WHITEPAPER

By Dr. Hossein Azodi, Managing Consultant & Deepak Andayil, Junior Consultant
FrontWell Intelligence Data AI Solutions (FIDAS)

Data Maturity in Life Sciences Manufacturing:

From Siloed Systems to Enterprise-Ready Data Mesh and AI at Scale

Data Maturity in Life Sciences Manufacturing:

From Siloed Systems to Enterprise-Ready Data Mesh and AI at Scale

By Dr. Hossein Azodi, Managing Consultant & Deepak Andayil, Junior Consultant

FrontWell Intelligence Data AI Solutions (FIDAS)

Executive Summary

The amount of data produced in the Pharmaceutical and Life sciences industries by labs, clinical trials, and manufacturing plants, is significant. As the data volume grows, organizations face the challenge to properly manage data to gain the process insights they envision from it. Data trapped in silos, slow compliance processes, and failed AI pilots are often the result of shortcomings in data management.

A structured Data Maturity Assessment provides the path forward. By implementing the data maturity model, organizations are benchmarked on where they are today and how they build a roadmap toward becoming truly data driven. They can go from a data immature/fragmented state to truly AI-Driven organizations. Combined with modern approaches such as Data Mesh and domain-driven data architecture, Pharmaceutical and Life science organizations can transform data from a burden into a business advantage.

"The future of Pharma manufacturing is data-fluent, domain-owned, and AI-ready."

FIDAS, FrontWell Intelligence, Data, and AI Solutions Business Unit, offers the methodology, the tools, and the expertise to help customers progress on moving forward on this path to make sure their data is reliable, compliant, and AI-ready.

1. Introduction

In today's manufacturing environments, critical data is scattered across ERP[1], MES[2], LIMS[3] and SCADA[4] systems (structured), and countless Excel sheets, PDFs, and Word files (unstructured). Even though each system performs its purpose, the underlying silos make it challenging to link data from different processes such as quality control, manufacturing, and planning together. Eventually, scaling of AI or advanced analytics becomes more difficult, if not impossible. Decision-making and compliance slow down due to manual efforts.

The typical silos of a production-IT environment are shown in Figure 1. To overcome these bottlenecks, FIDAS offers a domain-driven design **Data Mesh** approach [1, 2], where data is treated as products. Each "Data Product" is owned and defined - not necessarily by IT but - by the owner of the process where it originates. All the different data products are governed in a unified way to ensure compliance and interoperability. The result is that data from many different areas becomes available in a standardized way across domains.

[1] Enterprise Resource Planning
[2] Manufacturing Execution System
[3] Laboratory Information Management System
[4] Supervisory Control and Data Acquisition

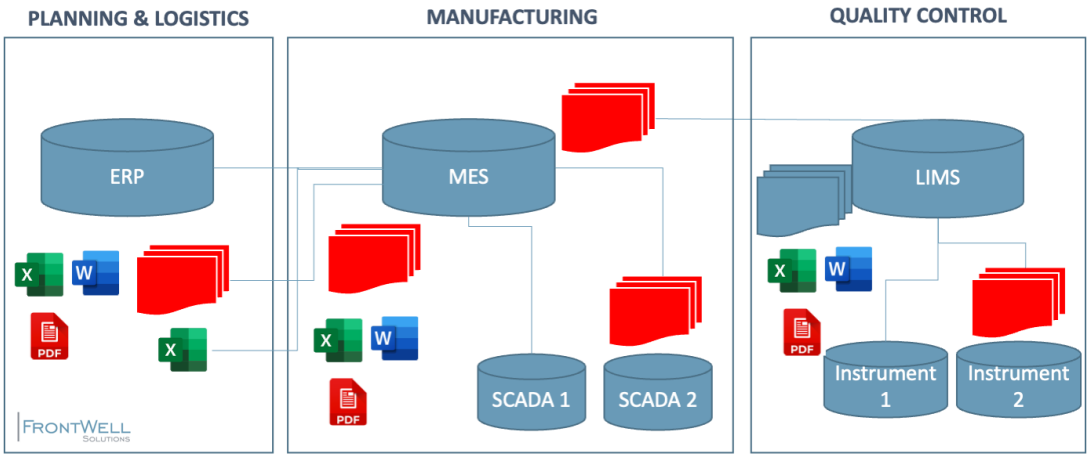

Figure 1 - Typical Data flow in Production IT/Operations Environment

2. Data Immaturity Challenges in Life Sciences Manufacturing

Pharma and Life Sciences organizations suffer from several challenges; they are facing with, due to data immaturity:

