GRIDS AND GUIDES TECHNOLOGIES

INDUSTRIAL SYSTEMS & AI ADVISORY

The Data Foundation Gap

Why industrial AI stalls before the first model ships, and the architecture that connects SCADA, PLC, MES, and ERP data well enough to trust it.

Industry Whitepaper
Manufacturing & Industrial Operations

August 2026
Grids and Guides Technologies Research

CONTENTS

What's in this report

A note on how this report was built. This is a secondary-research synthesis, not a primary survey. Every statistic is drawn from a publicly published, cited study; the architecture, framework, and recommendations are Grids and Guides Technologies' own point of view, developed through advisory engagements with manufacturers integrating OT and IT data for AI. Section 11 lists every source in full.

The AI project didn't fail. The data underneath it did.

Nearly every manufacturer is now running an AI pilot. Almost none of them are running on a data foundation built to support it.

Redwood Software's 2026 global survey of manufacturing professionals put a number on the gap directly: 98% of manufacturers are exploring AI, but only a fifth report their data and systems are actually prepared for it.[1] Cloudera's 2026 Data Readiness Index, surveying 1,270 enterprise IT leaders, found the same pattern from the other direction: 96% say they've integrated AI into core business processes, yet nearly 80% admit their AI initiatives are still constrained by limited, fragmented data access.[2] Gartner has projected that, through 2026, organizations will abandon 60% of AI projects that lack an AI-ready data foundation.[3]

This report is about that foundation specifically as it exists on a factory floor: the layer connecting PLCs, SCADA systems, historians, MES, and ERP into something a model can actually consume. That layer is where most industrial AI initiatives quietly stall, not in the model, not in the use case, but in the unglamorous work of getting clean, contextualized, real-time data out of decades-old control systems and into a form an algorithm can trust.

The stakes are not abstract. Unplanned downtime now costs the world's largest manufacturers an estimated $1.4 trillion a year, roughly 11% of their combined revenue.[4] Manufacturing has been the single most-attacked sector for cyberattacks for five years running, and industrial ransomware incidents rose 64% in 2025 alone.[5] And as of this month, the EU Cyber Resilience Act's first hard deadline is live: manufacturers must now report actively exploited vulnerabilities in connected products, including industrial control components, within 24 hours.[6] Every one of these pressures traces back to the same root cause: nobody has a complete, governed, real-time picture of what is actually running on the plant floor.

What this report covers

We walk through what the published data says about industrial AI readiness (Section 2), quantify what fragmentation actually costs in downtime and security exposure (Section 3), explain the specific technical reasons OT and IT data don't reconcile (Section 4), and lay out a reference architecture built around the Unified Namespace pattern, ISA-95, and Sparkplug B that closes the gap (Section 5). We close with what this foundation unlocks in predictive maintenance and quality outcomes, a phased implementation path, and a set of concrete recommendations for where to start.

What the data actually shows

Ask a plant head if their operation is ready for AI and the answer is usually yes. Ask their data engineering team the same question and the answer is usually no. Four independent 2026 studies converge on this exact split, and it is worth looking at all four together because each measures a different part of the same gap.

80%

of enterprises say AI initiatives are still constrained by limited data access across environments

Cloudera Data Readiness Index, 2026 (n=1,270 IT leaders)

54%

cite data quality and availability as the top barrier to industrial AI adoption

IIoT World Industrial Data & AI Readiness Survey, 2026 (n=272)

20%

of manufacturers report their data and systems are fully prepared for AI, despite 98% actively exploring it

Redwood Software Manufacturing AI & Automation Outlook, 2026 (n=300)

The readiness illusion

Cloudera's researchers call the pattern an "AI readiness illusion": 85% of leaders in their survey said they have a clear data strategy, yet the same population reported access, quality, and integration problems severe enough to constrain nearly every live AI initiative.[2] When Cloudera's respondents were asked why AI projects specifically fell short of expectations, data quality was cited most often at 22%, ahead of cost overruns (16%) and poor integration into existing workflows (15%). Manufacturing and healthcare respondents were the most likely of any sector to name workflow integration as the core blocker, not the model itself.[2]

Precisely's 2026 State of Data Integrity and AI Readiness study found a nearly identical split among data and analytics leaders specifically: 87% claimed they had the infrastructure, skills, and data readiness AI requires, while 41 to 43% of that same group separately named exactly those categories as their biggest obstacle.[7] Only 19% said they had "good enough" metrics and KPIs for their current stage of AI deployment.

