

Context Matters: Metadata Management in Industrial Intelligence





When it comes to industrial machines, context looms large. Especially when the difference in data signatures between various operating contexts, original equipment manufacturers (OEMs), maintenance histories, and productivity thresholds, reveals whether one asset urgently warrants investigation and another is performing optimally.

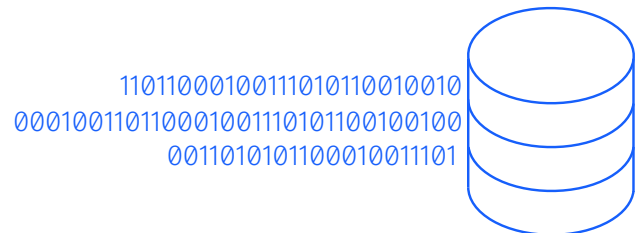
To borrow a pet analogy, it is the difference in knowing that a temperature of 102 degrees Fahrenheit is perfectly normal for a puppy and that a person has a fever.

Context matters, and it is the reason why process-intensive operations in industries like chemicals, oil and gas, mining and metals, manufacturing, and energy and utilities, are facing something like the puppy-or-person question at scale. Metadata is taking on an increasingly important role in industrial intelligence.



Potential issues on valuable industrial equipment loom over asset and personnel utilization decisions. Competitive pressures demand that companies make smarter decisions in each function of their business. And it's not just the balance sheet that shows it.

With environmental, social, and corporate governance (ESG) concerns top-of-mind, intelligence for decision-makers at each echelon of the organization begins with a precise understanding of the metadata — the context. For different stakeholders with various decisions to make, context is (excuse us) contextual.



Right now though, many companies have trouble seeing that context in existing datasets. Much of that difficulty owes to the original design of operational technology (OT) systems like supervisory control and acquisition (SCADA) systems or data historians.

These on-premise collection systems were originally made for plant staff with deep familiarity in asset performance and maintenance. They did not place a premium on context for those who knew how equipment operated. But keeping context, especially for those without a working knowledge of industrial assets, can help teams uncover valuable use cases in asset planning, monitoring, and reporting. It also opens up the possibility of multi-purpose solutions from a single source of industrial data for decision-makers.



More Power Users of OT Data Wanted

Industrial asset expertise is hard to come by and expensive to develop, but it is critical as organizations look to share and delegate decision-making responsibilities to a growing set of data consumers and stakeholders. With the equipment and site expertise necessary to put datasets and industrial assets in their proper context, process-intensive operations are counting on legacy OT systems to extract and make information useful for the rest of the organization.

Power users of these OT systems and IT teams are being called upon to make industrial data accessible to the masses. As a significant portion of the on-premise industrial workforce prepares to retire within the next few years across heavy industries, that dependence will be unworkable. Industrial organizations need a better way.



The payment model of traditional on-premise systems now jeopardizes that progress by favoring a narrow user base of OT data. Since on-premise collection systems require more hardware and licenses in order to process data for more users, many companies must make decisions about who has access to data. These licensing restrictions make it difficult for companies to train various departments in the management of data. It also limits organizations from taking advantage of the breadth of industrial applications available to them once their data is in the cloud.

For business analysts, data analysts, and executives, if a business unit or organization clears those high-cost barriers to an enterprise data lake strategy, they face another, similarly costly challenge of metadata loss. And for engineering and operations leaders looking over the performance of individual sites or facilities, metadata and data quality within and across sources are critical for building out best practices. Shared access to datasets remains expensive, in spite of organizational ownership over data. A new form of metadata management is in order: one that scales and adapts to data consumption needs as the priorities of industrial intelligence change.

““”

OT systems must cater to the growing appetite for smarter business decisions from all corners of the modern industrial organization: management, engineer, technician, data scientist, IT, business analyst.





