AI: A Case for Safety & Operational Excellence

Seizing the Agentic-AI Advantage Before the Gap Widens: The Operationally-Owned Management System as Strategic Lever

Bryan G. McMurray

June 2026

The views expressed in this paper are the author's own and do not necessarily reflect those of any organization with which the author is affiliated.

Abstract

Agentic AI is emerging as a powerful multiplier on operational performance, but the multiplier only works where operations are already run as a disciplined, documented system. Across industries, that discipline already lives in established management systems that drive performance, reliability, and risk reduction. In oil and gas, these include API RP 1173 (Pipeline Safety Management Systems), OSHA PSM (29 CFR 1910.119), IOGP Operating Management System frameworks, and ISO 9001/14001. Analogous frameworks govern information technology (ITIL, ISO/IEC 20000), information security (ISO 27001, NIST Cybersecurity Framework), financial controls (COSO, Sarbanes-Oxley), and engineering, legal, and small-business operations. Whatever the industry-specific vocabulary, these systems share a common structure: documented procedures, clearly defined measurable outputs with performance tolerances, root cause analysis following failures, and proactive audits to identify gaps before they result in adverse events.

However, traditional approaches remain resource-intensive and episodic, with major compliance audits often occurring on multi-year cycles. This limits the frequency and effectiveness of continuous improvement.

Agentic AI systems introduce a fundamentally new capability. These autonomous agents operate in closed-loop improvement cycles: planning and acting against established standards, observing real-time data from operational systems, detecting deviations or procedural shortcuts, performing RCA-style analysis, implementing or recommending improvements, and consolidating learnings through background memory processes. When bolted onto mature, operationally-owned management foundations, agentic systems enable continuous (versus episodic) assurance and system-level self-improvement while preserving essential human oversight for consequential decisions and regulatory compliance.

This paper argues that agentic AI is opening a fast-moving and widening competitive gap, and that the factor most determining which side of that gap an organization lands on is whether it has a strong, documented, operationally-owned management system for the agents to run on. Organizations that have built one, in any industry, are positioned to convert the agentic opportunity into compounding performance gains. Those that have not face an expensive, slower, lower-success scramble to build the foundation while better-prepared competitors pull away. A practical roadmap for implementation, education needs, governance considerations, and the competitive urgency of acting now is provided, with examples drawn primarily from midstream oil and gas operations where the management-system standards are most mature and the case evidence most detailed.

Keywords: Agentic AI, Competitive Advantage, AI Readiness, Operationally-Owned Management Systems, Closed-Loop Improvement, Operational Excellence, Continuous Improvement, API RP 1173, OSHA PSM, ISO 9001, ISO 27001, NIST CSF, ITIL, COSO

1. Introduction

Every organization that produces a result against a defined standard runs a management system. The accountant reconciling a ledger, the engineer designing a pressure vessel, the lawyer managing a litigation matter, the IT operator restoring a service, the security analyst responding to an intrusion, the small-business owner closing the books, and the midstream operator running a fractionator are all doing the same thing in their respective vocabulary: executing a procedure with a defined output against a defined measurement, with some kind of feedback loop when the output drifts from expectation. The most disciplined industries have codified those management systems into formal standards. The argument of this paper is that agentic artificial intelligence has turned that management system from a back-office discipline into one of the most decisive determinants of competitive advantage: it is the substrate the agents run on, so the organizations that own a mature one stand to capture the agentic opportunity and compound it, while those that do not fall rapidly and measurably behind.

The paper develops the argument using the case examples where the management-system standards are most mature and the data substrate most operationally detailed: midstream oil and gas operations, where the standards are codified, the operational data is abundant, and the early agentic case studies in adjacent industries are well-documented. The mechanics, however, are not midstream-specific. They apply wherever a documented standard, a defined measurement, a structured data substrate, and a culture of review exist. They apply to information technology operations under ITIL or ISO/IEC 20000, to information security under ISO 27001 or the NIST Cybersecurity Framework, to financial reporting and internal controls under COSO and Sarbanes-Oxley, to product engineering and project delivery under PMI or IEEE-style standards, to legal operations under matter-management frameworks, and to small-business operations wherever a disciplined set of operating procedures exists. Wherever the management system exists, the agentic opportunity exists. Wherever the management system is owned by the operating function that actually runs the work, the agentic opportunity is immediate. Wherever it is not, the opportunity is contingent on first putting the operating function in ownership.

By an *operationally-owned management system*, this paper means a documented body of standards, measurements, and review processes that is owned, used, and continuously refined by the operating organization, the people who actually do the work, rather than parked in a compliance binder maintained by a separate audit, ESH, quality, or controls function on a regulatory schedule. The standards are owned by the people who execute them. The measurements are owned by the people who run the equipment, write the code, close the books, ship the brief, or build the design. The audit is something the operating organization does to itself because the audit is how the operating organization runs the work, not because a regulator requires it. When ownership is in the operating function, the management system drives operational performance, and operational performance, in turn, produces the reliability, throughput, cost discipline, environmental performance, safety, data integrity, audit-readiness, security posture, or client outcomes the organization is in business to deliver. Those outcomes are not separate programs. They are *consequences* of an operating organization running its work to standard. Operational excellence, in the sense this paper means it, is not a separate program either; it is what an operating function produces when it owns and runs that management system.

The midstream sector is a useful demonstrated case because the industry has spent three decades building the *content* of its management systems. API Recommended Practice 1173 (American Petroleum Institute, 2015), OSHA's Process Safety Management standard at 29 CFR 1910.119, the IOGP Operating Management System Framework (International Association of Oil & Gas Producers, 2014a, 2014b), and the ISO 9001 and 14001 quality and environmental standards together encode the operating philosophy: define the standard, measure the result against a tolerance, investigate the deviation, audit for gaps, and improve. The frameworks are correct, and they have moved the industry forward. What separates the operators at the leading edge of operational performance from the operators who are not is not which standard they cite. It is who owns the system in their organization. The same pattern is observable across industries: organizations whose operating functions own ISO 27001, ITIL, or COSO as live operating practice outperform organizations that delegate those frameworks to internal audit or compliance groups outside the operating line.

The evidence that the industry has reached the ceiling of what its current implementation model can produce is in the federal data. PHMSA's twenty-year trend data shows that serious pipeline incidents have declined only slightly, while other reported incidents that damage the environment or property have increased (Pipeline and Hazardous Materials Safety Administration [PHMSA], n.d.). Independent analyses of the same federal data quantify the plateau: U.S. pipelines have averaged roughly 628 reported incidents annually since 2010, about 1.7 per day, with little movement in the recent rate (FracTracker Alliance, 2025), and the Congressional Research Service records an average of about eleven fatalities and forty-eight injuries a year from roughly twenty-six serious pipeline incidents annually over 2014–2024 (Congressional Research Service, 2025). The numbers describe an industry whose management systems have captured most of the gain available at the cadence and the ownership model under which they currently operate.

Agentic artificial intelligence changes both the cadence and the ownership model. Agentic systems are autonomous, goal-directed software systems that plan, act, observe, and adapt against a defined objective and a defined set of tools and data (Anthropic, 2024; OpenAI, 2024). Their defining behavior is a closed self-correction loop: the agent iterates against its own output, refines its approach, and consolidates what it learned for the next cycle. When that loop is layered onto a live, operationally-owned management system, the result is not a replacement of the management system but a multiplier on it. The same standards, the same tolerances, the same audit logic, the same RCA discipline, now executed continuously against real-time operating data, by an agent that is supervised by the operations organization that owns the standard. This is the central claim of the paper. The structural prerequisite for adopting agentic AI in any operation is the same thing that produces operational performance in the first place: a management system that the operating function owns and uses. Organizations that have that asset are positioned to multiply it. Organizations that do not are positioned to fall further behind across every outcome that the management system was supposed to produce, while the cost of catching up rises. The paper proceeds by examining the foundations of operationally-owned management systems and what makes them effective (Section 2), the limits of the implementation model the industry has converged on (Section 3), the mechanics of the agentic opportunity (Section 4), the resulting competitive asymmetry (Section 5), and a practical path forward (Section 6).

A note on what is and is not new in this argument. Line ownership of operating performance is established doctrine; every mature operating management system, from the IOGP OMS to the leadership element of API RP 1173, already assigns the system to the operating function. Likewise, the analyst literature has already converged on documented workflows, structured data, and governance as the preconditions for AI value. The contribution of this paper is the connection between the two: it identifies the locus of management-system ownership, operating function versus compliance function, as the variable that determines whether documentation becomes agentic capability, and it reframes decades of management-system investment as the substrate asset of the agentic transition. To the author’s knowledge, that connection has not been made explicit in either body of work.

