Factory Intelligence Platform Reference Architecture

Aldon P Smith

Published 15 May 2026

Executive Summary

Factories are not suffering from a lack of data. They are suffering from a lack of connected, trustworthy, usable context.

A modern manufacturing site may contain PLCs, SCADA systems, historians, MES platforms, QMS platforms, CMMS tools, ERP systems, laboratory systems, spreadsheets, validation documents, operator notes, and years of tribal knowledge. Each system may serve a legitimate purpose, yet the total picture is often fragmented. Production data lives in one place, quality records in another, maintenance history somewhere else, and validation evidence in static documents.

The Factory Intelligence Platform is a proposed open-source architecture for connecting these systems without replacing them. It is intended to act as a **Factory Intelligence Layer** above the existing manufacturing technology stack. Its job is to make factory data more understandable, contextual, explainable, and useful for the people responsible for production, quality, maintenance, validation, and continuous improvement.

This architecture is especially important in regulated and quality-critical environments. In those settings, intelligence cannot be a black box. Teams need source traceability, data integrity, cybersecurity, validation readiness, auditability, and clear human accountability.

In simple terms: > The Factory Intelligence Platform is an open-source intelligence layer for manufacturing operations. It connects existing systems, builds shared factory context, supports human decision-making, and helps teams adopt AI and analytics in a governed, explainable, and validation-ready way.

The Problem: Factories Have Data, But Not Shared Context

Most manufacturing sites already collect large amounts of information. A historian may contain years of process values. A QMS may contain deviations and CAPA records. A CMMS may contain maintenance history. An MES may contain batch execution information. Operators may know what actually happened during a difficult run, but that knowledge may never become structured data.

The challenge is that these systems usually do not explain one another.

A temperature tag may show an excursion, but the tag alone may not explain which equipment was involved, which batch was running, which phase was active, whether maintenance occurred recently, whether the instrument was calibrated, whether a related alarm occurred, or whether the same pattern has happened before. A quality event may describe the issue, but it may not

automatically bring together the process data, equipment state, historian trends, work orders, and validation context needed to investigate it efficiently.

The result is a factory that is data-rich but context-poor.

The Factory Intelligence Platform is designed to address that gap. It does not attempt to become the new system of record for everything. Instead, it creates a governed layer of context that connects the systems manufacturers already rely on.

Architectural Position

The Factory Intelligence Platform should not replace PLCs, SCADA systems, historians, MES, QMS, CMMS, ERP, LIMS, or other established systems. Those systems have specific responsibilities and, in many environments, validated roles. The platform should respect those boundaries.

The platform should instead provide a shared intelligence layer that can answer questions such as:

- What happened during this manufacturing event?
- Which equipment, batch, material, recipe, or process step was involved?
- What related alarms, maintenance records, quality events, or process changes occurred nearby?
- Is the supporting data complete and trustworthy?
- What source system did this information come from?
- What evidence supports the conclusion?
- What should a human review next?

The architecture is built around the idea that intelligence begins with context. AI and analytics may eventually provide powerful capabilities, but they are only useful if the underlying data relationships are reliable.

Reference Architecture

At a high level, the platform can be understood as a set of connected layers.

```
flowchart TD
    A[Existing Factory Systems] --> B[Integration and Edge Connectors]
    B --> C[Normalization and Contextualization]
    C --> D[Factory Knowledge Layer]
    D --> E[Events and Workflows]
    D --> F[Analytics and AI Services]
    E --> G[Human Review Applications]
    F --> G
    G --> H[Governance, Evidence, and Audit]
    H --> D
```

The lower layers connect to existing systems and preserve source traceability. The middle layers normalize data and build relationships between assets, processes, batches, events, quality records, maintenance activities, and validation artifacts. The upper layers provide applications, workflows, analytics, and AI-assisted features for human review.

The architecture is modular by design. A manufacturer should be able to adopt one component, such as a historian integration pattern or context model, without adopting the entire platform at

once.