The industrial-specific data tells a sharper version of the same story. IIoT World's 2026 survey found that only 7% of industrial organizations had AI embedded in core processes, versus 44% who expected to reach that point within three years, a gap that has persisted, largely unchanged in shape, across five prior annual editions of the same survey.[8] Legacy system integration and data silos were named a barrier by 48% of respondents, second only to data quality itself.

Why this matters for architecture, not just strategy. None of these studies found budget, executive sponsorship, or algorithm sophistication to be the leading blocker. In every case, the constraint sat in the data layer: whether the right data exists, in the right place, in a form a model or an analyst can actually use. That is an architecture problem before it is anything else.

Time is being spent in the wrong place

The clearest symptom of a missing data foundation shows up in how AI project time actually gets allocated. Enterprise data teams consistently report that 60 to 70% of AI project time goes to data preparation, cleaning, and reconciliation, not model development.[9] On the plant floor specifically, this shows up as engineers manually exporting historian tags into spreadsheets, reconciling three different inventory numbers across MES, ERP, and a supervisor's tracking sheet, and re-deriving context (what does this tag actually measure, on which line, in which unit) that should have been captured once and reused everywhere.

This is not a hypothetical inefficiency. OxMaint's 2025 Global Industry Report on manufacturing maintenance found that while roughly 70% of plants have implemented a CMMS or EAM system, 49% of them still run parallel spreadsheets alongside it, a strong signal that the system of record isn't trusted or complete enough on its own.[10] Every hour spent reconciling that shadow spreadsheet against the official system is an hour not spent building the model the plant actually wants.

What high performers do differently

Deloitte's 2026 manufacturing survey in Germany found 84% of manufacturers report measurable value from at least one AI use case, but only 20% have scaled that use case beyond its original pilot line or plant.[11] The barriers cited most were not technical sophistication but implementation cost, expertise gaps, and organizational friction, the same categories that show up when a pilot's data pipeline was hand-built for one line and cannot be replicated without redoing the integration work from scratch.

The pattern across all five studies

Every readiness study cited in this section, spanning general enterprise IT, industrial operations specifically, and manufacturing quality functions, converges on the same finding stated in different words: organizations are further along in AI ambition than in AI infrastructure, and the gap between the two is measured in data architecture, not data science.

96%

of enterprises report AI integrated into core business processes

Cloudera, 2026

7%

of industrial organizations have AI embedded in core processes today

IIoT World, 2026

60–70%

of AI project time spent on data preparation rather than modeling

Enterprise AI-readiness research, 2026

Fragmentation has a price tag, and it is already being paid

A missing data foundation is not a neutral, wait-and-see position. It shows up directly on two lines of the P&L: unplanned downtime, and cyber risk concentrated exactly where OT meets IT.

Downtime economics

Siemens' widely cited True Cost of Downtime research puts the aggregate figure at $1.4 trillion a year in lost output across the world's 500 largest manufacturers, equivalent to roughly 11% of their combined revenue.[4] Averaged across all manufacturing sectors, Aberdeen Research puts the cost of a single hour of unplanned downtime at approximately $260,000; ABB's 2026 survey of more than 3,200 global plant maintenance leaders found two-thirds of plants experience an unplanned stoppage at least monthly, at a median cost of $125,000 per hour.[12,13] In automotive assembly specifically, that figure climbs to roughly $2.3 million per hour, having doubled since 2019.[4] Deloitte's benchmarking puts the average plant's annual unplanned downtime at roughly 800 hours, more than 15 hours every week.[14]

These figures matter here for a specific reason: the leading cause of extended, expensive downtime is not the failure itself but the time it takes to diagnose it, and diagnosis time is a direct function of how much visibility a team has into machine condition data at the moment something goes wrong. A control system nobody is streaming from behaves, from a data standpoint, exactly like a control system that doesn't exist.

