Before the Incident

A Framework for Industrial AI Safety

Why Operators Must Lead the Standards Conversation
— Before Someone Else Does

Council for Industrial AI Safety (CIAS)

February 2026

Executive Summary

> *"You already have AI making decisions in your facility. Do you know which ones?"*

Artificial intelligence is no longer a pilot project in industrial operations—it is production infrastructure. According to McKinsey, global AI adoption reached 78% in 2025, with 23% of organizations already scaling agentic AI systems. The National Association of Manufacturers reports 51% of U.S. manufacturers are actively using AI. DNV finds 47% of energy organizations plan to deploy AI in operations within the year. The industrial AI market is projected to grow from $4.35 billion in 2024 to $280 billion by 2035.

This is not a technology problem. It is a governance vacuum.

No existing industrial standards body—not ISA, not API, not ISAGCA, not the OPC Foundation—has published a safety framework for AI in operational technology environments. IEC 62443 does not address AI. ISA-84/IEC 61511 cannot validate non-deterministic systems. The EU AI Act classifies industrial AI as high-risk but offers no implementation standards. Executive Order 14110, the most comprehensive U.S. AI governance effort, was rescinded after only fifteen months.

Operators are deploying AI into environments where failures have kinetic consequences—thermal runaway, toxic release, overpressure events—without shared standards, incident data, or a collective voice in the regulatory conversation.

This paper presents an operator-led framework for industrial AI safety: a governance structure, risk assessment methodology, lifecycle management approach, and implementation roadmap designed by and for the people who run these facilities. It draws on established practices from process safety—Management of Change, Layers of Protection Analysis, Safety Integrity Levels—and extends them to address the unique characteristics of AI systems: non-determinism, model drift, adversarial vulnerability, and emergent behavior.

The Council for Industrial AI Safety (CIAS) was founded to fill this gap. We are not waiting for an incident to define the standards. We are building them now.

Three things you can do today:

1. **Share this paper** with your operations, safety, and engineering leadership.
2. **Use the self-assessment framework** (Section 5) to map where AI is making decisions in your facility.
3. **Join CIAS as a founding member** to help build the standards your industry needs.

1 The Convergence: AI Meets Industrial Operations

1.1 A Scenario

It is 2:15 AM on a Tuesday. A predictive maintenance model at a Gulf Coast refinery flags an anomalous vibration signature on a critical compressor and automatically adjusts operating parameters to compensate. The adjustment is subtle—a 3% reduction in feed rate, a slight modification to a temperature setpoint. No alarm fires. No operator is notified. The DCS logs the change as a routine optimization.

Twelve hours later, a process engineer reviewing daily reports notices the feed rate deviation. She traces it to the AI system's intervention and realizes the model was responding to sensor drift, not an actual mechanical fault. The parameter adjustment, while small, shifted the process closer to a constraint boundary that the AI system was never trained to recognize. The safety instrumented system would have caught an actual exceedance—but the margin had been quietly eroded.

No one was hurt. No release occurred. But the facility's safety envelope had been narrowed by a system that no one in operations had reviewed through Management of Change, that had no defined operating limits, and whose decision logic could not be interrogated by the board operator who owned that unit.

This scenario is a composite. It is also, increasingly, unremarkable.

1.2 Deployment Velocity

The pace of AI adoption in industrial operations has outstripped every prior technology wave. Consider the data:

- **McKinsey (2024–2025):** Global AI adoption rose from 72% to 78% in a single year. Of those organizations, 23% are now scaling agentic AI—systems that take autonomous action, not merely provide recommendations.
- **National Association of Manufacturers:** 51% of U.S. manufacturers report active AI use in operations.
- **DNV:** 47% of energy organizations plan to deploy AI in operational roles within the coming year.
- **Market projections:** The industrial AI market is forecast to grow from $4.35 billion in 2024 to $280 billion by 2035—a compound annual growth rate of approximately 46%.

These numbers describe a transformation already underway. AI is optimizing crude unit yields, predicting heat exchanger fouling, scheduling turnaround maintenance, managing power grid loads, and—in an increasing number of facilities—writing directly to control systems.

1.3 The Use Cases Are Real

To be clear: this paper is not an argument against AI in industrial operations. The technology delivers genuine value:

- **Predictive maintenance** reduces unplanned downtime by identifying equipment degradation weeks or months before failure.
- **Process optimization** improves yield, energy efficiency, and throughput in ways that traditional model-predictive control cannot match.
- **Anomaly detection** identifies process deviations that human operators might miss in the flood of thousands of tag values.
- **Demand forecasting** in power and utilities enables more efficient dispatch and grid management.
- **Quality prediction** reduces off-spec production and rework.

