LNS Research

RESEARCH SPOTLIGHT

Advanced Industrial Analytics:

Building a Next-Gen Operational Architecture with a Data Hub Strategy

www.lnsresearch.com

CELEBRATING 30 years MESA INTERNATIONAL Where Manufacturing Meets IT

Executive Summary

The **2022 edition of the Analytics that Matter research** by LNS Research and MESA briefly introduced the emerging space of Advanced Industrial Analytics. Among other findings, the research revealed that one of the primary differentiators of the top-performing companies- the Analytics Leaders- was investing in data infrastructure and a robust operational architecture.

LNS Research has written extensively about Operational Architecture over the past several years, calling it out as one of the five critical phases of Industrial Transformation. However, many manufacturers pursuing transformation often skip this vital step for several reasons.

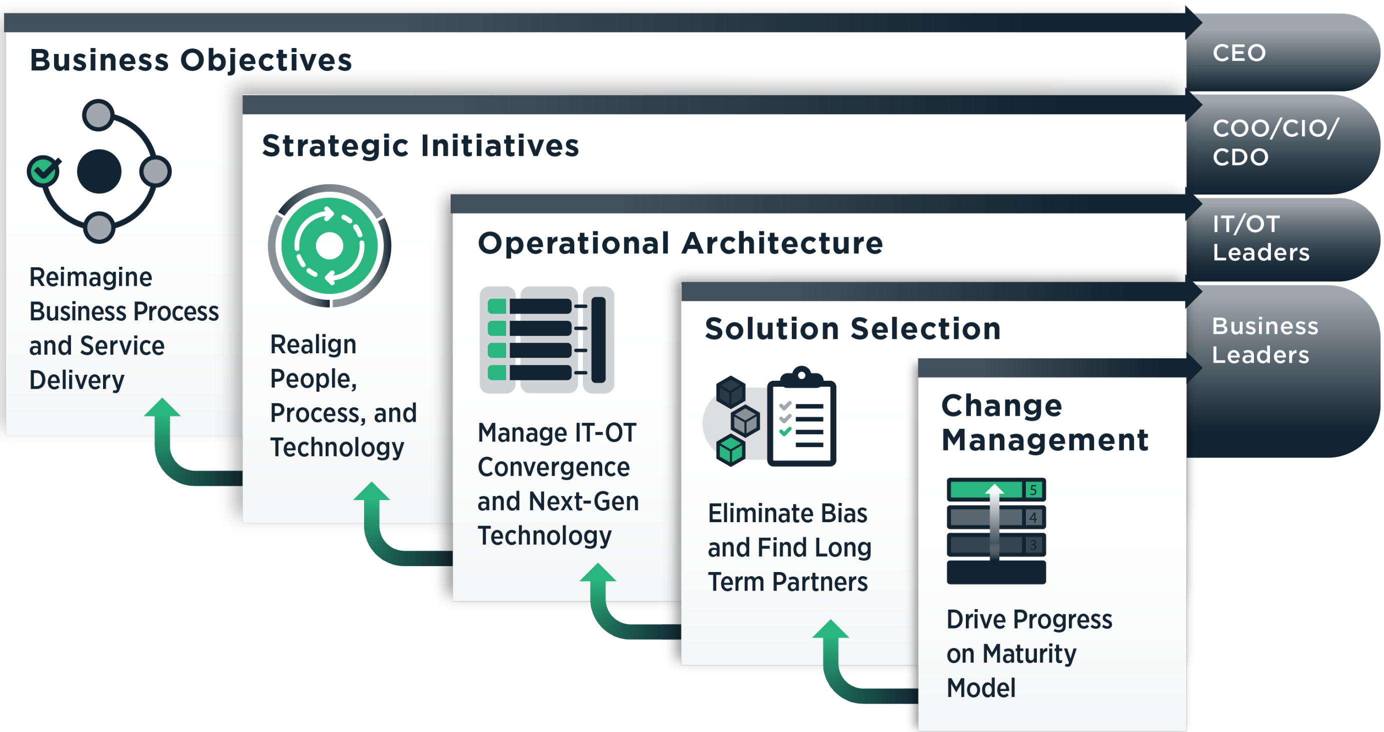


FIGURE 2 - LNS Research Framework

To begin with, one-off architectural investments don't produce immediate value; rather, it takes a series of these investments to move the needle. Even then, while they lay the foundation for future revenue streams, they don't directly yield ROI themselves. Many companies either fail to understand this or don't have the financial freedom to continuously invest in their architecture without immediate returns.