$1.4T

annual downtime cost across the world's 500 largest manufacturers, ~11% of combined revenue

Siemens, True Cost of Downtime, 2024

$125K–260K

average cost of a single hour of unplanned downtime across manufacturing

ABB (2026) / Aberdeen Research

800 hrs

average unplanned downtime per plant per year, over 15 hours a week

Deloitte Advanced Manufacturing Report

The same gap is the attackers' entry point

Fortinet's 2025 State of Operational Technology and Cybersecurity Report, surveying more than 550 OT security professionals globally, found that industrial ransomware incidents rose 64% in 2025 alone, affecting an estimated 3,300 organizations worldwide. Manufacturing was the single most-attacked sector for the fifth consecutive year, accounting for 27.7% of all recorded incidents, and half of Fortinet's respondents reported at least one OT security incident over the year.[5]

The mechanism is specific and consistent across the 2026 threat data: intrusions overwhelmingly begin at exposed, internet-facing IT systems, then move laterally into OT once the two networks converge without adequate segmentation.[15] IT/OT convergence is precisely the architecture pattern this report advocates for on the data side; the same convergence, done without governance, is also the reason the attack surface has grown. This is not an argument against integration. It is an argument for integrating deliberately, with segmentation, visibility, and inventory built in from the start rather than bolted on after an incident.

A live example. In August 2025, Jaguar Land Rover suffered what the UK Cyber Monitoring Centre later called the most economically damaging cyberattack in British history. Attackers exploited a vulnerability in a third-party supplier's software, moved laterally into JLR's production systems, and deployed ransomware that halted manufacturing across three countries for five weeks.[6] The entry point was IT. The damage was entirely on the OT side, where the loss was measured in weeks of stopped production, not stolen records.

This is also where the data foundation argument and the regulatory argument converge directly. You cannot report an actively exploited vulnerability to a regulator within 24 hours, as the EU now requires (Section 9), if you do not already have a governed, current inventory of what is running on your floor. An asset and data inventory is not a separate cybersecurity project. It is the same foundational work this report describes, viewed through a different lens.

The dual-purpose case

Every plant considering a data foundation investment is, whether framed this way internally or not, making two investments at once: one in AI readiness, and one in the visibility that OT security and regulatory reporting now require by law. Architected correctly, a single unified data layer serves both. Architected as two separate initiatives run by two separate teams, it typically serves neither well, and duplicates the integration cost twice.

Why the floor and the enterprise don't speak the same language

The data problem on a plant floor is rarely a shortage of data. Most plants are drowning in it. The problem is that none of it was ever designed to be understood outside the system that produced it.

Protocol fragmentation

A typical multi-decade plant runs a mix of Modbus, PROFINET, EtherNet/IP, and proprietary PLC protocols, layered with OPC UA implementations of varying vintages, none of which share a common payload format or semantic model out of the box. Every one of these was designed to solve a real-time control problem inside its own cell, not to publish meaning to the enterprise. Getting data out of any single one of them is solvable in isolation; getting them all out into one consistent model is where most integration budgets go to die.

Point-to-point integration spaghetti

Absent a shared architecture, the default pattern is direct, custom integration between every pair of systems that needs to exchange data: SCADA to MES, MES to ERP, historian to a BI dashboard, a robot cell to a quality system. Each connection is its own project, with its own failure mode, and the number of connections grows roughly with the square of the number of systems. Adding one new sensor line or swapping one vendor's historian can mean touching a dozen brittle point-to-point links that nobody fully documented.

Semantic drift

The same physical measurement is routinely named differently in every system that touches it: a temperature tag might be `TT-204` in the PLC, `Line3_Temp_Zone2` in the historian, and "Zone 2 Temperature" in a supervisor's spreadsheet, with the mapping between them held in one engineer's memory rather than in any system. ISA-95 (IEC 62264) exists specifically to standardize this hierarchy, from Enterprise down through Site, Area, Line, and Cell, but the standard describes the model; it does not enforce it. Most plants reference ISA-95 in a slide deck and diverge from it in practice within the first integration project.

Historian silos and the IT/OT organizational split

Proprietary historians hold years of process data in formats never designed for modern data platforms to query natively; extracting it requires vendor-specific connectors and often a specialist consultant. Layered on top is an organizational split: IT and OT report through different leaders and optimize for different metrics, uptime and safety versus patch cadence and governance, with little history of sharing a roadmap. IT/OT convergence, done well, dissolves this; done as a slogan without an architecture behind it, it just moves the silo problem one layer up.[16]

The net effect. None of these four problems is individually exotic. Together they produce a data layer that looks connected on a diagram and is not, in practice, trustworthy enough to build a model on without months of manual reconciliation, exactly the 60 to 70% figure from Section 2.