The question is not whether to deploy AI. The question is how to deploy it safely, in environments where the consequence of failure is not a crashed application or a wrong recommendation—it is a pressure vessel rupture, a chlorine release, or an arc flash fatality.

1.4 The Governance Gap

Every prior generation of industrial technology was adopted within an existing safety and governance infrastructure. When Distributed Control Systems replaced pneumatic loops, there were established practices for alarm management, operator training, and change control. When Safety Instrumented Systems were formalized, IEC 61511 defined the entire lifecycle from hazard analysis through decommissioning.

AI has no equivalent. It arrived faster than the governance infrastructure could adapt, and it possesses characteristics—non-determinism, opacity, continuous learning, emergent behavior—that existing frameworks were never designed to address.

As former DHS Secretary Mayorkas observed, *"AI offers a once-in-a-generation opportunity to improve the strength and resilience of U.S. critical infrastructure."* But as the Government Accountability Office warned in December 2024, *"[Deploying AI] may make critical infrastructure systems…more vulnerable."*

Both statements are true. The difference between them is governance.

2 The Risk Landscape: What Makes Industrial AI Different

Industrial AI is not enterprise AI. The differences are not matters of degree; they are differences in kind. A failed recommendation engine loses revenue. A failed industrial AI system can kill people.

2.1 Kinetic Consequences

Industrial processes involve energy, mass, and chemistry at scale. The consequence categories that operators manage daily—thermal runaway, overpressure, toxic release, explosion, environmental contamination—exist because the physics do not forgive errors. An AI system that subtly shifts a reactor temperature set-point, misclassifies a leak detection signal, or fails to flag a developing equipment failure operates in this consequence space whether its designers intended it to or not.

The process safety community learned this lesson through decades of incidents: Bhopal, Texas City, Deep-water Horizon. Each taught us that safety is not a feature of individual components—it is a property of the entire system, including its management, its culture, and its decision-making processes. AI is now part of that system.

2.2 The Purdue Model and AI's Ambiguous Position

Industrial control architectures are organized by the Purdue Enterprise Reference Architecture into hierarchical levels, from Level 0 (physical process) through Level 4 (enterprise). Security architectures built on IEC 62443 define zones and conduits that control data flow between these levels.

AI systems violate this architecture by design. A predictive maintenance model might ingest Level 0 sensor data, process it at Level 3 or in a cloud environment (Level 4+), and write recommendations—or actions—back to Level 1 or Level 2 controllers. The data flow crosses zone boundaries that were specifically designed to be restrictive. The model itself may reside outside the control system's security perimeter entirely.

No current standard addresses how to define zones and conduits for AI data flows, what security controls should govern AI model deployment into control system environments, or how to validate that an AI system respects the architectural boundaries that protect safety-critical functions.

2.3 Safety Instrumented System Interaction

Safety Instrumented Systems (SIS) represent the last automated line of defense against catastrophic events. They are designed, validated, and maintained under ISA-84/IEC 61511 with rigorous lifecycle requirements, defined Safety Integrity Levels (SIL), and proof-testing regimes.

AI introduces several challenges to SIS integrity:

- **Process variable manipulation:** An AI optimization system that adjusts process parameters may move the process closer to SIS trip points, reducing the effective safety margin without triggering any alarm.
- **Sensor dependency:** AI systems that share sensor inputs with SIS can create common-cause failure modes not accounted for in SIL verification calculations.
- **Demand rate changes:** AI-driven operational changes may alter the demand rate on safety functions, invalidating assumptions made during LOPA.
- **Bypass pressure:** When AI-recommended setpoint changes conflict with SIS trip points, there is organizational pressure to adjust the safety system rather than constrain the AI.

IEC 61511 was written for deterministic systems with characterizable failure modes. Non-deterministic AI systems cannot be validated under existing SIL verification methods. This is not a theoretical gap—it is a fundamental incompatibility that no standards body has yet addressed.

2.4 Adversarial Threats in OT

The OT security community is well acquainted with targeted attacks. Stuxnet (2010) demonstrated that nation-state actors could manipulate industrial control systems with surgical precision. Triton/TRISIS (2017) showed that attackers could target Safety Instrumented Systems directly. Colonial Pipeline (2021) illustrated how IT-side compromises could cascade into operational shutdowns.