2. Foundations of Operationally-Owned Management Systems

The shared logic of every management system standard reviewed in this section, and of every analogous framework across other disciplined industries, is a recursive feedback loop. The operator defines a standard or procedure for a work process with a defined output and a defined measurement; sets a tolerance around that output; executes the standard; measures the result; treats deviations as either failures (which trigger root cause analysis) or as findings from a proactive audit; and feeds those findings back into the standard. This is the Plan-Do-Check-Act (PDCA) cycle that underlies every major framework discussed below. PDCA is necessary but not sufficient. What makes PDCA produce performance is *who is in the loop*. When the operations function owns each step (defines the standard, runs the measurement, performs the verification, conducts the RCA, approves the improvement), the loop closes inside the organization that has the authority and the incentive to act on it. When any of those steps is owned by a function that is not running the equipment, the loop opens. Findings sit. Corrective actions languish. The audit becomes a compliance event.

2.1 API RP 1173: Pipeline Safety Management Systems

API RP 1173 was issued in July 2015 as the first U.S. pipeline-industry-specific safety management system standard (American Petroleum Institute, 2015). It originated from the National Transportation Safety Board's investigation of the July 2010 Enbridge crude oil pipeline rupture near Marshall, Michigan, which released approximately 843,000 gallons of crude into Talmadge Creek and the Kalamazoo River (National Transportation Safety Board [NTSB], 2012). In Safety Recommendation P-12-17, the NTSB called on API to develop a pipeline-industry-specific safety management system standard, and the resulting workgroup brought together API, the pipeline trade associations (including AOPL, INGAA, AGA, APGA, GPA Midstream Association, APCA, and DCA), and PHMSA (NTSB, 2015; American Petroleum Institute, 2022). The standard is organized as a PDCA cycle around ten essential elements: Leadership and Management Commitment; Stakeholder Engagement; Risk Management; Operational Controls; Incident Investigation, Evaluation, and Lessons Learned; Safety Assurance; Management Review and Continuous Improvement; Emergency Preparedness and Response; Competence, Awareness, and Training; and Documentation and Record-keeping (Pipeline SMS Industry Team, n.d.).

The first element, Leadership and Management Commitment, is the ownership element. The standard does not call for ESH leadership commitment. It calls for *operating* leadership commitment. The 2023 Pipeline SMS Annual Report records that operators representing nearly 85% of U.S. transmission and gathering mileage have adopted or are implementing PSMS (the pipeline Safety Management System of API RP 1173, distinct from OSHA's Process Safety Management standard, or PSM, addressed in Section 2.2), the third consecutive year at that level (Pipeline SMS Industry Team, 2023); the 2024 report, the program's ninth, documents continued implementation maturity (Pipeline SMS Industry Team, 2025). What the adoption numbers do not reveal is which of those operators have implemented PSMS as an operationally-owned discipline and which have implemented it as a compliance program housed under ESH. The difference is decisive for performance. It will be decisive again for agentic adoption. In March 2025, PHMSA issued Advisory Bulletin ADB-2025-01, urging all gas distribution, transmission, and hazardous liquid operators to implement safety management systems aligned with API RP 1173 (PHMSA, 2025). An advisory bulletin is not a rule, but it responds directly to NTSB Safety Recommendation P-24-002 and the PIPES Act of 2020, and it signals the direction of regulatory expectation. Whether or when a mandate follows, the operators that have run PSMS as a real, operationally-owned system are in a different position from the operators that have run it as a binder.

2.2 OSHA PSM: 29 CFR 1910.119

OSHA's Process Safety Management standard at 29 CFR 1910.119 has governed the prevention of catastrophic releases of highly hazardous chemicals since 1992. It comprises fourteen elements (including employee participation, process safety information, process hazard analysis, operating procedures, training, contractors, pre-startup safety review, mechanical integrity, hot work, management of change, incident investigation, emergency planning and response, compliance audits, and trade secrets) and requires that operators conduct a compliance audit at least every three years (OSHA, n.d.). The structure is sound, and the conceptual ownership is correct: every element on that list is naturally executed by an operations organization. Operating procedures, mechanical integrity, management of change, pre-startup safety review, incident investigation: these are operations functions in their substance, regardless of where they sit on the organization chart.

The OSHA enforcement data is the cleanest available evidence on what happens when the ownership is wrong. The most frequently cited PSM elements are not equipment failures but documentation and verification failures. Historical analysis of mechanical integrity citations has shown that within paragraph (j) of the standard, written procedures and inspection-and-testing failures dominate the citation profile, with citations issued not because the process equipment was deficient but because the methods of verifying and documenting mechanical integrity were insufficient (Inspectioneering, 1995). More recent reviews show the same pattern persisting: two decades later, mechanical integrity remained the most frequently cited PSM element, led by inspection-and-testing and written-procedure failures (Inspectioneering, 2019). A quantitative evaluation of OSHA PSM inspections found that the citation record tracks the root causes identified in major accident investigations (Luo, 2010): the inspections are finding real gaps. The diagnosis is consistent: the standard exists; the field execution does not match the standard; the audit is supposed to find the mismatch; the audit only runs every three years; and the people most likely to find the mismatch in real time, the operators running the equipment, are not the people who own the audit. That is what compliance-owned PSM looks like in the citation data.

2.3 IOGP Report 510 / 511: Operating Management System Framework

The International Association of Oil & Gas Producers' Report 510 defines an Operating Management System framework built on four fundamentals (Leadership, Risk Management, Continuous Improvement, and Implementation) and ten elements that span commitment and accountability; policies, standards, and objectives; organization, resources, and capability; stakeholders and customers; risk assessment and control; asset design and integrity; plans and procedures; execution of activities; monitoring, reporting, and learning; and assurance, review, and improvement (IOGP, 2014a). Report 511 is the implementation supplement (IOGP, 2014b). The framework is more explicit than its U.S. peers about ownership: the word *Operating* in *Operating Management System* is doing real work. The framework's own implementation guidance situates the OMS as the way the operations organization runs the business: quality, reliability, environment, and safety together, against one set of policies, one audit program, and one continuous improvement loop. The structural compatibility of the IOGP framework with ISO 9001, 14001, 45001, and 55001 has been mapped by ISO Technical Committee 67 (ISO TC 67, 2018). After three decades of independent development, the major management system standards have converged not only on PDCA and on roughly ten elements, but on the principle that the operating function owns the system.

2.4 ISO 9001 and ISO 14001 in the Integrated Management System

ISO 9001 (Quality) and ISO 14001 (Environmental) both employ the PDCA cycle and share, with ISO 45001 (Occupational Health and Safety) and ISO 55001 (Asset Management), a common Annex SL high-level structure. The 2015 revisions deliberately aligned terminology and top-level clauses across the family to make integrated management systems feasible (ISMS.online, n.d.). In a midstream context this matters operationally because it permits a single documented management system, owned by the operations function, to satisfy the quality, environmental, safety, and asset-integrity requirements simultaneously. One management system, one cadence, one set of standards, one audit program. Quality is produced by the same documented work that produces safety, reliability, and environmental performance. These are not separate programs sitting alongside the operation; they are the operation, executed to standard.

2.5 Equivalent Frameworks Across Industries

The same structural logic appears in every industry that has codified its operating discipline. ISO/IEC 27001 and the NIST Cybersecurity Framework structure information security as a managed system of documented controls, defined measurements, monitored deviations, and audit-driven improvement. ITIL and ISO/IEC 20000 do the same for IT service management. The COSO Internal Control Integrated Framework and the Sarbanes-Oxley Section 404 control regime do the same for financial reporting. Project Management Institute, ISO 21500, and IEEE bodies of practice do the same for engineering and project delivery. The American Bar Association's matter-management and legal-operations bodies of practice do the same for legal work. In every case the standard was designed to be owned by the operating function it governs (security operations, IT service operations, finance and accounting, engineering, legal), not by a separate audit, compliance, or controls group. The structural argument of the remainder of this paper applies to all of them. The midstream-specific examples that follow demonstrate the mechanics with the most operationally-detailed evidence available, but the substance generalizes.

2.6 The Quality Airplane

The mental model that captures what these systems do, when they are owned by operations and used as designed, is the *quality airplane* (Figure 1). One wing is failure analysis: root cause analysis after an event, structured and documented. The other wing is the proactive audit:

supervisor or peer inspection of standards against actual field execution, looking for the gap between paper and practice before a failure makes the gap visible. The fuselage is the catalogued, auditable body of procedures and measurements. The two wings together lift performance; either wing alone causes the plane to roll. The structural critique of the dominant implementation model, developed in the next section, is that the proactive-audit wing has been outsourced to a calendar and a compliance function, and the airplane therefore climbs only intermittently.