The four failure modes, side by side

Failure mode	What it looks like on the floor	What fixes it
1	**Protocol fragmentation** Modbus, PROFINET, EtherNet/IP, mixed-vintage OPC UA, no shared payload format	Edge gateways that normalize every protocol into one common format before it reaches the enterprise layer
2	**Point-to-point integration** Every system wired directly to every other system it needs to talk to, N-squared complexity	A single publish/subscribe hub (Unified Namespace) every system connects to once
3	**Semantic drift** The same tag named differently in the PLC, the historian, and the spreadsheet	An ISA-95 topic hierarchy and a governed tag dictionary maintained as a deliverable, not tribal knowledge
4	**Historian silos & org split** Years of process data locked in a proprietary format; IT and OT on separate roadmaps	A data platform layer that ingests from the namespace once, plus a joint IT/OT governance model with shared KPIs

Why this is an architecture problem, not a tooling problem

It is possible to buy software that addresses any one of these four failure modes in isolation, and most plants have. The recurring mistake is treating each purchase as the whole solution rather than as one layer of a stack that has to work together. An MES upgrade does not fix semantic drift. A new historian does not fix protocol fragmentation. A dashboard does not fix the IT/OT organizational split. Section 5 lays out the reference architecture that treats these as one connected problem, because in practice, that is what they are.

From sensor to model: a layer-by-layer view

The architecture below is built around three components that have become the de facto industrial standard for this exact problem: the **Unified Namespace (UNS)** pattern as the central publish/subscribe hub, **MQTT Sparkplug B** (Eclipse Foundation, now Sparkplug 3.0) as the standardized payload format, and the **ISA-95 (IEC 62264)** hierarchy, Enterprise / Site / Area / Line / Cell, as the topic structure everything is organized under.[16,17] None of this is proprietary to any single vendor; brokers implementing it include HiveMQ, EMQX, Mosquitto, and AWS IoT Core, which is precisely the point: it is a pattern, not a product.