AI introduces new attack surfaces specific to OT environments:

- **Training data poisoning:** Corrupting the historical process data used to train models, causing them to learn incorrect operating envelopes.
- **Adversarial inputs:** Manipulating sensor data in ways imperceptible to human operators but sufficient to cause AI misclassification.
- **Model extraction:** Stealing proprietary process models that encode sensitive operational knowledge.
- **Supply chain compromise:** Injecting vulnerabilities through third-party AI components or pre-trained models.

MITRE ATLAS documents adversarial tactics against AI systems, but contains no OT- or ICS-specific scenarios. The CISA/NSA joint guide published in December 2025 addresses the secure integration of AI into OT environments, but focuses on cybersecurity rather than process safety. The gap between security and safety in the AI context remains unaddressed.

2.5 Model Drift and Non-Determinism

Traditional control systems are deterministic: the same inputs produce the same outputs, every time. This property is foundational to hazard analysis, LOPA, and SIL verification.

AI systems are non-deterministic by nature. A machine learning model's behavior is a function of its training data, architecture, and learned weights—not a set of explicitly programmed rules. Model drift—the gradual

degradation of model performance as the relationship between inputs and outputs shifts over time—is not a bug; it is an inherent characteristic.

In a refinery, feedstock composition changes seasonally. Equipment degrades. Process conditions shift after turnarounds. A model trained on historical data becomes progressively less accurate, and the rate and direction of that degradation are themselves unpredictable. In enterprise applications, model drift causes revenue loss. In industrial operations, it can cause the model to misidentify a genuine safety-critical condition as normal.

Existing practice has no answer for this. There is no equivalent of proof testing for model accuracy. There is no defined frequency for model revalidation. There is no standard for what “acceptable degradation” means in a safety-critical context.

3 The Regulatory Reality

3.1 The Standards Gap

The following table summarizes the current state of standards and guidance relevant to industrial AI safety. The pattern is clear: existing frameworks either predate AI entirely or address it only at a sector-agnostic level that provides no actionable guidance for OT environments.

Standard/Framework	Scope	Industrial AI Gap
NIST AI 100-1	AI risk management (sector-agnostic)	No OT-specific guidance; no SIS interaction; no real-time control considerations
IEC 62443	Industrial cybersecurity	Does not address AI systems, AI-specific threats, or AI data flows
IEC 61511 / ISA-84	Safety instrumented systems	Cannot validate non-deterministic systems; no AI lifecycle provisions
ISO/IEC 42001	AI management systems	Sector-agnostic; no industrial process safety integration
EU AI Act	AI regulation (risk-based)	Classifies industrial AI as high-risk; no implementation standards for OT
CISA/NSA Joint Guide	AI in OT security	Security-focused; does not address process safety
DHS Safety Guidelines	AI in critical infrastructure	High-level principles; no operational implementation detail
MITRE ATLAS	Adversarial AI threats	No OT/ICS-specific attack scenarios

Table 1: Regulatory and Standards Gap Analysis for Industrial AI Safety

Not a single row in this table provides an industrial operator with a complete, implementable framework for safely deploying AI in an OT environment. This is the gap.

3.2 The Federal Landscape

The U.S. federal approach to AI governance has been characterized by ambition followed by retreat:

Executive Order 14110 (October 2023) represented the most comprehensive federal AI governance initiative, directing agencies across government to develop AI safety standards, risk assessments, and reporting requirements. It was rescinded on January 20, 2025—after only approximately fifteen months—leaving the directives it initiated in various states of incomplete implementation.

NIST AI 100-1 (January 2023), the AI Risk Management Framework, provides a useful taxonomy of AI risks

and a governance structure, but is explicitly sector-agnostic. It contains no guidance specific to operational technology, process safety, or safety-critical control systems.

DHS published safety and security guidelines for AI in critical infrastructure (April 2024) and a roles and responsibilities framework (November 2024). Both documents operate at a principles level that is valuable for policy but insufficient for operational implementation.

CISA and NSA released a joint guide on the secure integration of AI into OT environments (December 2025). This document is the closest existing guidance to what operators need, but its scope is limited to cybersecurity. Process safety—the discipline of preventing catastrophic releases and equipment failures—is not addressed.

NIST announced the launch of Centers for AI in Manufacturing and Critical Infrastructure in December 2025, signaling recognition of the gap but not yet producing implementable standards.