The Evolving Role of OT Systems

The modern database historian emerged in the 1980s. Industrial equipment, outfitted with better instrumentation and streaming higher and higher volumes of data, outmatched existing record-keeping. Earlier archival systems in individual spreadsheets, including (and sometimes still) those recorded with paper and pen, could not keep up with the volume and velocity of the newly available data. OT engineers adopted the on-premise historian as an archive of single-site information.

The new systems were a goldmine of information. OT environments, in addition to collecting asset-produced information, stamped each data point with descriptive information that distinguishes that figure among other strands of data for each tagged asset. Metadata included attributes like pressure, volume, location, and temperature. Together, they placed industrial asset data into perspective. It was this data about data that enriched the value of a dataset with greater granularity, more accurately reflecting the entire operating context of an asset.

Today, the story around the collection of data in OT systems is much the same. Each of these descriptive points about the data could paint a more holistic view of asset performance.

Except: hardware limitations kept these systems primarily as onsite technologies, suited to the analysis of on-premise assets. The data historian was a feasible solution for plant staff to quickly collect, query, and analyze OT data.

Even so, the traditional on-premise historian ran into a few core challenges to making data available for power users.

1 HIGH COST OF OWNERSHIP

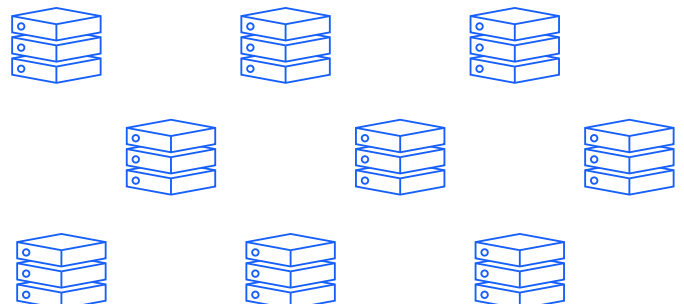
On-premise machines need to be maintained and kept up-to-date with security patches, and there needs to be procedures in case of hardware failures. More data to be processed, by bringing online new assets at an industrial site or more instrumentation, involves the purchase of more hardware. Constant upkeep of the infrastructure for data collection carries a high cost and limits scalability.

2 INFORMATION SILOS

Since they are on-premise systems, data collected in them stays local to the plant. That makes it difficult for industrial data science teams to be effective, especially when they are unable to be onsite – as with the COVID-19 pandemic and going forward too, as more businesses look to enable remote work.

3 LIMITED TOOLS AND ANALYTICAL APPROACHES

Consumers of historian data are limited in the tools they can use for analysis. They are primarily those provided by the vendors of the data historian and may not support analytical approaches or business goals. Because of the proprietary framework of these on-premise systems, this single vendor lock-in can be limiting.





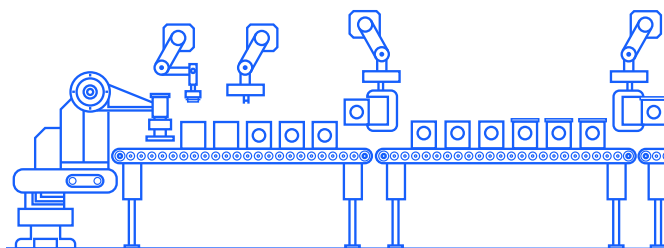
Different Metadata Management Standards Persist

Before the emergence of the cloud, much of the metadata from industrial control systems took on the individual characteristics of data collection at a single site. Since then, varied metadata collection practices have created a dizzying patchwork of different standards. That variation has obscured the visibility of operations into asset performance.

Before the emergence of the cloud, much of the metadata from industrial control systems took on the individual characteristics of data collection at a single site. Since then, varied metadata collection practices have created a dizzying patchwork of different standards. That variation has obscured the visibility of operations into asset performance.

Ideally, metadata management includes detailed and understandable definitions, code values, data quality metrics, and data profiles. For some departments and organizations, this proves to be the case.

As often happens, however, metadata standards are local — most understandable to power users and their on-premise control systems. Beyond the consistent users of these systems, there was no need for translation. Likewise with asset frameworks or taxonomies — individual facilities developed different norms for collection and aggregation.