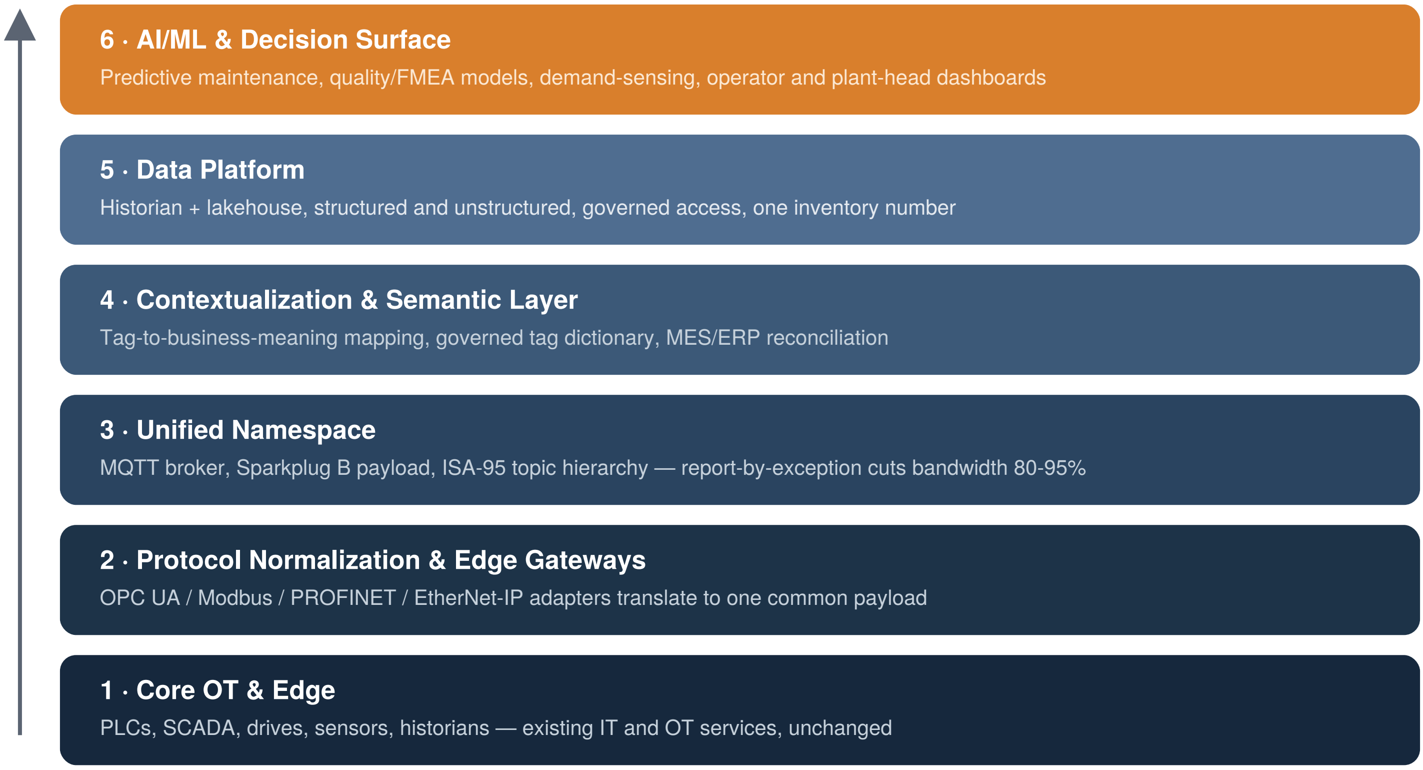

Figure 1. The Grids and Guides Technologies industrial data foundation reference architecture.

Why this pattern specifically

Report-by-exception, not constant polling

Sparkplug B's birth and death certificate model means every device publishes what data it provides the moment it comes online (an NBIRTH or DBIRTH message), and the namespace immediately knows when that device goes offline (NDEATH/DDEATH), so consumers never mistake stale data for live data. Ongoing updates use report-by-exception, publishing only changed values rather than the full dataset on every cycle, which reduces network bandwidth by a documented 80 to 95% compared to constant polling.[17] That reduction matters directly for retrofits: it is what makes it practical to stream from decades-old PLCs without overloading a control network that was never sized for continuous IT-grade traffic.

One topic hierarchy everyone agrees on

Structuring the namespace by ISA-95 levels (Enterprise/Site/Area/Line/Cell) means a predictive maintenance model, a scheduling system, and a plant-head dashboard can all subscribe to the same data at the level of granularity each one needs, without three separate point-to-point integrations being built to get it there. A new line coming online publishes into the existing hierarchy; it does not require renegotiating a dozen custom interfaces.

Context lives as a deliverable, not as memory

Layer 4 is where most reference architectures go thin, and where most of this report's emphasis sits. Publishing clean, normalized data is necessary but not sufficient; a temperature value with no unit, no line, and no business meaning attached is still not usable by an AI model or a human analyst. The semantic layer is the governed tag dictionary and business-context mapping that turns `Line3.Cell2.TT204` into "Extruder Zone 2 Temperature, Line 3, tied to Product Family A quality spec," maintained as an artifact the whole organization can query, not knowledge sitting in one engineer's head.

The data platform is where AI actually gets built

Layers 1 through 4 exist to feed Layer 5 reliably: a governed data platform combining historian time-series data with unstructured data (inspection images, maintenance logs, supplier documentation) that a data science team can build against without re-deriving context from scratch for every model. This is also the layer where the "three systems, three answers" inventory problem gets resolved into one number, because MES, ERP, and shop-floor data are reconciled once, upstream of every application that consumes them, rather than reconciled separately inside every dashboard that needs an inventory figure.

What changes for the AI/ML layer. With Layers 1 through 5 in place, a predictive maintenance model is no longer a bespoke integration project. It is a consumer of data that already exists, already has context, and already updates in real time. That is the difference between a six-month pilot and a two-week deployment for the second, third, and tenth use case, which is exactly the plant-to-plant, line-to-line scaling problem that stalls most industrial AI programs after their first pilot.

The outcomes a real data foundation makes possible

The published ROI figures for AI-driven predictive maintenance vary considerably depending on program maturity, and a responsible whitepaper should show that range rather than lead with only the most favorable number.