GAO published a report (GAO-25-107435, December 2024) on AI risks to critical infrastructure, explicitly warning that deploying AI "may make critical infrastructure systems…more vulnerable."

3.3 The International Picture

The **EU AI Act** (Regulation (EU) 2024/1689) classifies AI systems used in critical infrastructure management as high-risk, triggering requirements for conformity assessment, risk management, and human oversight. However, it delegates implementation standards to European standardization bodies, which have not yet produced sector-specific guidance for industrial operations.

The **International AI Safety Report** (2025) provides a comprehensive overview of AI risks across domains but does not address industrial process safety in operational detail.

The result: operators deploying AI in facilities subject to OSHA PSM, EPA RMP, COMAH, or the Seveso Directive have no regulatory-aligned framework for managing AI-specific risks. They are left to interpret sector-agnostic guidance through the lens of their own experience, without the benefit of shared standards, incident data, or industry consensus.

4 An Operator-Led Framework for Industrial AI Safety

This section presents a practical framework for governing AI systems in industrial operations. It is not a standard—CIAS does not yet have the consensus process to issue standards. It is a starting point: a set of recommended practices built on established process safety principles and adapted for the unique characteristics of AI.

4.1 Governance Structure

AI governance in industrial environments must integrate with existing safety management systems, not replace them. We recommend a three-tiered governance structure:

1. **Site AI Safety Committee:** A cross-functional team including operations, process safety, instrument and controls engineering, IT/OT security, and management. This committee reviews all AI deployments through the Management of Change process, defines AI operating envelopes, and authorizes AI interactions with control and safety systems.
2. **AI Model Steward:** A designated role (not necessarily a new hire—often an experienced instrument engineer or process engineer) responsible for the lifecycle management of each deployed AI model, including performance monitoring, drift detection, and revalidation scheduling.
3. **Corporate AI Safety Function:** For multi-site operators, a corporate function that maintains a registry of deployed AI systems, aggregates performance and incident data across sites, and ensures consistency of governance practices.

4.2 Risk Assessment: The Industrial AI Risk Matrix

We propose classifying AI deployments along two dimensions: the **level of AI autonomy** and the **criticality of the process** the AI interacts with.

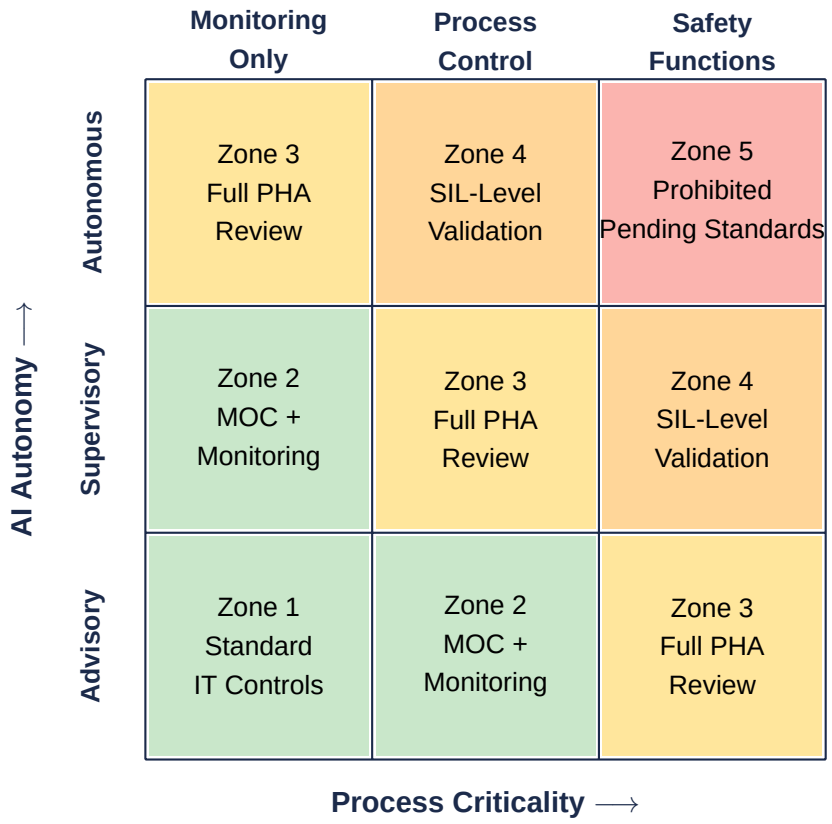

Figure 1: Industrial AI Risk Matrix — Governance requirements increase with both autonomy level and process criticality. Zone 5 (autonomous AI in safety functions) should be prohibited until validated standards exist.