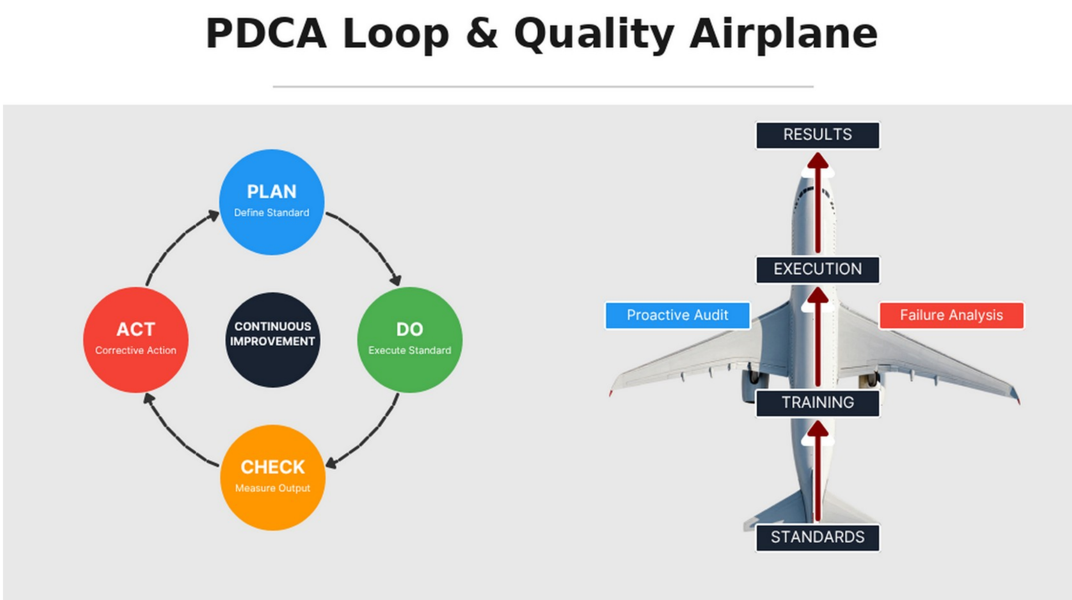

Figure 1. The PDCA cycle and the "Quality Airplane" model. The two wings, failure-triggered root cause analysis and proactive audit, lift performance only when both, together with the documented fuselage of standards between them, are owned by the operations function.

3. Challenges and Limitations of the Current Implementation Model

The structural critique of how the industry implements its management systems today is not a critique of the standards. The standards are correct. The critique has four parts, of which the first, ownership, is primary, and the other three are downstream of it.

The first limit is ownership. When the management system is owned by an ESH or compliance function rather than by the operations function, the system stops being a tool for running the business and becomes a document maintained for an audit. Operations stops being the customer of the standard and starts being the *subject* of it. Findings are issued to operations by a function that does not have operating authority. Corrective actions compete for operating attention with the work that operations leadership actually owns. The audit becomes an event rather than a discipline. Over time, the field practice and the documented practice drift apart, and the institutional knowledge of the gap lives only with the operators who have learned which corners to cut and which ones the auditor will look at. This is the failure mode the OSHA enforcement data describes. Citations are issued for documentation and verification failures, not equipment failures, because the work was done and the verification was not, because the verification was not the operating organization’s work to begin with (Inspectioneering, 1995, 2019). Ownership is a spectrum rather than a binary, and most real organizations are hybrids, with some elements owned by operations and others delegated; the failure mode described here strengthens in proportion as ownership moves away from the operating line.

The second limit is cadence. OSHA PSM compliance audits are required every three years (OSHA, n.d.). Pipeline SMS programs measure themselves on a similar multi-year horizon. Procedure-level audits inside a typical operating company occur annually at best, and many occur less often. The proactive-audit wing of the quality airplane therefore flaps at a frequency measured in years. Between those audits, the only formal correction loop is the incident itself: a release, an injury, an unplanned shutdown, or a regulatory finding. A triennial audit cannot detect drift that develops and matures inside the three-year window. A great deal of drift can develop in three years. Cadence is downstream of ownership because the cadence is set by the function that owns the audit. When operations owns the audit, the cadence is the operating cadence: daily walkdowns, shift turnover, weekly reviews, monthly performance against tolerances. When compliance owns the audit, the cadence is the regulatory cadence.

The third limit is resource intensity. A complete PSM or PSMS audit requires a multi-person team, weeks to months of fieldwork, document review, and interview cycles, and a corrective action plan that competes with operating priorities for the same staff. The gains achievable at the cadence and the resource model that a regulated, periodic audit cycle can sustain appear to be largely exhausted. The constraint is not under-investment in safety; operators sit at the frontier of what an episodic, human-intensive, paper-mediated audit cycle owned by a non-operating function can produce.

The fourth limit, observed across change management literature, is that large-scale management system transformations fail at high rates. McKinsey’s transformation practice estimates that approximately 70% of large-scale organizational transformations fail (McKinsey & Company, n.d.-b) — a widely cited figure whose empirical basis has been challenged (Hughes, 2011), though the qualitative pattern of installed-but-not-embedded programs is consistent across the literature. Kotter’s foundational 1995 *Harvard Business Review* article reached a similar observational conclusion across more than one hundred companies (Kotter, 1995). The failure mode is consistent with the ownership critique above: programs designed and installed by functions outside the operating line do not embed in daily management because the operating line was never the owner. The implication is not that operational excellence cannot be built, but that it cannot be installed from the outside on a calendar driven by a consulting engagement. It must be owned, used, and refined by the operations organization that runs the equipment. Every operator that has built an operationally-owned management system has done the hard part. Every operator that has installed a compliance-owned program has the hard part still in front of them.

These four limits together describe a real ceiling. The ceiling is not on the standards themselves. The ceiling is on the ownership model, the cadence, the resource model, and the verification gap of the dominant implementation. Lifting the ceiling requires putting the management system back in the hands of operations, raising the audit frequency by orders of magnitude, raising the verification fidelity, and doing all three without proportionally raising the human cost. That is the problem statement that agentic AI is uniquely suited to address.

4. The Agentic AI Opportunity: Closed-Loop Improvement Bolted onto the Operating Management System

Agentic AI systems, as defined in the introduction, plan, act, observe, and adapt against a defined objective and a defined set of tools (Anthropic, 2024). They differ from prior generations of operational software in two ways that matter for this paper. First, they reason. The current generation of reasoning models produces an explicit chain of thought before acting, learns to refine strategies, and recognizes its own mistakes (OpenAI, 2024). Second, they iterate. The defining behavior of an agentic system is the closed improvement loop: the agent acts, observes the result, evaluates the result against its objective, refines its approach, and consolidates what it learned for the next cycle.

That structure is, intentionally, the same structure as PDCA. The agent's *Plan* step references the operational standard. The *Do* step is the action taken against tools and data. The *Check* step compares the observed output to a defined tolerance. The *Act* step proposes a correction. Anthropic's engineering guidance describes the prerequisites bluntly: agents need structured tool definitions, structured inputs, and grounded context to act reliably (Anthropic, 2025a, 2025b). Anthropic's introduction of Agent Skills frames the prerequisite layer in language that will be immediately familiar to anyone who has built a management system: skills are "organized folders of instructions, scripts, and resources that agents can discover and load dynamically" (Anthropic, 2025c). That is, an operating procedure. A standard with a defined output, a defined measurement, and a defined tolerance is the technical interface through which an agent operates against the operation. A management system that is already used by operations to run the operation that way is, in the most practical sense, AI-ready.

The point worth emphasizing is that the substrate the agent needs is not the *documentation*. It is the *used* documentation. An agent loaded with a procedure that is filed but not followed will detect the discrepancy between the procedure and the operating reality and surface it as an audit finding. That is useful, but it is useful only if there is an operations organization that owns the procedure and is willing to act on the finding. In a compliance-owned model, the finding is generated and then it competes with the rest of the compliance backlog for operating attention. In an operationally-owned model, the finding is generated and the operations organization that owns the procedure is the one receiving it. The agentic loop closes inside the operating organization. The improvement happens. The next cycle starts.

4.1 Two Levels of Agentic Capability

Agentic AI applied to operations comes in two distinct levels, and the distinction matters for the kind of investment an operator should be making.

The first level is *task-level*. A task-level agent executes a defined task against a defined standard. An agent drafts a Management of Change package, runs an inspection routine, fills out a control test, reconciles a ledger entry, or drafts an incident report. This is what most enterprise agentic AI adoption looks like today, and it is what most people picture when they hear "AI agent." Task-level agents are useful, increasingly available, and increasingly easy to deploy. They are also fundamentally a productivity story: substitute agent labor for human labor on individual units of work.