Next, upgrading architectures is an extremely tedious process. As-is assessments and workshops to discuss the shortcomings of the current architecture are typically involved. Then there are the future-state decisions, such as which legacy systems will be replaced and which ones will have to be worked around, which technology vendors to choose, which systems integrators to partner with, etc. More often than not, people's day jobs running the plant get in the way and derail these long-term projects.

And finally, more recently, pandemic-induced shutdowns and suppy chain disruptions have also played their part in forcing manufacturers to prioritize short-term value without making substantial changes to the underlying infrastructure.

As a result, the mainstream manufacturing industry has avoided its architectural problems and continues to be constrained by paper-based processes, limited connectivity, and complex spaghetti architectures. In contrast, the frontrunners have dealt with this problem head-on and built the next generation of operational architectures. In this Research Spotlight, we will describe this next-gen architecture and explain how manufacturers have used a Data Hub strategy to get there.

> The mainstream manufacturing industry has avoided its architectural problems and **continues to be constrained by paper-based processes, limited connectivity, and complex spaghetti architectures.** In contrast, the frontrunners have dealt with this problem head-on and built the next generation of operational architectures.
>
> **- Vivek Murugesan**
> *Senior Research Associate*

The Evolution of Operational Architecture

Simply put, operational architecture extends the well-established principles of enterprise architecture to the operations level. It provides organizations with a framework for integrating manufacturing and control systems with enterprise-level business planning systems.

A robust enterprise and operational architecture is essential to optimize people, process, and technology resources across IT and OT as it helps organizations map business objectives and needs to long-term technology investments, and defines a process to evolve the architecture over time.

Most manufacturing companies today have built their architecture based on the ISA-95 framework, which itself was based on the Purdue reference model. Dating back to the mid-1990s, these models provide hierarchical frameworks to transfer data from the factory floor to the C-suite. Asset and sensor data could pass through control systems, which would feed into data historians and scheduling systems like MES, which then would integrate with ERP and the other enterprise systems.

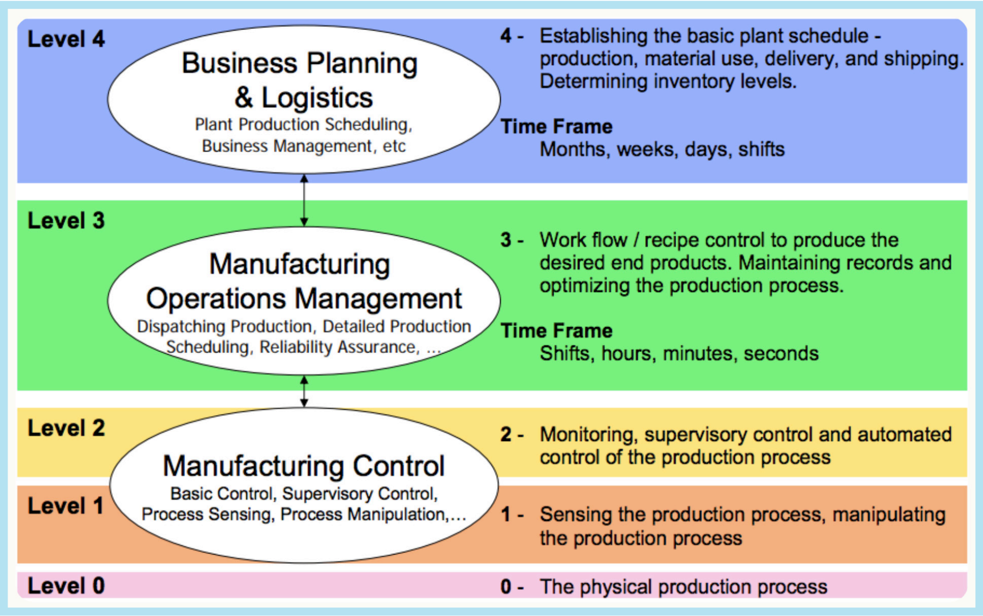


FIGURE 2 - ISA-95 Framework

While this type of hierarchical data flow was ideal for traditional manufacturing, the fourth industrial revolution changed the rules of the game. The proliferation of inexpensive sensors, cyber-physical systems, Industrial IoT, and next-gen automation technologies exponentially increased the volume, variety, and velocity of data generated from factories. Within a few years, factories dealt with terabytes and petabytes of industrial data, eventually paving the way for transformational strategic initiatives and use cases.