Zone definitions:

Zone 1 — Standard IT Controls

AI provides information only, in non-critical applications. Standard cybersecurity and data quality controls apply. Example: energy consumption dashboards.

Zone 2 — MOC + Performance Monitoring

AI provides recommendations or monitors critical processes. Requires Management of Change review and ongoing performance monitoring. Example: predictive maintenance alerts on rotating equipment.

Zone 3 — Full PHA Review

AI has supervisory influence on critical processes or autonomous action in non-safety contexts. Requires Process Hazard Analysis including AI-specific failure modes. Example: AI-driven process optimization that writes to DCS setpoints.

Zone 4 — SIL-Level Validation

AI interacts with safety-related functions in a supervisory capacity. Requires validation rigor equivalent to SIS design, including failure mode analysis, demand rate assessment, and defined proof-test intervals for model performance. Example: AI-assisted alarm rationalization that modifies safety alarm setpoints.

Zone 5 — Prohibited Pending Standards

Autonomous AI control of safety-critical functions. No validated methodology exists for ensuring the reliability of non-deterministic systems at SIL-rated levels. CIAS recommends this zone remain prohibited until industry consensus standards are developed.

4.3 AI Lifecycle Management

AI systems are not static. Unlike a control valve or a PLC program, an AI model's behavior changes over time as it encounters new data, as process conditions evolve, and—for continuously learning systems—as the model itself updates. Lifecycle management must account for this.

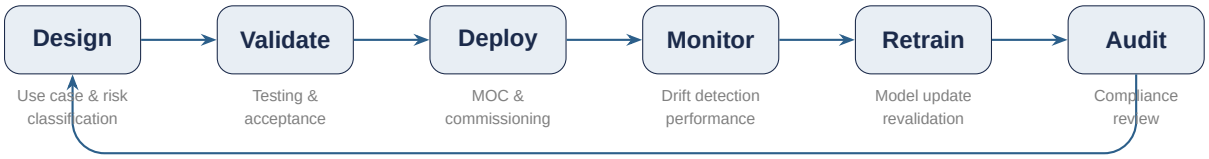

Figure 2: AI Governance Lifecycle — Continuous cycle mirroring the SIS lifecycle in IEC 61511 but adapted for non-deterministic systems.

Key lifecycle requirements by zone:

	Zone 1–2	Zone 3	Zone 4	Zone 5
MOC	Standard	AI-specific checklist	Full PHA integration	N/A
Validation	Functional testing	+ Process safety review	+ SIL-equivalent	Prohibited
Monitoring	Quarterly KPIs	Continuous drift detection	Real-time with alerting	N/A
Revalidation	Annual	Semi-annual or on trigger	Per proof-test schedule	N/A
Human Oversight	Exception review	Approval before action	Continuous supervision	N/A
Documentation	Model card	+ Hazard analysis	+ SIL verification record	N/A

Table 2: Lifecycle Requirements by Risk Zone

4.4 Zones and Conduits for AI Data Flows

IEC 62443 defines security zones and conduits to control data flows between network segments. AI data flows—training data ingestion, model inference, action outputs, telemetry—must be governed with the same

rigor.

We recommend extending the zone and conduit model to address:

- **Training data conduits:** What process data leaves the OT environment for model training? What controls ensure data integrity? What prevents training data poisoning?
- **Inference conduits:** Where does the model execute? If in the cloud or at Level 4, what latency, availability, and integrity controls govern the conduit to the control system?
- **Action conduits:** When AI outputs write to control system setpoints, what validation, rate-limiting, and bounding controls exist on the conduit? What happens when the conduit fails?
- **Model deployment conduits:** How are updated models delivered to the production environment? What change control governs this conduit?

Each conduit should have a defined owner, security classification, monitoring requirement, and failure mode. The AI system's conduit map should be reviewed as part of the IEC 62443 zone and conduit assessment and updated whenever the AI system architecture changes.

4.5 Human-in-the-Loop Requirements by SIL Level

The degree of human oversight required for AI systems should correlate with the safety criticality of the function the AI influences. We propose the following minimum requirements:

SIL Context	Minimum Human Oversight	AI Role Boundary
Non-SIL rated	Periodic review of AI outputs	Advisory or supervisory
SIL 1	Operator approval before AI action	Advisory only; operator executes
SIL 2	Continuous operator supervision; independent verification	Advisory only; dual confirmation
SIL 3–4	AI prohibited from direct interaction	No AI in SIF loop; advisory to engineering only

Table 3: Human-in-the-Loop Requirements by Safety Integrity Level

These requirements are conservative by design. As validated methodologies for AI reliability assessment emerge, the boundaries may be relaxed through industry consensus—but until those methodologies exist, caution in safety-critical applications is not conservatism; it is competence.