The second level is *system-level*. A system-level agentic capability does not execute tasks inside the management system. It manages the management system. It observes the standards, the measurements, the audit outputs, the corrective actions, and the operating outcomes; looks for gaps between the documented system and the live operation; surfaces refinements to the standards themselves; and consolidates what it learns across cycles into a refined model of the operation. A system-level agent is what turns an episodic, periodic, human-paced management system into a continuously self-improving one.

The distinction is the difference between an agent that *does the work* and an agentic system that *manages the management system*. The first is useful; the second is the transformation. This paper is about the second. And the second is only possible when there is a coherent, owned management system for the agent to manage. A task-level agent can be deployed against an ad hoc instruction. A system-level agent has nothing to manage if no one in the organization has built and embedded a management system in the first place. That is the structural prerequisite operationally-owned management systems satisfy and that compliance-owned or absent systems do not.

4.2 How the System-Level Loop Runs

The system-level agent operates entirely on top of the existing operational infrastructure rather than replacing any part of it. Standards live where standards already live: in the document management system, the PSM document base, the procedure library. Data lives where data already lives: in the distributed control system, the historian, the SCADA system, the lab information management system, the work order system, the procedural compliance records. The agent reads from those systems. It does not own them, and it does not duplicate them.

That read posture rests on a deliberate security architecture. In critical-infrastructure operations, agent access to operational technology data is read-only and indirect: the agent consumes replicated data from a historian mirror or data platform in the demilitarized zone between the corporate and control networks, consistent with the network segmentation operators already maintain under their OT cybersecurity programs and, for pipeline operators, the TSA Security Directives issued after the 2021 Colonial Pipeline incident. The agent has no presence on the control network, no path to a controller, and no write access to any system that touches the process. Its write access is confined to the business-side systems where findings and drafts belong: the CMMS, the document management system, the MOC tool. The security review for this kind of agentic deployment is therefore an extension of the OT/IT segmentation review the operator already performs, not a new class of exposure.

The agent's continuous work is to observe the operating state against the standard. When the output drifts toward or outside the tolerance, the agent performs root-cause-style reasoning across the data it can see, surfaces the most likely contributors with citations to the underlying data, and drafts a corrective action into the system that already manages corrective actions: a work order in the CMMS, a procedure flagged for review in the document management system, a Management of Change package drafted in the MOC tool, a follow-up inspection scheduled on the inspection plan. The agent contributes to the existing workflow; it does not create a parallel one.

For low-consequence, reversible items, the agent can be authorized to execute the action directly within a tightly scoped action space. For consequential or irreversible items (anything touching primary process containment, safety instrumented systems, or anything that requires formal Management of Change), the agent prepares the recommendation, documents the evidence, and routes it to the responsible human through the existing review channel. The human reviews, approves or modifies, and authorizes. The authority structure of the management system is unchanged; the agent participates in it as a continuously-available analyst.

The same logic governs the regulatory record. An agent-drafted Management of Change package, incident report, or regulatory submission is a draft until the accountable human reviews and signs it, and the record of compliance is the human-authorized record, with the agent's evidence trail retained beneath it. Nothing in the current regulatory framework contemplates an agent as a responsible party, and nothing in this operating model requires it to. The signature authority, and the accountability it carries, stays exactly where the management system already places it.

After each cycle, the agent consolidates what it learned across the cycle: refining its internal model of the unit, sharpening the heuristics it watches for, and improving the pattern base it draws on for the next cycle. This consolidation step gives the agent its memory across cycles. It is the source of the system-level learning that distinguishes agentic operations from rule-based automation. The work itself is not new (proactive audit, deviation detection, RCA, corrective action, management review), but it now runs continuously against live data instead of episodically against a sampled paper trail.

This consolidation step is why Figure 2 draws the agentic loop as a flywheel rather than a circle. A loop repeats; a flywheel accumulates momentum. Each cycle leaves the system better than it found it: a sharper model of the unit, a tighter standard, cleaner data, a more precise watch list for the next pass. The output of one cycle becomes the input of the next, so the improvement compounds rather than merely recurs, and the longer the flywheel runs, the more operation-specific knowledge it embodies and the harder the resulting advantage is to replicate. That compounding is the mechanism behind the widening competitive gap described in Section 5: the leaders are not just ahead, their flywheel is already turning.

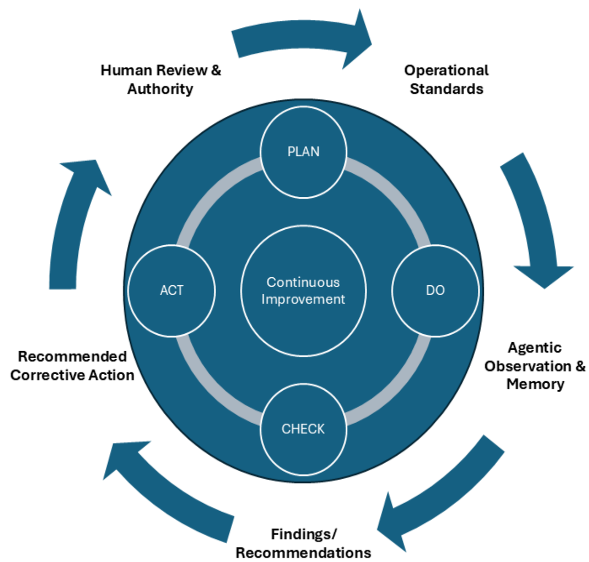

Figure 2. The agentic flywheel accelerating the operational management system. The outer, system-level loop (Operational Standards → Agentic Observation & Memory → Findings & Recommendations → Recommended Corrective Action → Human Review & Authority) sits on top of the existing PDCA infrastructure, reading from where standards and data already live and routing through existing review channels, to accelerate the continuous improvement the management system already calls for.

4.3 Why This Is the Right Moment

Two recent industry surveys say the moment is here. The 2025 MIT Sloan Management Review / BCG study of agentic enterprise adoption found that more than a third of surveyed companies were already deploying agentic AI systems and another 44% were planning to (MIT Sloan Management Review & BCG, 2025). McKinsey's State of AI 2025 survey reports that 88% of organizations now use AI in at least one business function, 23% are scaling an agentic AI system somewhere in the enterprise, and 39% report that AI has begun to register a measurable impact on EBIT (McKinsey & Company, 2025). The technology is past the proof-of-concept phase. The competitive differentiation is now in the operating model.

4.4 What the Data Substrate Has to Look Like

The analyst research is consistent about where agentic implementations founder: not on model capability, but on the data, governance, and workflow substrate underneath the agents. Gartner's 2025 AI-Ready Data research finds that 63% of organizations either lack the right data management practices for AI or are uncertain whether they do, and Gartner forecasts that organizations will abandon 60% of AI projects through 2026 for want of AI-ready data (Gartner, 2025). BCG's *Closing the AI Impact Gap* (2025a) reaches a similar conclusion, finding that only about a quarter of executives report that their companies have created significant value from their AI initiatives.

Perhaps the most telling statistic for an operations audience comes from a 2025 survey of roughly 2,200 knowledge workers: only 16% report that their workflows are extremely well documented (Lucid Software, 2025). The operators who have built and embedded an operationally-owned management system have already done what only 16% of the broader economy has done. They have documented their work, defined their outputs, specified their tolerances, and catalogued their procedures. They have a baseline against which an agent can audit and reason, and, crucially, an operating function that is in the habit of acting on the findings. They are, in the language of the analyst literature, AI-ready by accident of decades of investment in operational discipline.

One caveat keeps the claim honest. Management-system maturity and data-substrate readiness are related but not identical. An operating function can own its standards, measurements, and audit discipline and still hold much of that knowledge in unstructured documents and siloed historians that an agent cannot yet consume. Documentation maturity supplies the standards, the defined outputs, and the operating habit of acting on findings; it does not, by itself, supply the structured, integrated, machine-readable data and the AI-specific governance that an agent also requires. What the operationally-owned management system closes is the larger and harder gap: the organizational and cultural one. The data-engineering and integration work remains real, which is precisely why even an AI-ready operator should treat it as a deliberate step rather than an accident of good documentation. A related objection holds that rapidly improving models will simply read the unstructured documents and siloed historians directly, making the substrate investment unnecessary. Models are indeed improving at retrieval, but retrieval cannot supply authority: an agent can find a procedure in any format, while only a management system can tell it which version is current, what tolerance applies, and who owns the deviation. The binding constraint is not format; it is correctness and ownership.