Source Systems and Connectors

The first layer of the architecture is the existing manufacturing environment. This includes shop-floor systems, enterprise systems, quality systems, maintenance systems, laboratory systems, and human-generated records.

The platform should connect to these systems through controlled integration patterns. In many environments, the safest starting point is read-only access. A read-only historian connector, for example, can retrieve process data for investigation without modifying source records or influencing process control.

This read-only-first approach is important because it allows the platform to create value without immediately introducing the risk associated with writeback, automated actions, or closed-loop control. Over time, more advanced integrations may be possible, but they should be treated as higher-risk capabilities requiring stronger governance, security review, and validation evidence.

Each connector should have a clear responsibility. It should identify the source system, the type of data it retrieves, the access method used, the authentication model, the configuration baseline, the error handling behavior, and the intended use of the data. Connectors should preserve source references rather than hiding where data came from.

Contextualization and the Factory Knowledge Layer

Raw data becomes useful when it is connected to meaning.

A historian tag by itself may be difficult to interpret. When that tag is mapped to a piece of equipment, an instrument, an engineering unit, a process step, a batch window, a maintenance event, and a quality record, it becomes part of a usable operational picture.

The Factory Knowledge Layer is the part of the platform that manages these relationships. It represents the factory as a connected model of sites, areas, lines, equipment, instruments, tags, materials, products, recipes, batches, procedures, events, work orders, quality records, validation artifacts, and users.

This layer may be implemented using a relational database, a graph database, a semantic model, a document store, or a combination of approaches. The implementation choice is less important than the architectural principle: factory intelligence depends on governed relationships between data, systems, assets, and workflows.

For example, a deviation investigation should be able to start with a quality event and navigate outward to the related batch, equipment, process step, historian tags, alarms, maintenance work orders, calibration records, and relevant procedures. The platform should make that context easier to gather, review, and explain.

Events, Workflows, and Human Review

The platform should treat events as first-class objects. A manufacturing event may come from a source system, such as an alarm, state change, or quality record. It may also be derived by the platform, such as a detected process excursion, repeated equipment trip, or data quality issue.

This distinction matters. A derived event is not the same as an original record from a system of record. The platform should preserve provenance so users can understand whether they are looking at source data, transformed data, inferred context, or AI-assisted output.

The workflow layer connects these events to human review. It should help people investigate, triage, annotate, and act while preserving the boundaries of existing quality and operational systems. In regulated manufacturing, the platform should support human accountability rather than replace it.

A useful early version of the platform might support an investigation workspace where engineers and quality personnel can view time-series data, alarms, batch context, equipment history, maintenance activity, and AI-assisted summaries in one place. The final quality decision would still remain with authorized personnel and the appropriate system of record.

Analytics and AI Services

The intelligence layer can support trend comparison, event correlation, anomaly detection, process review, natural language search, AI-assisted troubleshooting, and validation documentation support.

However, the architecture should be careful about how AI is introduced. AI should not be used to obscure uncertainty, invent context, or make regulated decisions without appropriate controls. Early AI capabilities should focus on helping humans find, summarize, and review relevant information.

The strongest initial use cases are human-in-the-loop capabilities such as retrieving related events, summarizing a process window, identifying missing data, suggesting possible investigation paths, or preparing a review packet. More decision-impacting use cases require higher levels of assurance, validation, governance, and oversight.

The platform should also treat prompts, model versions, retrieval sources, evaluations, and output retention as controlled configuration items when the intended use requires it.

Governance, Evidence, and Auditability

A factory intelligence system must be trustworthy. That trust does not come only from good software architecture. It comes from governance.

The platform should support user roles, permissions, audit logs, configuration history, release evidence, validation artifacts, test evidence, security records, data lineage, model and prompt versions, and known limitations. These capabilities are especially important in cGMP and other regulated environments, where conclusions often need to be supported by objective evidence.

Open-source governance is also part of the architecture. The project should publish its contribution process, security policy, release discipline, issue triage practices, code review expectations, licensing approach, and validation support documentation. This does not make the software automatically validated, but it gives regulated adopters a more credible foundation for their own assessment.