4.6 Management of Change for AI

Every industrial operator has an MOC process. AI deployments must go through it. But standard MOC checklists were not designed for AI systems. We recommend augmenting the MOC process with AI-specific elements:

AI-Specific MOC Checklist (Minimum Elements)

- What process variables does the AI system read? Write? What is the source and quality of each input?
- What is the AI system's operating envelope? Under what conditions is it valid? What happens outside those conditions?
- Has the AI system been tested against the process's known upset conditions, abnormal operations, and emergency scenarios?
- Does the AI system interact with or share inputs with any Safety Instrumented Function? If so, has the SIF's SIL verification been updated to account for the AI?
- What is the model revalidation frequency? What triggers unscheduled revalidation?
- Who is the designated Model Steward? What is their escalation path?
- Can the AI system be bypassed or overridden by the board operator without additional authorization?
- Has the AI system's decision logic been documented in a form that operations personnel can interpret?
- What is the rollback procedure if the AI system must be removed from service?

5 Implementation Roadmap

Implementing a comprehensive AI safety framework does not happen overnight. The following phased approach allows organizations to build capability progressively while managing risk from day one.

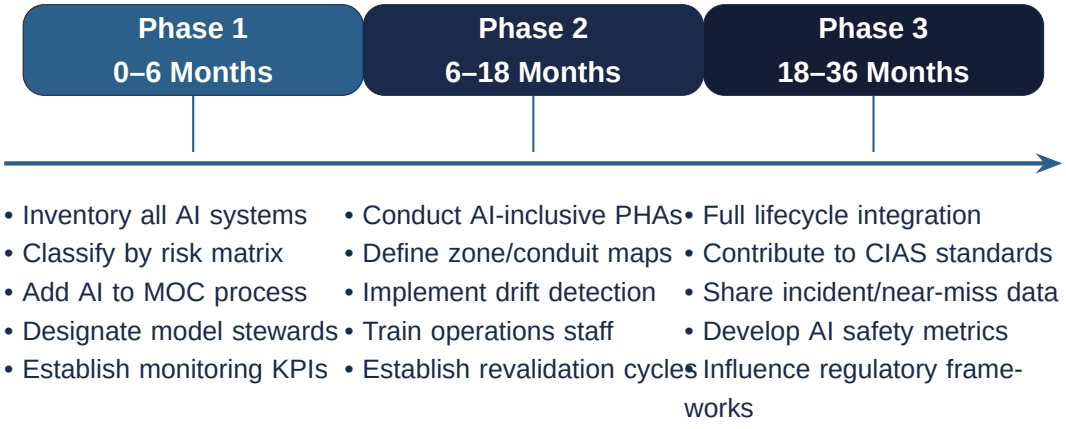

Figure 3: Implementation Roadmap — Three phases from immediate actions to industry leadership.

5.1 Phase 1: Foundation (0–6 Months)

Objective: Establish visibility and basic governance.

The first and most important step is deceptively simple: *find out what AI is already running in your facility.* In many organizations, AI deployments have arrived through vendor upgrades, IT initiatives, or engineering pilot projects without formal process safety review.

1. **Conduct an AI inventory.** Identify every system using machine learning, neural networks, or AI-driven decision-making that interacts with process data, control systems, or safety systems. Include vendor-embedded analytics that may not be explicitly labeled as "AI."
2. **Classify each system** using the Industrial AI Risk Matrix (Figure 1). Identify any systems operating in Zones 3–5 that have not undergone appropriate review.
3. **Integrate AI into MOC.** Update your MOC procedure to include AI-specific elements (see Section 5.6). Ensure that future AI deployments—including vendor updates that modify AI functionality—trigger MOC review.
4. **Designate Model Stewards.** Assign ownership for each deployed AI model. Define monitoring responsibilities, escalation paths, and revalidation triggers.
5. **Establish baseline KPIs.** Define measurable performance indicators for each AI system: accuracy, drift metrics, false positive/negative rates, availability, and any safety-relevant metrics.

