41(8), 1998.
7. Senge, P.M. The Fifth Discipline: The Art and Practice of the Learning Organization. Doubleday, 1990.
8. Argyris, C. & Schön, D. Organizational Learning: A Theory of Action Perspective. Addison-Wesley, 1978.
9. Mayer, R.C., Davis, J.H. & Schoorman, F.D. "An Integrative Model of Organizational Trust." Academy of Management Review, 20(3), 1995.
10. Edmondson, A.C. "Psychological Safety and Learning Behavior in Work Teams." Administrative Science Quarterly, 44(2), 1999.
11. ISACA. "Effective Capability and Maturity Assessment Using COBIT 2019." 2020.
12. ISACA. "Leveraging COBIT for Effective AI System Governance." 2025.
13. HHS OIG. "Medicaid Agencies Made Millions in Unallowable Capitation Payments to MCOs for Deceased Enrollees." 2025.
14. HHS OIG. "Medicare Payments for E&M Services Provided on the Same Day as Eye Injections." May 2025.
15. McKinsey & Company. "2024 Payers Outlook: Opportunities Abound." December 2023.
16. McKinsey & Company. "Rewiring Healthcare Payers: A Guide to Digital and AI Transformation." March 2025.
17. Padmanabhan, G. Autonomize AI published perspectives: "The Unseen Shift: Why AI-Native Is the Only Future for Healthcare" and "Compliance Is the Baseline. Advantage Is the Opportunity." 2026. autonomize.ai/blog.
18. Redman, T.C. Data Quality: The Field Guide. Digital Press/Elsevier, 2001.
19. CMS. CERT (Comprehensive Error Rate Testing) Report, FY 2025. https://www.cms.gov/data-research/monitoring-programs/improper-payment-measurement/cert
20. Obermeyer, Z. et al. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." Science, 366(6464), 2019. https://doi.org/10.1126/science.aax2342
21. Machinify. "Responsible AI in Healthcare: Why Governance, Transparency, and Control Matter." March 2026. machinify.com/resources/responsible-ai-in-healthcare.
22. Concert. "Policy That Computes." Version 1.0, March 2026. concert.co. https://concert.co
23. Autonomize AI. Published case studies: "How a Fortune 100 Health Plan Cut Prior Authorization Processing Time by 55%" (2025); "How Altais Cut Prior Authorization Review Time by 45%" (2026); "How a National MA Plan Cut Appeals & Grievances Case Preparation Time by 80%" (2026). autonomize.ai/case-studies.
24. Machinify. Company disclosures, product documentation, and client testimonials. machinify.com, 2025–2026.
25. Dierickx, I. & Cool, K. "Asset Stock Accumulation and Sustainability of Competitive Advantage." Management Science, 35(12), 1989.
26. Deloitte Center for Health Solutions. "Future of Health." 2025. https://www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/future-of-health.html
27. Dhawan, E. Get Big Things Done: The Power of Connectional Intelligence. St. Martin's Press, 2015.
28. Little, R. (Machinify). "How Healthcare Leaders Can Adopt New Technology Without Disrupting Operations." March 2026. machinify.com/resources/adopt-new-tech-without-disrupting-ops.

29. Sahni, N.R., Carrus, B. & Cutler, D.M. "Administrative Simplification and the Potential for Saving a Quarter-Trillion Dollars in Health Care." JAMA, 326(17), 2021. https://doi.org/10.1001/jama.2021.17315
30. Kenney, C. Transforming Health Care: Virginia Mason Medical Center's Pursuit of the Perfect Patient Experience. CRC Press, 2010.
31. GAO. "Financial Audit: FY 2024 and FY 2023 Consolidated Financial Statements of the U.S. Government." GAO-25-106928. February 2025.
32. Premier Inc. "Trend Alert: Private Payers Continue Gamesmanship Despite Calls for Reform." March 2024. premierinc.com/newsroom
33. March, J.G. "Exploration and Exploitation in Organizational Learning." Organization Science, 2(1), 1991.
34. NACD (National Association of Corporate Directors). "Governance Practices for a Changing Environment." 2023. nacdonline.org
35. Robinson, H. "Why do most transformations fail? A conversation with Harry Robinson." McKinsey & Company, July 2019. mckinsey.com/capabilities/transformation/our-insights
36. Teece, D.J., Pisano, G. & Shuen, A. "Dynamic Capabilities and Strategic Management." Strategic Management Journal, 18(7), 1997. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z
37. Kong, E. & Ji, Y. "Provision of Hospital Price Information After Increases in Financial Penalties for Failure to Comply With a US Federal Hospital Price Transparency Rule." JAMA Network Open, 6(6), 2023. https://doi.org/10.1001/jamanetworkopen.2023.20694
38. CMS. "Interoperability and Prior Authorization Final Rule." CMS-0057-F. January 17, 2024. https://www.cms.gov/newsroom/fact-sheets/cms-interoperability-and-prior-authorization-final-rule-cms-0057-f
39. Pierre, D. (Machinify). Interview with Jonathan Palmer, Vanguards of Health Care by Bloomberg Intelligence. July 10, 2025. Transcript preview at machinify.com/resources/how-machinify-s-ai-platform-is-rewriting-the-rules-of-payer-ops.
40. Little, R. (Machinify). “Why AI in Healthcare Operations Fails Without Experience.” March 26, 2026. machinify.com/resources/ai-in-healthcare-operations-requires-experience.