4.5 Evidence from Comparable Operations

The early agentic case studies in oil and gas come from upstream and from the service-company side, but their mechanics translate directly to midstream operations. Equinor disclosed in January 2026 that AI created $130 million in value in 2025, bringing the cumulative value realized from AI in its industrial processes to more than $330 million since 2020 (Equinor, 2026). Three contributors illustrate the disclosure. Predictive maintenance on more than 700 rotating machines instrumented with 24,000 sensors accounts for $120 million of that cumulative total since 2020. AI-driven well and field-development planning on Johan Sverdrup Phase 3 found a solution that no one had considered, saving the partnership $12 million. AI seismic interpretation delivered a tenfold increase in interpretation capacity, with two million square kilometers interpreted in 2025. Equinor's executive vice president for technology, digital, and innovation describes the operating posture: employees use AI tools (copilots, chatbots, and agentic AI) to solve tasks and work in new ways. What the disclosure describes, in other words, is AI in the hands of the workforce, with the value reported where the operating assets are: the rotating machinery, the field-development plans, the seismic volumes.

A *Journal of Petroleum Technology* case study from April 2026, authored by an ONGC engineer, documents an agentic AI framework for large-scale offshore well modeling. Across two cases covering more than 600 offshore wells and 370 tubing-sensitivity scenarios, the framework saved more than 1,000 combined engineering hours, with all 370 simulations of one case running in under one hour of wall-clock time. The authors are precise about the operating philosophy: the central objective is not to replace engineering judgment, but to remove repetitive manual steps so that engineers can focus on interpretation and decision-making rather than model construction (Society of Petroleum Engineers, 2026). That is the same operating philosophy this paper recommends for midstream: operations authority, agent acceleration. SLB and Shell announced a strategic agreement in December 2025 to develop agentic AI for subsurface operations, well construction, and production, built on SLB's Lumi data and AI platform (SLB, 2025). Baker Hughes' Leucipa system has reported uplifts in mature-field production driven by data integration and continuous optimization (Baker Hughes, 2025).

A reader applying the two-level distinction of Section 4.1 will notice that every one of these deployments is task-level: predictive maintenance, model construction, production optimization. That is the honest state of the published evidence. The system-level capability this paper describes, the agent that manages the management system, has not yet been demonstrated in a published industrial deployment. The claim of this paper is not that it exists today. The claim is that the task-level deployments are converging on it, that its technical prerequisites are the same substrate the task-level agents already require, and that the operators who own their management systems are the ones positioned to run the first demonstrations.

The industry gap is not technological. The technology is in production at adjacent operators. The gap is the ownership and the substrate. Which operators have a management system that the operations function uses to run the operation, populated with current standards, current measurements, current data, and a current cadence of review? Those operators can put an agent to work against the management system tomorrow. The operators that do not have the operationally-owned substrate cannot.

The same diagnosis generalizes to the industries surveyed in Section 2.5: an IT operation under ITIL, a finance function under COSO and Sarbanes-Oxley, or a security operation under ISO/IEC 27001 that owns its documented controls, measurable outputs, and live telemetry can do the same against its operating model. The constraint is everywhere the same: live, operationally-owned documentation, measurable outputs, structured data, and an operating function that will act on what the agent surfaces.

5. Competitive Advantage and the AI-Readiness Gap

The competitive consequence of the agentic transition is asymmetric, and the asymmetry runs along the line of who owns the management system. Operators with mature, operationally-owned management systems can begin deploying agents against their existing documentation and data substrate immediately, into an operations function that knows how to act on the findings. Operators with compliance-owned programs have the *appearance* of the substrate without the operating practice that makes the substrate valuable. Operators without either are looking at a multi-year, multi-million-dollar program of work simply to qualify to evaluate the technology.

The strongest evidence for this asymmetry comes from the analyst and consulting research published over the last twelve months. McKinsey's State of AI 2025 finds that only about 6% of organizations qualify as AI high performers, defined as those reporting an EBIT impact of 5% or more attributable to AI, and that these high performers are at least three times more likely than peers to report scaling agents. The McKinsey synthesis identifies the differentiator: high performers run a full-stack management playbook across strategy, talent, operating model, technology, data, and adoption, with fundamental workflow redesign among the factors correlating most strongly with EBIT impact (McKinsey & Company, 2025). BCG's Build for the Future 2025 study reports that AI leaders expect twice the revenue increase and 40% greater cost reductions than laggards in the areas where they apply AI, and BCG's measured comparisons show the leaders delivering 1.7 times the revenue growth of laggards (BCG, 2025b). Deloitte's 2025 AI ROI analysis finds that only about one in five surveyed organizations qualify as AI ROI leaders, and those that do treat AI as an enterprise transformation rather than as a tool implementation (Deloitte Global, 2025). All of these findings rest on self-reported survey data and are best read as directional rather than precise; what gives them weight is their convergence across independent samples and methodologies.

The reading of this evidence is direct and applies across industries. The organizations positioned to be AI high performers are those that already operate by the discipline that the analyst literature is now identifying as the precondition for AI value. They write down what they do, measure what they produce, audit their own work against the standard, and investigate deviations, all within a culture of management review and continuous improvement. The operating function owns the system, uses the system, and refines the system. None of that is new. All of it is now also a competitive moat that produces operational performance and every outcome that depends on it as a consequence.

The cost side of the asymmetry is equally direct. The dominant cost of agentic implementation is not AI model work; it is data engineering, governance, and workflow integration against the operating substrate, the same substrate deficiencies the analyst research identifies as the leading reason AI projects are abandoned (Gartner, 2025). For the operator with an operationally-owned management system, much of that work is *already substantially reduced*. For the operator without one, that substrate work is now an additional, unfunded program of work that must be built before the agentic tooling can be put to use. It is built, when built externally, through large consulting engagements in the millions of dollars (industry-standard ranges for end-to-end operational excellence transformations run wide, with the upper end common for global integrated programs; cited consulting practices generally publish the *value captured* rather than the *cost incurred*; see McKinsey & Company, n.d.-a, and BCG, n.d.). It is built, when built internally, on a multi-year timeline that competes for the same operating attention that the existing priorities already absorb. Both routes are expensive, both are slower than the agentic adoption curve, and both face the transformation-failure record discussed in Section 3 (Kotter, 1995; McKinsey & Company, n.d.-b). The failure mode, in the literature, is the same failure mode the OSHA citation pattern shows: the program exists on paper but is not embedded in the daily management of the operation. Embedding takes time, repetition, and operational ownership. It is not a deliverable that an external program can hand over.

A temporal trap is built into this asymmetry. The agentic adoption curve is steep and accelerating; the adoption and scaling figures cited in Section 4.3 describe where the market already is, not where it is headed. The work of building an operationally-owned management system from scratch (documenting standards, defining measurements, establishing audit cadences, embedding the discipline in the operating culture) is a multi-year program even under ideal conditions. The work of adopting agentic AI on top of a mature management system is a matter of quarters. An organization that begins both programs simultaneously is starting the slower one against competitors who are already several years ahead on the foundation and are using that lead to compound through the faster one. The catch-up problem is not one of effort or budget alone but of *time the market is not giving back*. Each quarter that passes without the foundation is a quarter that the leaders are pulling further ahead and that the cost of closing the gap rises. This is what makes the disadvantage critical rather than merely competitive: the gap is opening during the catch-up itself (Figure 3).

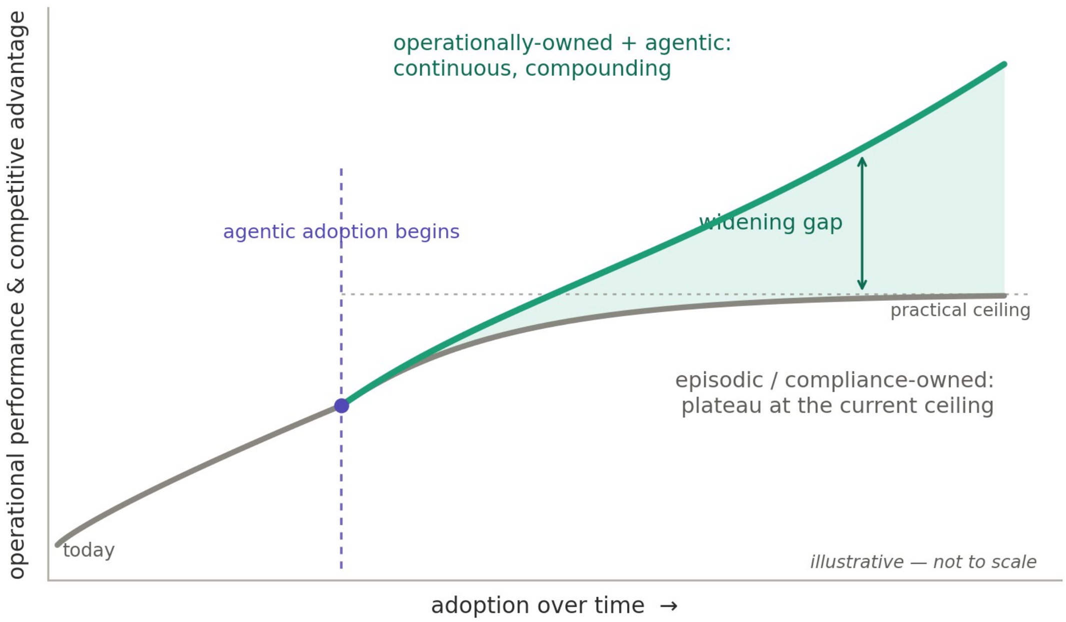

Figure 3. Diverging improvement trajectories (illustrative). Operationally-owned operations that adopt agentic AI move from episodic, periodic improvement to continuous, compounding improvement, while organizations without that foundation plateau at the ceiling of the current model. The gap is small at first and widens over time.