These strategic initiatives, such as Factory of the Future, Quality 4.0, Autonomous Plant, etc., promised dramatic step change benefits to operations but required fundamental changes to the existing architecture. Capabilities, such as seamless data integration, IIoT-based machine connectivity, and advanced analytics, suddenly became table stakes.

As a result, the early adopters of these transformation programs began gradually updating their traditional architecture to a flatter, network-based architecture, as shown below.

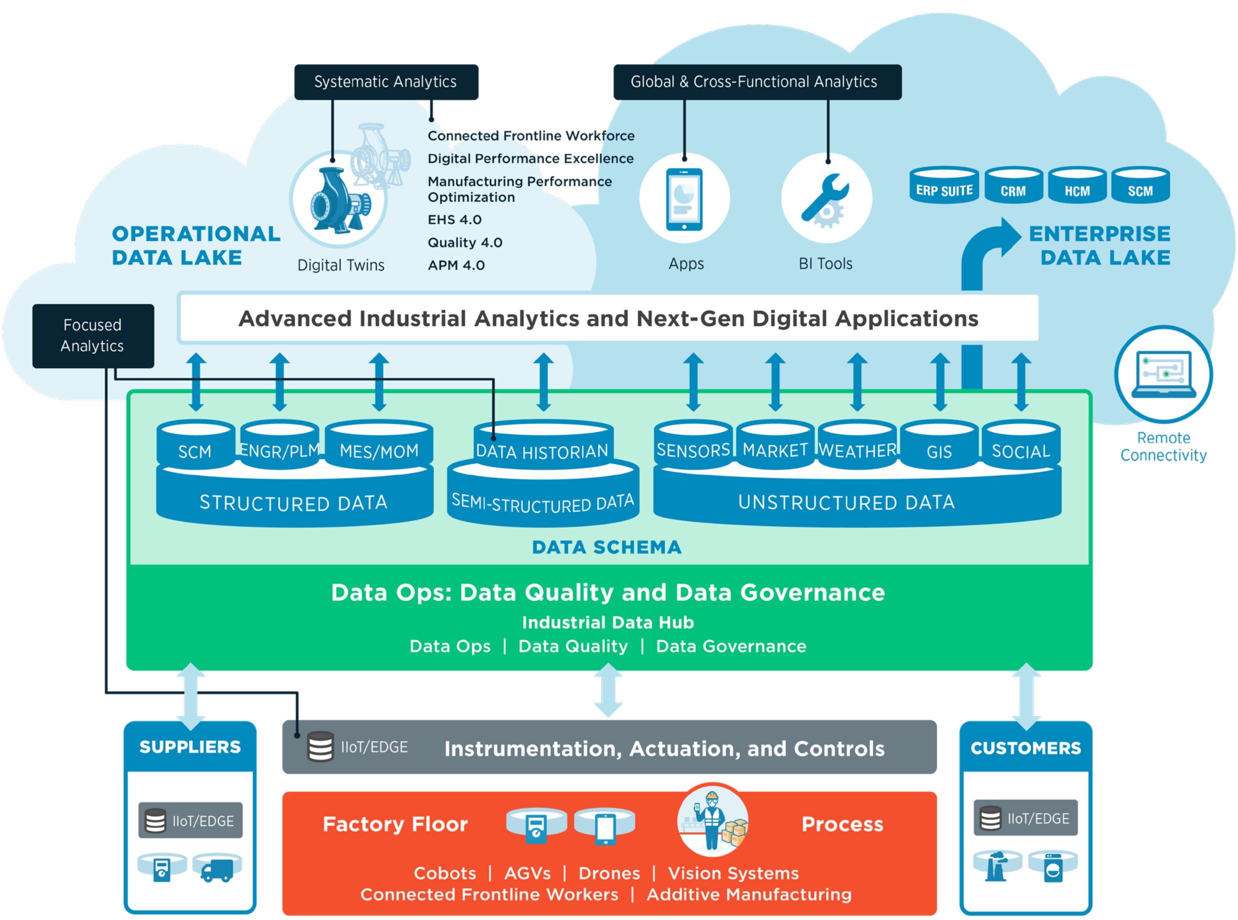


FIGURE 3 - Industry 4.0 Operational Architecture

This newer architecture incorporated significant changes across several levels. Cloud-based applications and platforms were deployed at the enterprise level to enable efficient storage, API-based integration, and advanced batch and streaming data processing and analytics. On the factory side, traditional automation systems, such as PLC, HMI/SCADA, etc., were supplemented with next-gen hyper-automation technologies, including robots/cobots, automated guided vehicles, and vision systems.