While these data collection, tagging, and naming conventions are familiar and useful for the power users of on-premise historians, many interested data consumers themselves lack the onsite or engineering backgrounds to know what to make of the metadata. That specificity can render datasets meaningless for people outside the environment.

It would be difficult, for example, to associate one value or another with a specific pump. For others without asset expertise, the difference in temperature readings between a puppy, person, and pump is not so readily apparent. Without a reference system for context (or context for context, if you will), metadata can contribute to imprecision and confusion for the enterprise.

Metadata Compression and Asset Hierarchy Losses

Hardware limitations and differences in metadata management between sites complicated the sharing of datasets by limiting expert access and distribution. The loss of context can be an even more fundamental issue to power users of on-premise systems.

In many historians, preconfigured settings around time-series data compression save licensees from having to buy additional hardware required to handle more metadata. Instead, the historian reduces quality, storing metadata like temperature and pressure at lower granularity — often by reducing the frequency of collection or by having a predefined level of parameters for each of those data tags based on groupings of assets.



In other cases, the historian removes select metadata values entirely from storage. Compression leads to the loss of important variables in the entire operating context. It leaves the enterprise short of the full picture.

Real-time values are just a fraction of the data available to operators. Most organizations also have many years of historical data — and metadata — in their on-premise systems. With enterprise imperatives to migrate that data to the cloud, they face similar challenges to preserve high-resolution metadata and historical data.



Though cloud connectivity add-ons from on-premise collection systems providers have eased the migration of data beyond individual sites, operators have faced challenges with the depreciation of metadata through compression and through lacking support for coexisting data models. Recurring payments for cloud migration and integration services contribute to higher data management costs. And those one-off fees fail to treat the basic issue of metadata loss – the sacrifice of quality for quantity of data.

Asset framework hierarchies are also regular casualties of OT data transfer to the cloud. Users of on-premise systems with cloud connections must consistently rebuild these hierarchies in the cloud. Even so, many data historians cannot support coexisting data models. In the cloud, multiple coexisting data models allow different consumers to see the same dataset within a context that makes sense for their decision-making responsibilities.

Metadata compression and single data models restrict visibility into operations. Integrations and plug-ins in the cloud from on-premise providers to promote data integrity are unable to treat the root issue of unscalable metadata management. And then with different metadata standards, companies lose the context critical for the development of industrial intelligence.

Scaling Metadata for Enterprise Use

As process-intensive businesses turn to the cloud and data lake strategies to leverage the value of their data, the preservation of metadata in the movement of OT data to their cloud environment represents a significant opportunity to optimize the maintenance, productivity, sustainability, and safety of critical assets.

The loss of metadata has been among the most severe limiting factors in the value of OT data. By one estimate, industrial businesses are losing out on 20-30 percent of the value of their data from regular compression of metadata or losses in their asset hierarchy models.

More smart sensors and frequent data collection make data and metadata loss a growing problem. With an expertise shortage sweeping across process-intensive operations, many companies will need to digitize and conserve institutional (puppy-or-person) knowledge, beginning with their own data.

Through an overhead view that organizes and orchestrates that data, data consumers can widen the seemingly narrow differences in asset behavior and bring operational visibility into a crisper, more comprehensive image – with the convenience of a digital solution. When data is connected to the cloud and made more widely available, metadata ensures that datasets are useful, valuable, and reusable for the company at large.



Best Practices for Metadata Management

When industrial intelligence helps asset-intensive organizations make better business decisions, operating conditions are visible at a glance. That clear sense of direction results when metadata, along with traditional OT and IT data, are integrated into tailored, persona-specific views in the cloud. But without defined goals of a metadata management policy to strengthen industrial intelligence, business units run the risk of not meeting their needs. Or they can undermine team confidence in intelligence powered by metadata. Both put metadata management in industrial settings at risk.