One caveat tempers the urgency without dissolving it. The same agentic tooling the leaders are deploying also lowers the cost of building the foundation in the first place. Task-level agents, the productivity tier described in Section 4.1, can accelerate the unglamorous work of establishing a management system: drafting procedures, normalizing them to a house format the agent learns once and then reapplies, reconciling existing documentation against actual practice, and instrumenting measurements. A procedure that once took an engineer the better part of a day to write and route may take an hour to draft and review. This genuinely lowers the barrier. What it does not change is the nature of the work that has to be done. The standards still have to be made correct by the people who own them; the measurements still have to be defined and trusted; the data still has to be integrated; and, hardest of all, the operating function still has to take ownership and embed the discipline in its daily management. Agents can compress the drafting. They cannot compress the *embedding*. The catch-up is therefore less daunting than a pre-AI consulting timeline would suggest, but it is not negligible, and the time it still takes is time the leaders are using to compound their head start.

A second version of the leapfrog argument holds that the substrate can be bought rather than built, through vendor platforms that arrive with embedded workflows and data models. The platforms are real and increasingly capable; the SLB-Shell agreement is evidence of how heavily the vendor ecosystem is investing. But a purchased workflow is a documented standard that no one in the operating function yet owns, and the change-management failure record is precisely the record of installed-but-not-embedded systems. What can be bought is the drafting.

What cannot be bought is the ownership, and ownership is the variable this paper's entire argument turns on.

The exposure is particularly acute for small and mid-sized organizations. Most small businesses, professional service firms, regional operators, and emerging companies do not have formal management systems and have not historically seen the need for them. The economics of the consulting model that built those systems for the large industrial operators put the entry cost out of reach. The agentic transition is changing the calculus. Those same organizations now face the prospect of AI-native competitors entering their markets with documented, agent-driven operations from the start, competitors whose unit economics will reflect the productivity multiplier the agentic substrate produces. The competitive consequence for a small or mid-sized organization without a management system is not "slower growth." It is "displaced by a competitor whose cost structure cannot be matched without the foundation." Formal operational excellence has historically been the domain of large regulated industrial operators with compliance budgets to fund it. The agentic transition is generalizing the requirement to every disciplined business (accounting firm, law practice, engineering consultancy, IT service provider, security operations team, manufacturing shop) that intends to remain competitive over the next several years. For these organizations the entry cost has also dropped: agentic tooling itself can help draft procedures, document standards, instrument measurements, and embed the audit cadence, *if* the operating function is the one driving the build. The window in which an entrant organization can build the foundation before the AI-native competition arrives, however, is shorter than the historical management-system implementation timeline assumes.

The most important corollary of this argument is that the relationship between operational performance and the outcomes the organization is accountable for becomes more visible at the agentic horizon. The implication is direct: an operating organization that is running its management system continuously, with agentic support, is producing reliability, throughput, cost discipline, environmental performance, safety, data integrity, security posture, and client outcomes from the same disciplined work. None of those are parallel programs competing for attention with operational priorities. They are the output of an operating organization running its own management system to standard, at the cadence the work actually moves at. When the cadence rises from triennial to continuous, every output rises with it. The competitive advantage the AI-ready organization captures is a competitive advantage across all of these dimensions at once, and the competitive disadvantage the AI-blocked organization suffers is a disadvantage across all of them at once. Figure 4 summarizes the resulting positions as a two-by-two of management-system ownership and agentic deployment intent.

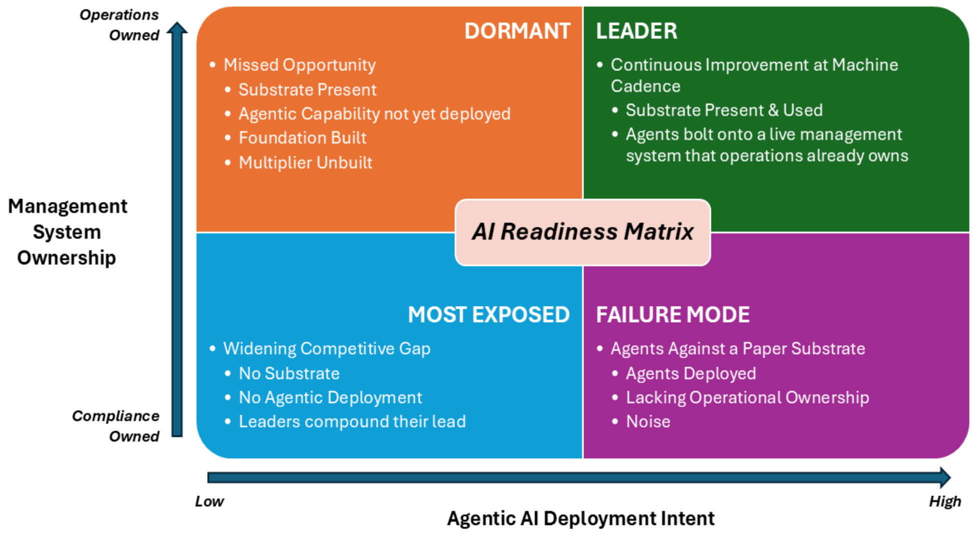

Figure 4. The AI-readiness matrix. Management-system ownership (compliance-owned to operationally-owned) on the vertical axis and agentic deployment intent (low to high) on the horizontal. Only the upper-right quadrant (operationally-owned, high intent) achieves continuous, system-level improvement at machine cadence; the lower-right, where agents are deployed against a paper substrate the operating organization will not act on, is the failure mode the analyst literature documents.

6. Path Forward: Roadmap, Ownership, Education, and Governance

The path forward is practical rather than conceptual: three phases, and three requirements that cut across all of them.

The first phase is foundation. Before any agentic work, the operating leader has to answer the ownership question honestly. *Who in this organization owns the management system?* The question has observable answers. Five markers separate an operationally-owned system from a compliance-owned one: who chairs the management review; who writes and approves procedure revisions; who closes corrective actions, and how long they stay open; whether the audit cadence exceeds the regulatory minimum; and where findings route first. An organization that answers “operations” to all five owns its management system; an organization that answers “ESH,” “quality,” or “compliance” to most of them owns a binder. If the answer is the operations function, the foundation is in place and the work is in maintaining it: keeping standards current, keeping the audit cadence at the operating cadence rather than the regulatory one, keeping the deviation-investigation discipline live, keeping the management review meaningful, and ensuring the operational standards are stored in a form an agent harness can consume reliably, typically a machine-readable mirror (markdown or structured text) maintained alongside the authoritative document. That last step is a small modification to the existing system, but it is the difference between an agent that can parse and reason against the standards and one that cannot. If the answer is an ESH or compliance function, the foundation work is to move ownership to where the work is done. That is a leadership move, not a documentation move. It requires the operations function to take responsibility for the standards that govern its work, to take responsibility for measuring its outputs, and to take responsibility for closing its own gaps. Audit, compliance, ESH, controls, and quality functions become subject-matter partners and assurance providers; they stop being the system's owners. External forcing functions help drive this internally: PHMSA's March 2025 advisory bulletin urges all U.S. pipeline operators to implement PSMS-aligned safety management, responding to an NTSB recommendation and to the PIPES Act of 2020, and signals where regulatory expectation is heading even before any mandate (PHMSA, 2025); parallel forcing functions are emerging in other industries, from cybersecurity-disclosure rules to the EU AI Act. The same investment that makes an organization regulator-ready makes it AI-ready, but only if the operating function is the owner.