30–50%

reduction in unplanned downtime for mature AI-driven predictive maintenance programs

McKinsey & Company, 2025 manufacturing analytics research

8–12%

more conservative, defensible savings versus preventive maintenance for early-stage programs

U.S. Department of Energy; Deloitte (5-10% cost reduction, 10-20% uptime gain)

20–40%

extension in useful equipment life reported by mature predictive maintenance deployments

McKinsey & Company, 2025

IoT Analytics' research across predictive maintenance adopters found 95% report a positive overall return, with 27% reaching full payback within 12 months, a spread wide enough to make clear that the data foundation, not the algorithm, is usually what separates the fast payback cases from the slow ones.[18] Programs built on ad hoc, single-line integrations tend to land in the conservative range; programs built on the layered architecture in Section 5, where the second and third use case reuse the same data pipeline, tend to land closer to the McKinsey figures because the marginal cost of each new model drops sharply.

Quality and traceability

Deloitte's 2026 Germany manufacturing survey found 84% of manufacturers already report measurable value from AI in quality applications, but industry analysis is converging on hybrid approaches, pairing traditional statistical process control with AI-driven prediction, rather than AI replacing inspection outright, specifically because SPC provides the auditability and regulatory traceability that a pure AI model does not easily replicate in regulated sectors.[11] This is precisely why the semantic layer in the reference architecture matters for quality and FMEA use cases: a non-conformance traced back to a specific line, station, and supplier lot requires the same governed context and lineage that a predictive maintenance model needs, built once and reused by both.

The compounding effect

The single most consistent finding across every readiness and ROI study cited in this report is that the data foundation, not any individual use case, is what determines whether an organization's fifth AI deployment is faster and cheaper than its first, or exactly as slow. Plants that treat the namespace and semantic layer as shared infrastructure see that compounding effect. Plants that rebuild a bespoke pipeline for every new model do not, regardless of how good any single model is.

A composite scenario

Illustrative only. This is a composite scenario built from patterns commonly reported across the industry sources cited throughout this report. It does not describe a specific named client, and the figures below are illustrative estimates within the ranges documented in Sections 3 and 6, not verified outcomes from a real engagement.

A mid-size discrete manufacturer, three plants, one unreconciled inventory number

A composite of a common starting point: three plants running a mix of 15-year-old PLCs, two different SCADA vendors, and a historian none of the newer engineering staff had been trained on. MES, ERP, and a supervisor's spreadsheet each reported a different finished-goods inventory figure, and reconciling them consumed roughly half a day of a planner's week at each site. Unplanned downtime was tracked manually after the fact, with root cause typically assigned days after the event, once someone had time to review paper logs.

What an architecture like Section 5 typically changes: edge gateways normalize each plant's PLC and SCADA output into a common payload; a Unified Namespace, structured by ISA-95 site and line, becomes the single source every downstream system reads from; and a governed tag dictionary replaces the tribal knowledge previously held by two senior engineers. Inventory reconciliation moves from a manual weekly exercise to a continuously updated single number. A predictive maintenance model for the plants' highest-downtime asset class, typically the first use case deployed on the new foundation, becomes a data consumer rather than a fresh integration project, consistent with the two-week versus six-month contrast described in Section 5.

Illustrative outcome range, based on the McKinsey and ABB benchmarks cited in Sections 3 and 6: 30 to 50% reduction in unplanned downtime on the targeted asset class, and a single reconciled inventory figure replacing three conflicting ones. Actual results vary by plant condition, asset age, and program maturity.

A four-phase path to an AI-ready plant

The architecture in Section 5 is not built in one project. Every engagement Grids and Guides Technologies has run follows some version of the same four phases, in the same order, because each phase produces the input the next one needs.

Phase 1 · Assess

A full asset and protocol inventory across every PLC, SCADA system, historian, and network segment; a gap map against the ISA-95 hierarchy; and a security posture review, since the asset inventory this phase produces is the same artifact the Cyber Resilience Act now requires (Section 9). Typical duration: 3 to 6 weeks per plant.

Phase 2 · Architect

Design the Unified Namespace topic structure, select the broker and edge gateway components, and, critically, build the first version of the governed tag dictionary with the engineers who actually know what each tag means today. This phase produces the blueprint, not yet running infrastructure.

Phase 3 · Build

Deploy edge gateways and the namespace on one line or one plant first, not the whole enterprise at once. Ship one AI use case against it, typically predictive maintenance on the highest-downtime asset class, to prove the pipeline end to end before asking the organization to trust it with a second use case.