Manufacturers who collectively invested in these capabilities were able to address most of the shortcomings of the traditional hierarchical architecture and reaped several benefits along the way, some of which are listed below.

- Asset/sensor data can directly feed into the enterprise layer, bypassing control systems altogether.
- Easier integrations between IT and OT systems. In other words, no more spaghetti diagrams.
- Better data quality by eliminating/reducing manual data entry and paper-based processes.
- Run advanced analytics at multiple levels, along with role-based dashboards and visualizations.
- Efficient asset and process monitoring and collaboration through remote operations centers.
- A digitally enabled, Connected Frontline Workforce with higher levels of safety and productivity.

Moving on to the next section, we will take a closer look at three critical and complementary elements of this next-gen architecture: the common data model – the cornerstone that the architecture is built on, DataOps processes, and the organizational structure that governs it.

Readers should be aware that while these three elements have several interpretations and different ways to architect them, such as Data Hubs, data fabrics, data mesh, unified namespace, etc. For the purpose of this report, we will focus on a specific architecture strategy – the Industrial Data Hub.

The Industrial Data Hub

LNS Research defines the Industrial Data Hub as a software system that collects, stores, and governs industrial data, along with its associated metadata, in a centralized location that natively connects to several data sources across the plant and enterprise, through multiple proprietary and open-source protocols.

While not replacing traditional data storage systems, the Data Hub facilitates symbiotic relationships by enabling interoperability between existing applications and platforms. Primarily sold as a platform, the Data Hub has a set of pre-packaged applications but also provides development tools and libraries for citizen developers to build their own applications quickly.