The second phase is pilot. The right pilot is data-rich and lower-consequence. Midstream examples include continuous procedural compliance monitoring against documented standards; automated drafting of management of change packages for human review; automated synthesis of incident investigation reports from operational data; surveillance of process variable drift against tolerances over time horizons longer than a control loop but shorter than an audit cycle; and synthesis of regulatory reporting drafts from operational and inspection records. The analogues in other industries are direct: continuous internal-controls monitoring in finance, change-advisory and incident-postmortem drafting in IT, and continuous control-evidence collection in security. The objective is not to demonstrate the technology but to show that the existing operationally-owned management system, instrumented agentically, produces an order-of-magnitude improvement in audit frequency and verification fidelity at a fraction of the human cost. Anthropic's published engineering guidance is explicit: start with simple prompts, optimize them with comprehensive evaluation, and add multi-step agentic systems only when simpler solutions fall short (Anthropic, 2024). That is the same discipline that governs a well-run management of change program. The pilot's purpose is to demonstrate the loop closing inside the operating organization, not to showcase the most agentic system possible.

The third phase is scaling. Scaling is where the management system itself begins to evolve. The agent's consolidation step, what it learned across many cycles, surfaces patterns in the operating data, gaps in the standard, and opportunities to refine the procedure. The management review element of the standard takes those recommendations as input. The standard is updated by the responsible operations leader. The operating procedure becomes a living document under continuous evidence-informed review rather than a static document under triennial audit. This is the system-level improvement loop the thesis claims, and it is the point at which the agentic capability stops being a productivity tool and becomes a capability of the operationally-owned management system itself.

Across all three phases, three requirements hold. The first is ownership, discussed above. The second is governance. The agent acts within a defined action space. The boundary of that action space is set by the consequence and the reversibility of the action. Safety-critical, irreversible, or material-cost actions require human authority, specifically the operations authority who would have made the decision in the absence of the agent. The agent prepares the evidence; the human decides. The OpenAI o1 system card documents that even the most capable reasoning models hallucinate on a non-trivial fraction of responses, with a meaningful share of those hallucinations being errors the model could have caught from its own chain of thought (OpenAI, 2024). The governance posture, in any high-consequence context (process safety, financial reporting, cybersecurity, legal advice, patient safety), is that a non-zero hallucination rate is incompatible with autonomous action on consequential systems. Gartner’s research on AI-ready data is blunt about where AI projects fail: the binding constraint is the strength of an organization’s data and governance foundations rather than the capability of the model, and weak foundations are why a large share of AI initiatives are abandoned before they deliver value (Gartner, 2025). The operator that treats governance as a constraint rather than as an enabler will not get past the pilot phase.

The third crosscutting requirement is education, and it is broader than it first appears. The workforce skills required to operate, supervise, and improve an agentic system are not the same as the skills required to write a procedure, run a control test, manage a service incident, or close the books, but they overlap substantially with the skills required to think clearly about a management system, and they belong to the same people who do the work. The decisive mistake an organization can make here is to treat agentic design as an information-technology function. IT understands a knowledge worker’s workflow no better than it understands that worker’s expense report; the person who runs the work is the person who knows what an agent should watch for, what ‘good’ looks like, and when the agent is wrong. When agentic capability is walled off inside IT, or restricted by default out of an understandable fear of the technology, the organization rebuilds the very barrier this paper warns against, this time between the operating function and its own improvement engine.

The right mental model is the spreadsheet. Knowledge workers build their own spreadsheets every day; they do not file an IT ticket for every formula. The same division of labor should govern agentic systems: IT stands up the platform, the security model, and safe sandboxes; legal and risk define the policy envelope of what may and may not be automated; and the operating function builds and runs its agents within it. Require IT approval for every agent and the predictable result is the agentic equivalent of handwritten ledgers and notes kept in a text file: workers routing around the system while the value never materializes. Within a generation of tools, fluency in building and supervising agents will be a baseline expectation of a knowledge worker, exactly as fluency in office software is today; there is no meaningful difference. An organization that builds that fluency broadly, across the operating function rather than in a central team, turns its workforce into the engine of its own improvement; one that does not has left a real and compounding barrier to competitiveness in place. Internal capability held widely is more sustainable than external dependency, and far cheaper over time than the consulting model.

The workforce dimension of continuous assurance requires the same attention. An agent that continuously compares field execution to the documented standard can be perceived by the workforce as surveillance, and the perception is legitimate wherever the deployment model invites it. Ownership is again the difference. In a compliance-owned deployment, continuous monitoring is imposed on the operating workforce from outside, and it meets the resistance that externally imposed audit regimes have always met. In an operationally-owned deployment, the crew supervising the agent is the crew whose work it observes; findings route to the people who own the procedure rather than to a scorecard; and the agent is tuned to answer the operator's own question: is the standard wrong, or is the practice wrong? Field experience produces both answers. A management system that treats a gap between paper and practice as information rather than as an infraction is the cultural precondition for placing an agent on the audit wing at all.

7. Conclusion

Agentic AI is the largest near-term performance opportunity most operating organizations will see this decade, and the gap between those positioned to act on it and those who are not is already widening. The case made in this paper is not that agentic AI is a substitute for the operational discipline the disciplined industries have built over the last several decades. The case is the opposite. Operationally-owned management systems are the substrate. The agentic system is the multiplier. The substrate without the multiplier is what most disciplined industries already have, and most have reached the ceiling of what their current cadence and ownership model can produce. The multiplier without the substrate is what the analyst literature is documenting as the dominant failure mode of enterprise AI adoption: agents deployed against inadequate data, undefined outputs, and undocumented workflows, producing noise and governance failure rather than improvement.

The limits of the argument deserve the same directness as the argument itself. The evidence for the system-level claim is structural and analogical rather than experimental: the published deployments are task-level, the survey findings are self-reported, and no controlled comparison of agentic adoption across ownership models yet exists. The ownership construct itself needs empirical work; the five markers offered in Section 6 are a working hypothesis, not a validated instrument. And the pace of the underlying technology guarantees that any specific capability claim made here will age faster than the structural argument. These are the places where the thesis can be tested, and testing it is the research agenda this paper is intended to provoke.

The deepest reframe in this argument concerns the relationship between operational performance and the outcomes the organization is accountable for. In a process industry that means safety; in finance, data integrity and audit-readiness; in IT, uptime and incident-free service; in security, a defensible posture; in legal, client trust; in any industry, cost discipline and throughput. None of these are parallel programs. They are what an operating organization produces when it runs its work to standard. The same management system, owned by the same operating function, produces them all from one disciplined set of practices. When agentic AI multiplies that management system, it multiplies all of those outputs at once. Organizations that have invested in operationally-owned management systems have, without intending to, built the most important operating asset of the next decade. Organizations that have left their management systems with audit, compliance, ESH, controls, or quality functions, or have never built them at all, are looking at years of expensive foundational work simply to qualify for the operating model that competitors are already deploying, and they are watching the performance, reliability, and outcome gap widen at the same time. That gap is not theoretical, and it is not symmetric in time. For large organizations, it shows up as widening performance and cost differentials against AI-leader peers. For small and mid-sized organizations, it shows up as displacement by AI-native competitors whose unit economics cannot be matched without the foundation. In both cases the time available to close the gap is being consumed by the agentic adoption curve itself. The disadvantage is therefore critical, not merely competitive: organizations without a management system are not only behind, they are losing ground faster than they can build the substrate that would let them compete.

The resulting agenda is direct and industry-agnostic. It begins with the ownership question, honestly answered: who owns the management system, and is it the function that does the work? Where the answer is a separate audit, compliance, ESH, controls, or quality function, the first work is to move ownership to where the work is done. The work that follows is to close the documentation and data gaps in the highest-consequence operations first (the procedures, the controls, the measurements, the decision rights), to identify two pilot opportunities in data-rich, lower-consequence areas, and to build internal capability in agentic operations alongside the existing operating capabilities of the function. The industries that built the operating playbooks for process safety, for quality, for information security, for IT service management, and for financial control did it by giving the operating function the standards, the measurements, and the authority. The organizations that write the operating playbook for agentic operations will do it the same way. The work begins by treating the management system as the operating asset that it is.

About the Author

Bryan G. McMurray has more than 25 years in leadership roles spanning three domains that rarely converge in one person: technology and data, operational management, and environment, safety, and health (ESH). His career began at Texas A&M University in College Station, where he earned a B.S. in Manufacturing Engineering Technology and started in programming and database management at the university's Public Policy Research Institute (PPRI), an applied policy research center.