Phase 4 · Scale

Extend the same namespace pattern line by line, plant by plant. Because the tag dictionary and topic hierarchy are already defined, each additional line is materially faster to onboard than the first; this is the phase where the compounding effect described in Section 6 becomes visible in the numbers.

The most common failure pattern we see. Skipping Phase 1 and 2 and going straight to a Phase 3 pilot on a single line without a namespace or tag dictionary behind it. The pilot works, gets a great demo, and then cannot be replicated on line two without redoing the integration from scratch, because nothing reusable was actually built. This is the mechanism behind the Deloitte finding in Section 6: 84% see value, only 20% scale it.

The regulatory clock is now running

The EU Cyber Resilience Act, Regulation (EU) 2024/2847, entered into force on 10 December 2024 and is the first EU law to embed mandatory cybersecurity requirements directly into product compliance, applying broadly to any "product with digital elements," a scope that explicitly includes industrial control components, connected sensors, and embedded systems, not just consumer IoT.[19]

This deadline is not hypothetical; it is this week. As of 11 September 2026, the CRA's first binding obligation is live: manufacturers must report actively exploited vulnerabilities in in-scope connected products to ENISA and their national CSIRT within 24 hours of becoming aware of them, with a fuller notification due at 72 hours and a final report at 14 days.[19] This applies even to legacy products already on the market. Full CRA compliance, including CE marking tied to cybersecurity and conformity assessment, becomes mandatory on 11 December 2027.

Why this is a data foundation problem, not just a legal one

The CRA requires manufacturers to maintain a software bill of materials, an SBOM, identifying the components in every in-scope product, so that when a vulnerability is disclosed, the organization can determine within hours whether it is affected.[19] That requires, as a prerequisite, an accurate, current inventory of what is actually running across the plant floor, exactly the Phase 1 asset inventory described in Section 8, and exactly the kind of governed visibility the Unified Namespace architecture in Section 5 produces as a byproduct of solving the AI readiness problem.

Products fall into three CRA risk tiers. The default, lower-risk tier, covering roughly 90% of in-scope products including most industrial sensors and general embedded controllers, permits self-assessment. Higher-risk "Important" and "Critical" categories require third-party conformity assessment through a notified body, and industry guidance is already flagging that notified bodies expect significant demand backlogs ahead of the 2027 deadline, making earlier engagement the practical choice for anything that might land in those tiers.[20]

Complementary standards worth tracking

IEC 62443 remains the primary reference standard for securing industrial automation and control systems specifically, and is frequently used as the technical basis for demonstrating CRA conformity in OT environments where no CRA-specific harmonized standard yet exists. The EU's NIS2 Directive, meanwhile, places obligations on the operators of critical infrastructure that use these products, a complementary but distinct set of duties from the CRA's product-manufacturer obligations.

Where to start

1. **Inventory before you integrate.**
   Build the asset, protocol, and data-flow map before selecting any platform or vendor. This single artifact serves both the AI readiness case and the CRA reporting obligation in Section 9; do it once.

2. **Standardize on a namespace pattern, not a tool.**
   Adopt the Unified Namespace / ISA-95 / Sparkplug B pattern as an architectural commitment independent of any single vendor's product, so a broker or gateway swap later does not mean rebuilding the foundation.

3. **Treat context as a deliverable.**
   Fund the governed tag dictionary and semantic mapping explicitly, with named owners, rather than assuming it will emerge as a side effect of the technical integration. It will not.

4. **Pick one asset class, not the whole enterprise, for the first AI use case.**
   Predictive maintenance on the highest-downtime asset class is the most common and best-supported starting point in the published research, and it proves the pipeline before asking for a second use case's budget.

5. **Fold cybersecurity reporting readiness into the same program.**
   The asset inventory and vulnerability visibility the CRA now requires overlaps directly with the data foundation work in Section 5. Running these as two separate initiatives duplicates cost for no benefit.

6. **Measure the foundation, not just the model.**
   Track data completeness, latency, and quality as explicit KPIs alongside model accuracy metrics. A model can look correct in a demo and still be unusable in production if the pipeline feeding it is not monitored the same way the model is.