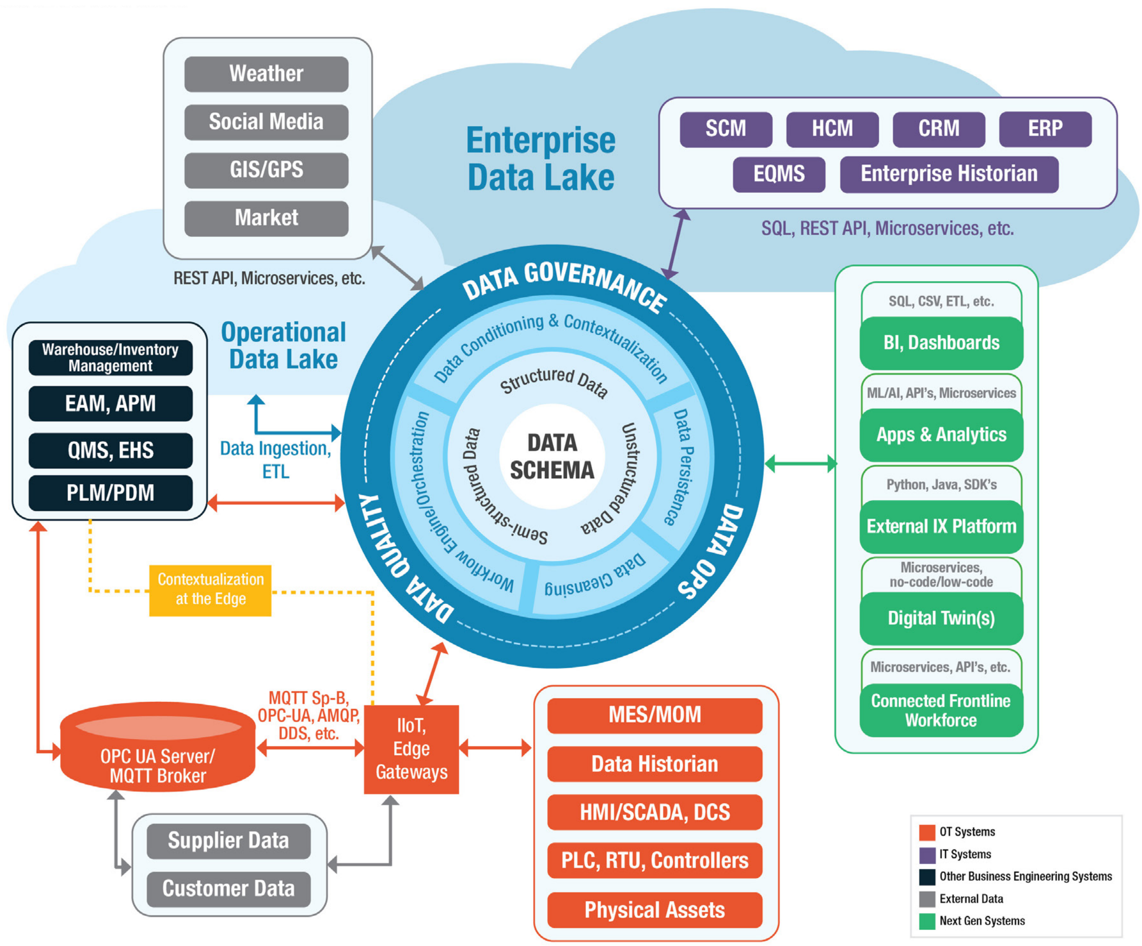


FIGURE 4 - Industrial Data Hub Architecture

As you can see here, the cornerstone of this Data Hub architecture is a data schema of structured, semi-structured, and unstructured data. It also supports governance processes, such as DataOps, to manage high levels of data quality, fidelity, and other attributes. Let's examine these functions of the Data Hub in the next section.

1. The Common Data Model and Schema

Data models are an abstract set of rules, requirements, and visual representations of how data should be defined and structured. Each enterprise, control, and execution system in an organization will have an underlying data model that describes how its data is organized.

By that definition, a common data model is a common set of rules that bring together a diverse set of IT and OT systems. As mentioned previously, this common data model is essential for the systems to integrate and interoperate with each other.

Building this common data model is relatively straightforward on the enterprise side since transactional systems mostly follow standard relational database structures and are often managed by centralized, shared IT resources. However, the operational side is where companies begin to encounter the complexity of bringing together industrial data from control, automation, and execution systems.

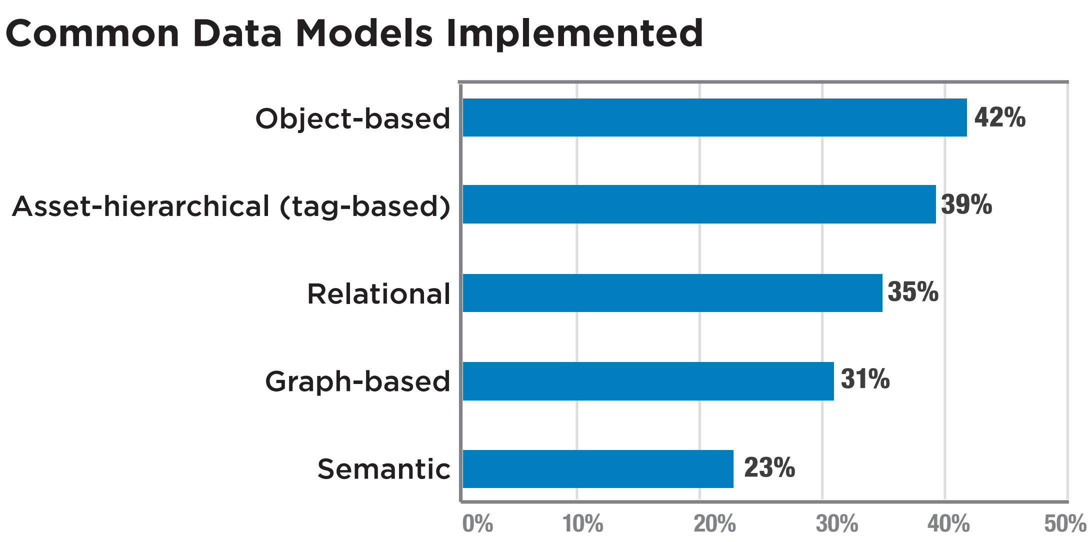


FIGURE 5 - Commonly Implemented Data Models

The main challenge in bringing industrial data together in a common model is its variety. In a typical manufacturing environment, one can expect to find:

- **Object-based** and relational databases in the execution layer – MES, plant-level quality, analytics, etc.
- **Tag-based systems** – the data historians, automation, and control systems that structure data based on timestamps. These timestamped data points typically include material flow, process data, events and alarms, and other units of work based on the mode of manufacturing.
- **Unstructured data** from vision systems (video and image), robots, geospatial, weather, traffic, etc.