5.2 Phase 2: Integration (6–18 Months)

Objective: Embed AI safety into existing process safety infrastructure.

1. **Conduct AI-inclusive PHAs.** For all Zone 3+ AI systems, perform or update Process Hazard Analyses to include AI-specific failure modes: model drift, adversarial inputs, training data quality, common-cause failures with SIS, and loss-of-AI scenarios.
2. **Map AI zones and conduits.** Document all AI data flows in the context of your IEC 62443 zone and conduit model. Identify gaps in security controls for training data, inference, and action conduits.
3. **Implement continuous monitoring.** Deploy drift detection, input validation, and output bounding for all Zone 3+ AI systems. Define automated alerts when model performance degrades beyond acceptable thresholds.
4. **Train operations personnel.** Board operators, shift supervisors, and field operators need to understand what AI systems are active in their units, what those systems do, how to override them, and what abnormal AI behavior looks like.
5. **Establish revalidation cycles.** Define proof-test-equivalent intervals for model performance revalidation, based on the risk zone and observed drift rate.

5.3 Phase 3: Leadership (18–36 Months)

Objective: Contribute to industry standards and collective learning.

1. **Achieve full lifecycle integration.** AI governance is embedded in all relevant management systems: MOC, PHA, training, incident investigation, audit, and mechanical integrity programs.
2. **Participate in CIAS standards development.** Contribute operational experience, anonymized incident and near-miss data, and subject matter expertise to the development of industry consensus standards.
3. **Share data.** The absence of an AI incident database for industrial operations is itself a risk. Contributing to a shared, anonymized repository of AI performance data, near-misses, and lessons learned accelerates the entire industry's learning curve.
4. **Develop AI safety metrics.** Move beyond model accuracy to operationally meaningful metrics: safety margin erosion rate, AI-influenced demand rate on SIFs, time between model revalidation events, and AI-related MOC rejection rate.
5. **Engage regulators.** Operators who have implemented structured AI safety programs are in the strongest position to influence emerging regulatory frameworks constructively—with operational credibility rather than lobbying.

6 The Case for Collective Action

6.1 Why No Single Operator Can Solve This Alone

The process safety community has a strong tradition of pre-competitive collaboration. The Center for Chemical Process Safety (CCPS), founded after Bhopal, has spent four decades developing guidelines that the entire industry uses. The American Petroleum Institute (API) publishes recommended practices that are adopted globally. These organizations exist because their founders recognized that safety is not a competitive advantage—it is a shared obligation.

AI safety in industrial operations presents the same imperative:

- **No single operator has enough data.** AI safety requires understanding failure modes across diverse processes, operating conditions, and AI architectures. No one facility generates sufficient data to characterize these risks comprehensively.
- **Standards require consensus.** A framework developed by one operator is a policy. A framework developed by the industry is a standard. Regulators, insurers, and the public respond to consensus, not individual initiatives.
- **Vendors respond to collective demand.** When operators individually request AI transparency, explainability, or safety documentation, they are feature requests. When the industry speaks collectively, they become requirements.
- **Regulatory engagement requires a unified voice.** Regulators are going to address industrial AI—the EU already has. The question is whether those regulations will be shaped by the people who run these facilities or by people who do not.

6.2 What CIAS Provides

The Council for Industrial AI Safety was founded to give operators that collective voice. CIAS is:

- **Operator-led.** Founding membership is limited to eight seats, reserved exclusively for industrial operators. This is not a vendor consortium or an academic exercise.
- **Practice-focused.** CIAS develops recommended practices, not position papers. Every deliverable is designed to be implementable by a site safety committee.
- **Pre-competitive.** Like CCPS and API, CIAS operates in the space where safety and competition do not intersect. AI safety frameworks benefit every operator and no operator's competitive position.
- **Standards-oriented.** CIAS aims to develop the recommended practices that evolve into the consensus standards that ISA, API, and other bodies adopt—the same path CCPS blazed for process safety.

6.3 The Window Is Open

Let us be honest about the current moment. No catastrophic incident caused by AI in an industrial environment has occurred—yet. The AI Incident Database tracks hundreds of AI failures across domains, but the industrial sector has not yet experienced its Bhopal, its Texas City, its Deepwater Horizon moment for AI.

This is precisely the window of opportunity. The process safety community knows what happens when standards are written after an incident: they are reactive, shaped by political pressure rather than operational wisdom, and they arrive too late for the people who were harmed.

We have the chance to do this differently. To build the governance infrastructure *before* the incident that makes it urgent. To define the standards while there is still time for thoughtful, operator-led consensus rather than emergency rulemaking.