11 — METHODOLOGY & SOURCES

How this report was built

This report is a secondary-research synthesis, not an original survey. Every statistic above is drawn from a publicly available, independently published study, dated where possible to 2025 or 2026, and attributed to its source inline and below. Where sources reported different figures for a similar metric, such as predictive maintenance downtime reduction, we present the range rather than the single most favorable number. The reference architecture, phased implementation model, and recommendations are Grids and Guides Technologies' own analytical framework, developed through advisory work on OT/IT data integration, presented as a point of view rather than an empirical finding.

1	Redwood Software. *Manufacturing AI and Automation Outlook 2026.* Global survey of 300 manufacturing professionals, fielded by Leger Opinion. January 2026.
2	Cloudera. *The Data Readiness Index: Understanding the Foundations for Successful AI.* Survey of 1,270 global IT leaders, fielded by Researchscape, Jan–Mar 2026. Published April 2026.
3	Gartner, AI project abandonment projection through 2026, as cited in industry analysis of enterprise AI-ready data requirements, 2026.
4	Siemens. *The True Cost of Downtime 2024.* Fortune Global 500 manufacturers, downtime cost and revenue-share analysis.
5	Fortinet. *2025 State of Operational Technology and Cybersecurity Report.* Survey of 550+ OT security professionals globally.
6	UK Cyber Monitoring Centre; industry cybersecurity reporting on the August 2025 Jaguar Land Rover ransomware incident.
7	Precisely. *2026 State of Data Integrity and AI Readiness.* Global survey of data and analytics leaders.
8	IIoT World. *Industrial Data and AI Readiness Survey 2026.* Survey of 272 industrial professionals across manufacturing, energy, pharma, and oil & gas.
9	Enterprise AI-readiness industry analysis, 2026, on the proportion of AI project time spent on data preparation.
10	OxMaint. *The State of Manufacturing Maintenance: 2025 Global Industry Report.* Consolidating data from Siemens, Deloitte, Fluke, McKinsey, ABB, MaintainX, and Aberdeen Research.
11	Deloitte. 2026 Germany manufacturing survey on AI value realization and scaling, as reported in Quality Magazine's 2026 trends analysis.
12	Aberdeen Research. Unplanned downtime cost benchmarking across manufacturing sectors.
13	ABB. Survey of 3,200+ global plant maintenance leaders on unplanned downtime frequency and cost, 2026.
14	Deloitte. *Advanced Manufacturing Report,* plant-level unplanned downtime benchmarking.
15	IBM and industry OT/IT threat analysis, 2026, on intrusion entry points and lateral movement patterns.
16	HiveMQ / Eclipse Foundation. Sparkplug specification documentation and ISA-95/Unified Namespace architectural guidance.
17	TEEPTRAK and HiveMQ. Unified Namespace implementation guidance, MQTT/Sparkplug B bandwidth and architecture benchmarks.
18	IoT Analytics. Predictive maintenance adoption and ROI research.
19	Regulation (EU) 2024/2847 (Cyber Resilience Act); official text and compliance timeline analysis from multiple 2026 industry legal and cybersecurity guidance publications.
20	Industry legal analysis on CRA conformity assessment tiers and notified body capacity, 2026.

12 — ABOUT THIS REPORT & GRIDS AND GUIDES TECHNOLOGIES

Plants run on decisions, not dashboards.

About the study

This report is a secondary-research synthesis on the OT/IT data foundation industrial AI depends on. It draws on more than a dozen independently published 2025 to 2026 studies spanning enterprise AI readiness, industrial data infrastructure, downtime economics, and manufacturing cybersecurity, each cited in full in Section 11, combined with Grids and Guides Technologies' own advisory framework for building that foundation. It is not a primary survey.

About Grids and Guides Technologies

Grids and Guides Technologies builds the production, maintenance, quality, and data systems that connect SCADA, PLC, MES, and ERP into the foundation predictive maintenance, quality AI, and demand-sensing actually require. We don't replace the control layer; we build the data foundation and the AI on top of it.

We work two ways. Bring us the problem and we bring the manufacturing SME, embedded on your floor to learn your systems and find the real gap before anything gets built. Or bring your own SME, and our ML, data science, and engineering team builds directly against their input. The build team is always ours; the floor expertise can be either.

Get more insights

For the full source list behind this report, the reference architecture diagrams, or to start a four-week diagnostic, get in touch or visit our site.

Website: www.gridsandguides.com · **Email:** info@gridsandguides.com