Additionally, these systems are mostly unstandardized from plant to plant and have alarming levels of manual entry points and severe data quality issues, making the process of building a common model even more challenging. LNS Research believes that companies can avoid most of these challenges by taking an agile, use case-based approach.

To begin, identify initial use cases and common elements of data across the different systems for those particular use cases. These could be units of assets, materials, time, or work. Then map a data schema for these common elements using network models, such as graphs or semantic models, that are better equipped to manage these many-to-many relationships.

By taking an agile approach here, companies can add or remove elements from the schema as the scope and maturity of the use cases change. This ensures transparency in the data model over time and provides some early wins for the architecture while avoiding spending substantial amounts of time mapping hundreds of different systems upfront.

2. DataOps

While there are numerous definitions for the emerging concept of DataOps in today's industrial market, on a high level, DataOps combines Agile, Lean, and DevOps methodologies to enable collaboration between multiple stakeholders and manage the lifecycle of an organization's data.

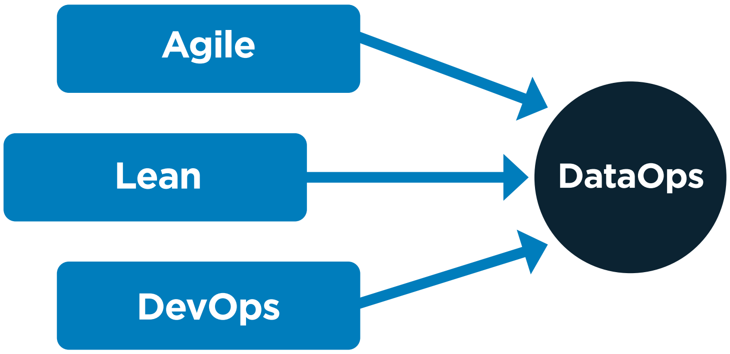


FIGURE 6 - DataOps: Combined Methodologies

Whichever definition you choose to go with, DataOps is nonetheless a critical function of the Industrial Data Hub and the next-gen operational architecture. While most companies who implement DataOps (26% according to a recent LNS Research survey) tend to limit it to just connectivity, it is much more than that.

A well-established DataOps program establishes connectivity but also builds other critical attributes of the data, such as availability, accuracy, trustworthiness, relevance, and timeliness. Some of the common DataOps functionalities that address these issues are:

- **Data availability and accessibility:** This ensures that the right data is available to the right personnel in the right format at the right time. For example, while corporate-level data scientists would require as much raw data as possible, the frontline workforce, on the other hand, would only need limited processed data and specific, prescriptive insights.
- **Data conditioning:** This includes data screening, filtering, interpolating, anomaly detection, etc., to ensure high levels of data quality, which is the number one challenge facing analytics initiatives, according to the Analytics that Matter research.

- **Data Contextualization:** An emerging area gaining quick momentum in the industry, contextualization involves establishing relationships between the data and its metadata. Contextualization can and should be done at multiple levels in the architecture (control systems, edge, cloud), as it helps to close the gap between correlations and causations.
- **Data Persistence:** Persistence maintains the data and its relationships as the source data changes over time. Since most advanced analytics use cases require a rich supply of batch and streaming data, it is necessary to find efficient ways to store historical data and metadata that can be retrieved later.
- **Data Orchestration:** Orchestration is the process of building data pipelines and breaking down existing silos, and organizing them based on the architecture's needs.

3. Organizational Structure & Governance

While DataOps or Data Hub technology vendors provide most of these technological capabilities, manufacturers must supplement them with internally built organizational capabilities as well. In most cases, this begins with the creation of a cross-functional digital council or a steering committee that houses IT, OT, subject matter experts, data scientists, data engineers, and other roles as necessary.