That window will not stay open indefinitely. AI deployment is accelerating. Every month without shared standards is another month of ungoverned risk accumulation across the industrial sector.

> **Three things you can do today:**
>
> 1. **Share this paper freely.** Forward it to your VP of Operations, your PSM coordinator, your I&E manager, your site safety committee. The conversation needs to start.
> 2. **Use the self-assessment framework.** The risk matrix and MOC checklist in Section 5 are designed to be used immediately. Map where AI is operating in your facility. Classify the risk. Identify the gaps.
> 3. **Join CIAS as a founding member.** Eight seats. Reserved for operators. Help us build the standards your industry needs—before someone else builds them for you.
>
> **thecias.com**

References

[1] National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, January 2023. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

[2] Cybersecurity and Infrastructure Security Agency & National Security Agency. *Principles for the Secure Integration of Artificial Intelligence in Operational Technology*. December 2025. https://www.cisa.gov/resources-tools/resources/principles-secure-integration-artificial-intelligence-operational-technology

[3] European Parliament and Council of the European Union. *Regulation (EU) 2024/1689 — Artificial Intelligence Act*. 2024.

[4] International Electrotechnical Commission. *IEC 62443: Industrial Communication Networks — Network and System Security*. ISA/IEC.

[5] International Electrotechnical Commission. *IEC 61511: Functional Safety — Safety Instrumented Systems for the Process Industry Sector*. (Also published as ISA-84.)

[6] International Organization for Standardization. *ISO/IEC 42001:2023 — Artificial Intelligence — Management System*. 2023.

[7] U.S. Department of Homeland Security. *Safety and Security Guidelines for Critical Infrastructure Owners and Operators*. April 2024. https://www.dhs.gov/sites/default/files/2024-04/24_0426_dhs_ai-ci-safety-security-guidelines-508c.pdf

[8] U.S. Department of Homeland Security. *Roles and Responsibilities Framework for Artificial Intelligence in Critical Infrastructure*. November 2024. https://www.dhs.gov/publication/roles-and-responsibilities-framework-artificial-intelligence-critical-infrastructure

[9] U.S. Government Accountability Office. *Artificial Intelligence: DHS Needs to Improve Risk Assessment Guidance for Critical Infrastructure Sectors*. GAO-25-107435, December 2024. https://www.gao.gov/products/gao-25-107435

[10] McKinsey & Company. *The State of AI in 2024* and *The State of AI in Early 2025*. McKinsey Global Surveys.

[11] National Association of Manufacturers. *AI in Manufacturing Survey*. 2024.

[12] DNV. *Energy Industry Insights: AI Adoption Survey*. 2024.

[13] AI Incident Database. Partnership on AI. https://incidentdatabase.ai/

[14] MITRE Corporation. *ATLAS: Adversarial Threat Landscape for AI Systems*. https://atlas.mitre.org/

[15] National Institute of Standards and Technology. *NIST Launches Centers for AI in Manufacturing and*

Critical Infrastructure. December 2025. https://www.nist.gov/news-events/news/2025/12/nist-launches-centers-ai-manufacturing-and-critical-infrastructure

[16] International AI Safety Report. 2025. https://internationalaisafetyreport.org/

[17] Di Pinto, A., Dragoni, Y., Carcano, A. *TRITON: The First ICS Cyber Attack on Safety Instrument Systems*. Black Hat USA, 2018.

[18] Langner, R. *Stuxnet: Dissecting a Cyberwarfare Weapon*. IEEE Security & Privacy, 2011.

About CIAS

Council for Industrial AI Safety

The Council for Industrial AI Safety (CIAS) is an operator-led, 501(c)(6) trade association dedicated to developing frameworks and recommended practices for secure AI deployment in industrial environments. Founded in 2026, CIAS brings together operators, technologists, and regulators to ensure AI systems in industrial environments are safe, secure, and resilient.

Founding membership is limited to eight seats, reserved exclusively for industrial operators. This structure ensures that the standards developed by CIAS reflect the operational realities of the facilities they are designed to protect.

Mission: To develop operator-led frameworks for the safe, secure, and resilient deployment of artificial intelligence in industrial operations.

Focus Areas:

- Recommended practices for AI governance in OT environments
- Risk assessment methodologies for industrial AI
- AI lifecycle management standards
- Anonymized incident and near-miss data sharing
- Regulatory engagement and standards development

thecias.com