Unified Namespace and Historian Compatibility

The platform should be compatible with modern Unified Namespace concepts while respecting the continued importance of manufacturing historians.

A historian is often the trusted system for high-volume time-series history. A Unified Namespace can provide real-time, structured, publish/subscribe access to operational state and events. These patterns should be treated as complementary, not mutually exclusive.

The Factory Intelligence Platform can consume data from a namespace, publish curated contextual events back into the namespace, and use historian data for historical investigation. It should not encourage manufacturers to publish every raw historian tag into a broker without context or governance.

The goal is to make data more useful, not merely more available.

Responsible Initial Scope

A responsible early version of the Factory Intelligence Platform should focus on observation, context, and human review rather than autonomous control.

The strongest initial scope includes public architecture documentation, open-source governance, a simulated manufacturing dataset, a read-only historian integration pattern, an asset and equipment context model, a tag mapping registry, an event model, human-reviewed investigation workflows, validation support templates, and security documentation.

Capabilities such as direct writeback to PLCs, automated batch release decisions, autonomous deviation closure, electronic signature system-of-record functionality, or closed-loop process control should be treated as out of scope for early releases unless they are separately designed, reviewed, governed, and validated.

Practical Example

Consider a batch process where a temperature excursion occurs during a critical process phase.

In a fragmented environment, the investigation may require manually pulling historian trends, checking alarms, reviewing batch records, asking maintenance about equipment history, confirming calibration status, and searching the QMS for similar events.

With a Factory Intelligence Layer, the user could start with the event and immediately see related process trends, equipment context, batch phase timing, relevant alarms, maintenance activity, calibration references, prior similar events, and data quality notes. An AI assistant could summarize the context and suggest areas for human review, while still showing the source data behind its summary.

The platform does not replace the investigator. It reduces the time required to gather context and improves the quality of the evidence available for review.

Success Criteria

The success of the Factory Intelligence Platform should be measured by whether it helps real manufacturing teams work more effectively and responsibly. A useful platform should make it easier to understand what happened, trace data back to source systems, connect quality and operational context, identify missing or poor-quality data, preserve validation evidence, and introduce AI in a way that is explainable and governed.

The platform should also support incremental adoption. A manufacturer should not need to replace its entire manufacturing stack to benefit from a better intelligence layer.

Conclusion

The Factory Intelligence Platform is a reference architecture for an open-source manufacturing intelligence layer. Its purpose is to connect existing systems, contextualize operational data, support human decision-making, and provide a responsible foundation for analytics and AI in manufacturing environments.

The Open Factory Initiative's architectural position is simple:

> The future of manufacturing intelligence should be open, secure, governed, validation-ready, human-centered, and grounded in real factory systems.

This reference architecture is a starting point for building that future.

Appendix A: Architecture Layers at a Glance

Layer	Primary Role
Source systems	Existing PLC, SCADA, historian, MES, QMS, CMMS, ERP, LIMS, document, and human-input systems
Integration and edge connectors	Controlled access to source systems, preferably read-only at first
Normalization and contextualization	Adds equipment, process, batch, units, source, and quality context
Factory Knowledge Layer	Represents relationships between assets, systems, events, records, and workflows
Events and workflows	Connects operational signals to human review and decision processes
Analytics and AI services	Supports search, summarization, anomaly context, and advisory intelligence
Human review applications	Gives users investigation, context exploration, and review tools
Governance, evidence, and audit	Preserves traceability, security, validation evidence, and change history

Appendix B: Recommended MVP Boundary

A responsible MVP should prioritize read-only integrations, simulated data, context modeling, human-reviewed workflows, validation support templates, security documentation, and transparent open-source governance. It should avoid autonomous quality decisions, closed-loop control, unrestricted writeback to source systems, and unreviewed AI conclusions.

Notes

1. This article was drafted and reviewed by the Open Factory Initiative team. AI tools may have been used for editing, organization, or drafting assistance; final content reflects human review and responsibility.