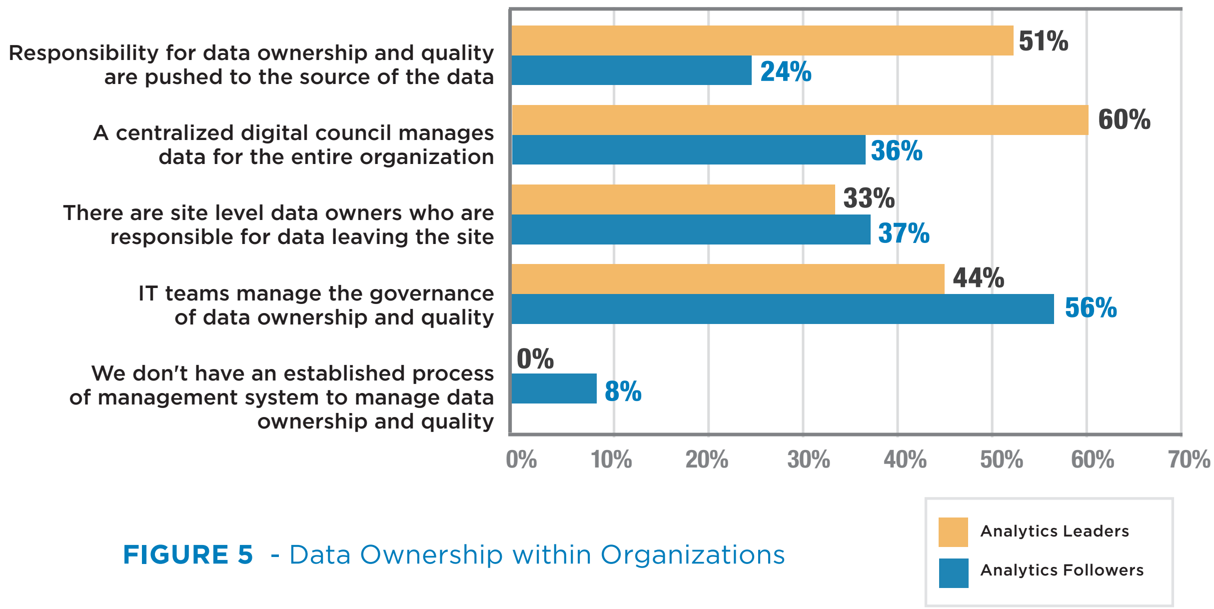
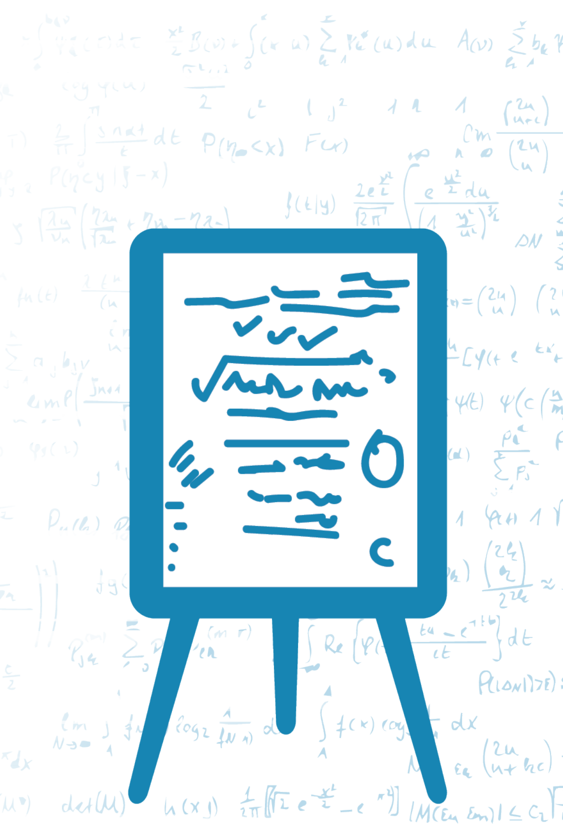


FIGURE 5 - Data Ownership within Organizations

This council or committee should take an active role in nurturing a data-centric culture and making industrial data a valuable asset to the organization. An essential part of promoting this culture is ensuring that providing good data is everyone's responsibility and not just the job of IT or digital teams. While each company should take its own way to build this culture based on its strategic objectives and competitive advantage, some commonalities do exist:

- Pushing data quality responsibilities as close to the source of the data as possible, thereby increasing traceability and ensuring the timely flagging of inconsistencies within the data. While this applies to most cases, there are situations where the source of the data is not often where it gets used the most. In such cases, pushing data quality to the active users of the data would be a better alternative.
- Embedding and integrating data analytics into the decision-making management systems of the company (TPM, TQM, WCM, Lean, etc.). This ensures that the production and operations teams actively leverage the data and analytics in their day-to-day duties.
- Developing a common vocabulary across IT and OT. Surprisingly, the majority of organizational challenges in transformation arise from miscommunication issues. For instance, while IT and digital teams are more aligned and able to effectively communicate using tech TLAs and buzzwords, the rest of the organization might not be. A clearly defined, common vocabulary of transformation terms, such as pilots, minimum viable products, agile, data models, etc., will help streamline communication.
- Bringing the frontline along the journey. It is common for seasoned frontline operators to prefer qualitative ways to run their machines to data-driven decisions. More often than not, this reluctance arises from fear of their jobs being replaced by digital tools. This can be avoided by empowering them through upskilling and career options in a digitally transformed manufacturing environment.