On the technology side, his programming and database work has continued for more than two decades and includes two companies he founded: XTEKI, LLC (formerly JB Web Technologies), which he owns and through which he continues to provide small-business consulting, and Desktop Disposal, LLC, a national electronics-recycling firm he later sold. In heavy manufacturing, steel, and oil and gas, he has led operations as a maintenance supervisor and plant superintendent and advanced through a progression of ESH leadership roles: ESH Coordinator, Environmental Manager, ESH Manager, Safety & Health Director, and ES&H Director.

Much of that leadership has centered on building and implementing management systems: ISO 14001 at Koppers, the Gerdau Business System and ISO 14001 at Gerdau, and the safety, health, and environmental programs at Targa Resources that he has helped design, build, and lead the implementation of over the past eleven years. He currently serves there as a Director of ES&H. That combination of operational leadership, ESH discipline, and hands-on software and data experience is the vantage point from which this paper is written. Additional detail is available on his LinkedIn profile.

References

American Petroleum Institute. (2015). *ANSI/API Recommended Practice 1173: Pipeline Safety Management Systems* (1st ed.). https://www.api.org/~/media/files/publications/whats%20new/1173_e1%20pa.pdf

American Petroleum Institute. (2022, July 7). *New Pipeline SMS report shows ongoing industry commitment to pipeline safety*. https://www.api.org/news-policy-and-issues/news/2022/07/07/pipeline-sms-report-shows-outgoing-industry-commitment-to-safety

Anthropic. (2024). *Building effective agents*. https://www.anthropic.com/research/building-effective-agents

Anthropic. (2025a). *Effective context engineering for AI agents*. https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents

Anthropic. (2025b). *Writing effective tools for AI agents*. https://www.anthropic.com/engineering/writing-tools-for-agents

Anthropic. (2025c). *Equipping agents for the real world with Agent Skills*. https://www.anthropic.com/engineering/equipping-agents-for-the-real-world-with-agent-skills

Baker Hughes. (2025). *Using unified data to boost oil production — no drilling required* (Sarir Oil Operations case study). https://www.bakerhughes.com/company/energy-forward/using-unified-data-boost-oil-production-no-drilling-required

Boston Consulting Group. (n.d.). *End-to-end operational excellence*. https://www.bcg.com/capabilities/operations/end-to-end-operational-excellence

Boston Consulting Group. (2025a). *Closing the AI impact gap*. https://www.bcg.com/publications/2025/closing-the-ai-impact-gap

Boston Consulting Group. (2025b, September 30). *AI leaders outpace laggards with double the revenue growth and 40% more cost savings* (press release). https://www.bcg.com/press/30september2025-ai-leaders-outpace-laggards-revenue-growth-cost-savings

Congressional Research Service. (2025, May 7). *DOT's federal pipeline safety program: Background and issues for Congress* (CRS Report R44201). https://crsreports.congress.gov/product/pdf/R/R44201

Deloitte Global. (2025). *AI ROI: The paradox of rising investment and elusive returns*. https://www.deloitte.com/global/en/issues/generative-ai/ai-roi-the-paradox-of-rising-investment-and-elusive-returns.html

Equinor. (2026, January 7). *Use of artificial intelligence saved Equinor USD 130 million in 2025*. https://www.equinor.com/news/20260107-artificial-intelligence-saved-equinor-usd-130-million

FracTracker Alliance. (2025, April 7). *Pipeline incidents are a daily occurrence*. https://www.fractracker.org/2025/04/pipeline-incidents-are-a-daily-occurrence/

Gartner. (2025, February 26). *Lack of AI-ready data puts AI projects at risk* (press release). https://www.gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk

Hughes, M. (2011). Do 70 per cent of all organizational change initiatives really fail? *Journal of Change Management, 11*(4), 451–464.

Inspectioneering. (1995, May). *OSHA PSM citations for mechanical integrity*. https://inspectioneering.com/journal/1995-05-01/60/osha-psm-citations-for-mechani

Inspectioneering. (2019, January 3). *A quick review of recent OSHA PSM citations*. https://inspectioneering.com/blog/2019-01-03/8177/a-quick-review-of-recent-osha-psm-citations

International Association of Oil & Gas Producers. (2014a, June). *Report 510: Operating Management System Framework for controlling risk and delivering high performance in the oil and gas industry*. https://www.ourenergypolicy.org/wp-content/uploads/2014/06/OMF.pdf

International Association of Oil & Gas Producers. (2014b, June). *Report 511: OMS in practice — A supplement to Report 510*. https://www.iogp.org/bookstore/wp-content/uploads/sites/2/woocommerce_uploads/2017/01/511.pdf

ISMS.online. (n.d.). *ISO 14001 for the oil and gas industry*. https://www.isms.online/sectors/iso-14001-for-the-oil-and-gas-industry/

ISO TC 67. (2018). *Management system standards for the oil and gas industry*. https://committee.iso.org/files/live/sites/tc67/files/Workgroup2/Management%20system%20standards%20for%20oil%20and%20gas%20industry%20v2018-12.pdf

Kotter, J. P. (1995, March–April). Leading change: Why transformation efforts fail. *Harvard Business Review*. https://www.researchgate.net/publication/246663143_Leading_Change_Why_Transformation_Efforts_Fail

Lucid Software. (2025). *AI readiness report* [Survey of approximately 2,200 knowledge workers]. https://lucid.co/blog/ai-readiness-survey-2025

Luo, H. (2010). The effectiveness of U.S. OSHA process safety management inspection — A preliminary quantitative evaluation. *Journal of Loss Prevention in the Process Industries, 23*(3), 455–461. https://www.sciencedirect.com/science/article/abs/pii/S0950423010000239

McKinsey & Company. (n.d.-a). *How we help clients — Operations* (oil and gas practice). https://www.mckinsey.com/industries/oil-and-gas/how-we-help-clients/operations

McKinsey & Company. (n.d.-b). *Why do most transformations fail? A conversation with Harry Robinson*. https://www.mckinsey.com/capabilities/transformation/our-insights/why-do-most-transformations-fail-a-conversation-with-harry-robinson

McKinsey & Company. (2025, November). *The state of AI in 2025: Agents, innovation, and transformation*. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

MIT Sloan Management Review & Boston Consulting Group. (2025, November). *The emerging agentic enterprise: How leaders must navigate a new age of AI*. https://sloanreview.mit.edu/projects/the-emerging-agentic-enterprise-how-leaders-must-navigate-a-new-age-of-ai/

National Transportation Safety Board. (2012). *Pipeline Accident Report PAR-12-01: Enbridge Incorporated hazardous liquid pipeline rupture and release, Marshall, Michigan, July 25, 2010*. https://www.ntsb.gov/investigations/AccidentReports/Reports/PAR1201.pdf

National Transportation Safety Board. (2015). *Safety Alert SA-095: Pipeline Safety Management Systems — Vital for the safe operation of pipelines*. https://www.ntsb.gov/Advocacy/safety-alerts/documents/SA-095.pdf

Occupational Safety and Health Administration. (n.d.). *29 CFR 1910.119: Process safety management of highly hazardous chemicals*. https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XVII/part-1910/subpart-H/section-1910.119

OpenAI. (2024, September 12). *OpenAI o1 system card*. https://cdn.openai.com/o1-system-card.pdf

Pipeline and Hazardous Materials Safety Administration. (n.d.). *Pipeline incident 20-year trends*. https://www.phmsa.dot.gov/data-and-statistics/pipeline/pipeline-incident-20-year-trends

Pipeline and Hazardous Materials Safety Administration. (2025, March 25). *Pipeline safety: Pipeline safety management system* (Advisory Bulletin ADB-2025-01; 90 FR 13658). *Federal Register*. https://www.federalregister.gov/documents/2025/03/25/2025-04960/pipeline-safety-pipeline-safety-management-system

Pipeline SMS Industry Team. (n.d.). *Fact sheet — RP 1173 Pipeline Safety Management Systems*. https://pipelinesms.org/fact-sheet-rp-1173-pipeline-safety-management-systems/

Pipeline SMS Industry Team. (2023). *2023 Pipeline SMS Annual Report*. https://pipelinesms.org/2023-pipeline-sms-annual-report/

Pipeline SMS Industry Team. (2025, July 22). *2024 Pipeline SMS Annual Report*. https://pipelinesms.org/2024-pipeline-sms-annual-report/

SLB. (2025, December 11). *SLB enters collaboration agreement to accelerate new digital solutions* (Shell collaboration press release). https://www.slb.com/news-and-insights/newsroom/press-release/2025/pr-2025-1211-slb-shell-collab

Society of Petroleum Engineers. (2026, April). Case study: An agentic AI framework for large-scale well modeling of offshore field developments. *Journal of Petroleum Technology*. https://jpt.spe.org/case-study-an-agentic-ai-framework-for-large-scale-well-modeling-of-offshore-field-developments