A metadata strategy is a key to effectively leveraging industrial intelligence in business decisions. Models of corporate data and metadata governance, like the Data Management Association (DAMA) International's [Guide to the Data Management Body of Knowledge](#), give companies and individual teams management practices to generate the most value from their metadata. Building off of this framework, we share seven fundamental principles that ensure industrial intelligence is backed by equally high-quality metadata in the cloud.



Build a Metadata Management Team

As with any initiative, an effective metadata management strategy depends on leadership. This team is responsible for setting and enforcing metadata management processes that, by working within and across business units, establish the business objectives to guide metadata management, as well as the selection of tactics, tools, and user adoption practices to inspire the improvement necessary.

The team responsible for metadata management would ideally represent the people who will benefit from the tool: senior management, business analysts, engineers and technicians, data scientists, and cybersecurity specialists. For industrial metadata teams, having a three-pronged approach between subject matter experts, data specialists, and business analytics is critical to tying all aspects of the business through unified data management.



Establish Business-Aligned Metadata Goals

- Identifying and setting key business goals are important next steps once the metadata team has the necessary expertise in place.

Benchmarks against productivity, cost avoidance, risk mitigation, and sustainability may guide goal-setting, or in more fundamental measures like the volume of data processed, the integrity and readiness of data for industrial intelligence, or the amount of industrial data rescued for Industrial AI and other advanced analytics. There may also be softer metrics to measure initiative success, including the creation of an agile, nimbler development culture, or even more exploratory projects with in-house industrial intelligence development.



Outline Industrial Intelligence Use Cases

Regardless of the goals the team decides to pursue, their establishment is key to deciding how the team will get there - the use cases for industrial intelligence, including intermediary steps like an enterprise data lake. Among the many possible technologies and analytics that teams can take advantage of are reporting, monitoring, and planning for critical assets. Other opportunities use Industrial AI from the data collected, including metadata, to develop applications like advanced industrial analytics, asset performance management (APM), and digital twins.



Define a Metadata Strategy

Once the team assembled has defined goals and the use cases that will help their business unit and the rest of the organization achieve them, outlining a metadata strategy will help the team set and communicate guidelines for improved metadata management. Given their goals and industrial intelligence use cases, the team should consider the following priorities in their metadata management:

- Metadata requirements (What do different data consumers need to see and how do they need to see it in order to develop, apply, or consume industrial intelligence?)
- Data infrastructure requirements (What will it take to transfer and preserve metadata in a format that is usable for the necessary data consumers?)
 - Current metadata location (Where is the metadata today?)
 - Metadata acquisition and access (How can the team transfer metadata to enable data consumers to use it?)
 - Future metadata location (Where should the metadata be once the strategy is implemented?)
- Ongoing metadata management requirements (What policies and resources are in place to ensure that data consumers continue to develop and use industrial intelligence as easily as possible?)