Summary & Recommendations

Traditionally, data and analytics have been built on hierarchical client-server architectures constrained by the ISA-95 framework. While it served its purpose, it also created gaps in the architecture and data silos, which are now addressed by modern network-based architectures in an Industry 4.0 world.

In this Research Spotlight, we discussed the evolution of this operational architecture and how companies are taking an Industrial Data Hub approach to building it. We also explored how companies implement common data models and DataOps to support this next-gen architecture.

These next-gen architectures will not only build a strong foundation for seamless data transfer, IT-OT Convergence, and Advanced Industrial Analytics but also serve as a stepping stone for the next steps in a transformation journey – a truly autonomous manufacturing system.

While some of the early adopters have begun implementing robotic process automation (RPA) technologies to automate repetitive physical processes in factories and back-office functions, a significant hurdle to autonomous manufacturing is the automation of knowledge work.

To address this knowledge gap, LNS Research believes that a robust next-gen operational architecture, supported by widespread connectivity, high-speed networking, a digital thread across the manufacturing network, and a digitally connected frontline workforce, can enable a truly autonomous plant.

Some closing recommendations for manufacturers:

- **Operational architectures don't provide immediate financial returns but don't let that stop you from investing in them.** To overcome this financial constraint, the leaders in this space have turned to co-funding models with contributions from both corporate and OpEx budgets.
- **Factories of the Future don't run just on PLCs and SCADA.** Redefine your factory automation strategy by populating your plant floor with IIoT and edge connectivity, vision systems, automated guided vehicles, robots/cobots, and a connected frontline workforce.
- **Take an agile, use case-based approach to a common data model.** The early adopters of transformation have tried and (mostly) failed in mapping hundreds of different IT and OT systems. Instead, map your data model and schema and iteratively scale it with maturity.
- **A next-gen architecture requires both good data and good culture.** While you can acquire data capabilities from technology providers, a data-centric culture needs to be built from within. Take a DataOps approach to build the required organizational structure and a data-centric culture.
- **There's more than one road to this next-gen architecture.** While some of the larger enterprises can afford to build it themselves, others don't have to. Companies could identify a trusted partner among their PLM, ERP, or automation providers to develop most of the architecture. Another alternative would be to partner with the pureplay Industrial DataOps and Data Hub solutions that are vendor-agnostic and take a best-of-breed approach from there.

RELATED RESEARCH ON INDUSTRIAL TRANSFORMATION (IX)

RESEARCH | Pivot to Value for Industrial Transformation Success →

RESEARCH | Advanced Industrial Analytics: Four Proven Strategies to Scale Transformation During Uncertain Times →

RESEARCH | People in Industrial Transformation (IX): Leadership, Culture, and Organizational Best Practices →

RESEARCH | Taking Control of Quality Transformation: Strategic & Cultural Imperatives for the Quality Executive →

RESEARCH | Next-Generation Sustainability: Risk, Opportunity, and Competitive Advantage →

RESEARCH | Connected Workforce: Enable a Competent, Agile Industrial Workforce →

RESEARCH | Driving Continuous Improvement Through Digital Lean Tools →

RESEARCH | Digital Continuous Improvement in an IX World →

EBOOK | Enable Operational Agility with a Digitally Connected Workforce →

BLOG | Introducing the Industrial Transformation (IX) Reference Architecture →

BLOG | Understanding Industrial Transformation: Definition and Framework for Success →

RESEARCH | Industrial Transformation: Architecture and Analytics Just the Beginning →

RESEARCH | IX Architectural Paths 1 of 3: Three Paths & Understanding IX Infrastructure →

RESEARCH | IX Architectural Paths 2 of 3: Evaluating IX Platforms and IX Applications & Analytics →

RESEARCH | IX Architectural Paths 3 of 3: Looking at IX Strategic Partners →

RESEARCH | Industrial Transformation Success: How to Secure Operations' Buy in to Create Effective Leadership →

RESEARCH | IX Digital Readiness →

AUTHORS:

Vivek Murugesan
Senior Research Associate
vivek.murugesan@lns-global.com

Matthew Littlefield
President & Research Lead
matthew.littlefield@lns-global.com