Adopt Metadata Standards

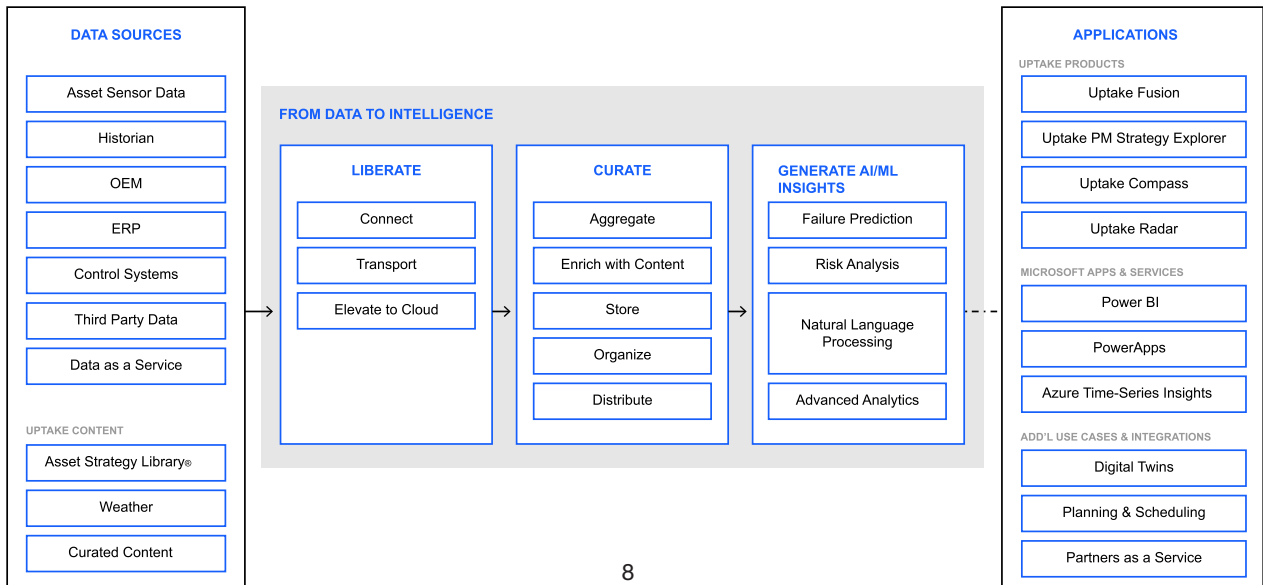
In an industrial organization, metadata standards are critical, and especially for those enterprises with wide-ranging operations from manufacturing to renewables and petrochemical processing. Subject matter experts like reliability engineers, industrial data scientists, and frontline technicians and equipment analysts, are important here as they ensure operation-specific requirements are incorporated in the overall approach.

At the same time, consistent and intuitive metadata entry forms, asset hierarchy, and object models are key to encouraging adoption. Data dictionaries and business term glossaries can help codify the standards for metadata management, as well as verify that these standards are in place once the team has a metadata strategy.



Deploy a Unified Data Management Tool

With agreed-upon goals, resources, requirements, and standards, the next best practice is to select a data management tool that carries out the metadata strategy. Unified data management packages and delivers enterprise data for consumption by business units and individual actors, typically arranging data and metadata for easy query and retrieval. Criteria in the selection of a tool reflect on organizational needs: licensing requirements, the number of users, cybersecurity and adoption specs, use cases, business goals, and management sponsorship.





 **Scale Metadata Management to Meet Enterprise Goals**

Through unified data management, metadata becomes the rich context that data consumers and decision-makers need to see in order to maximize operational performance and minimize business risk. Beyond the initial business unit that leverages robust metadata in their industrial intelligence, the data management tool and team should meet evolving business priorities. Periodic re-evaluation and the resetting of goals enable the team to drive innovation in industrial intelligence forward.

By integrating rich data types from metadata to OT and IT, industrial companies can now build the basis for intelligence free from proprietary constraints. They can answer the puppy-or-person question across their business. Industrial AI, digital twins, operational orchestration – the possibilities to maximize performance and minimize business risk abound once data is in the hands of the enterprise through the cloud.

The days of limited OT data use and tough choices around data preservation, compression, and asset hierarchy losses are through. Unified data management for industrial intelligence is at last enabling teams to make smarter, faster business decisions with Industrial AI – cost-effectively and at scale.

Metadata Powers the New Era of Industrial Intelligence

Industrial organizations have identified the value of their data, know the scale of better decisions possible through cloud computing, but have often struggled with the steps in between. Metadata has been key under lock, stopping organizations from switching on the fuller value chain of industrial intelligence.

Ready to translate your data into smarter operations?

Talk to an Industrial Intelligence Expert

Fusion@Uptake.com or +1 (780)-862-9699

Copyright © 2021 by Uptake Technologies, Inc. All rights reserved. No parts of this document may be distributed, reproduced, transmitted, or stored electronically without Uptake's prior written permission. The "Uptake" mark and related marks are trademarks of Uptake Technologies, Inc. Uptake makes no warranties, express or implied, in this document. Development plans may change, and any discussion of potential features is not a promise of future functionality.