A. **Data Silos & Their Risks:**

A Data silo exists when critical information is locked within a single department, system, or platform and cannot be easily shared. In pharma and life science manufactory environments, this is common: clinical trial data sits in EDC systems, manufacturing records in MES, and lab results in LIMS.

A unified view of operations is missing which slows down trial consolidation, delays manufacturing release, and impairs providing cross-function insights.

B. **Data Quality**:

Inconsistent or incomplete data creates risks, especially when adopting AI or advanced analytics. Beyond inconsistency and incompleteness, pharma data suffers from duplication, outdated records, missing metadata, and unclear lineage. Reports are often stacked on top of older reports, hiding the original source and creating technical debt. This slows new report development and complicates audits.

C. **Heavy manual effort**:

Many processes in pharma and life science industries still rely heavily on manual reporting and reconciliation. A common case is when compliance or quality teams build new reports on top of old ones; over time, no one can clearly trace where the numbers originally came from.

Regulations such as FDA, EMA, GxP, and GDPR intensify the challenge. Compliance requires complete, consistent, and traceable data, but siloed systems make this hard to achieve. For example, when preparing a New Drug Application (NDA) for the FDA, companies must integrate data from clinical, manufacturing, and safety systems. If those datasets are spread across disconnected platforms, months of manual reconciliation may be needed, delaying approval and time-to-market. A modification in a piece of data requires another iteration of the entire process, typically manual.

In clinical trials, data managers may copy outputs from different EDC systems into spreadsheets, leading to mismatched results and weeks of reconciliation before an FDA submission.

In manufacturing master batch records are often compared and corrected manually across MES and paper-based logs, delaying product release. Together, these practices not only create unclear data lineage but also build up technical debt, making every new report or submission slower, riskier, and more costly.

D. Impact on AI and ML:

Poor-quality data undermines the value of AI and ML. Predictive models trained on biased or incomplete trial data may generate unreliable risk assessments, while inconsistent manufacturing data can produce false alarms.

Data drift adds another challenge. For example, if an AI model is trained on data from one type of patient group or manufacturing process, it may give poor results when the data drifts over time. Without clear data history, guidance from subject matter experts, and strong governance, companies struggle to move AI from small tests to real, reliable use.

Figure 2 shows a summary of these challenges in Manufacturing environment.

Figure 2 - Challenges in Manufacturing

3. Data Maturity Journey

Data maturity is a continuous journey, not a one-time project. Inspired by the Gartner model [3], we adapt the staged approach to regulated environments like Pharma, showing how organizations can move step by step from scattered data to domain-driven maturity.

A. Descriptive: From Silos to Connected Insights

In Life Sciences similar to other sectors, digital systems like MES, SAP, EDC, and QMS[5] are not typically connected to each other. Clinical trials stay in EDC, production data stays in MES, and so on. When regulators request a consolidated submission such as an NDA or eCTD, the gaps become clear. Teams spend weeks reconciling spreadsheets, reports, and email attachments, with unclear lineage and low confidence in results. Furthermore, creating even an end-to-end view on efficiency excellence report on processes becomes very difficult.

At this stage, data is mainly **descriptive**. It explains what happened, but only after heavy manual work. Insights are slow, compliance is reactive, and critical processes like trial submissions or batch release face unnecessary delays.

Granting access to several data platforms for reporting does not address this issue. As even if the governance in a heterogeneous system becomes manageable, the business know-how of the system as well as a deep insight into columns are missing.

B. Diagnostic: Building Governed Foundations

The first step toward maturity is to bring order to this complexity. Organizations start introducing unified governance: data catalogues that define what the data means, metadata to track lineage, and master data to unify product codes, site identifiers, and trial numbers. Unstructured content like PDF certificates, lab reports, or scanned records is digitized, tagged, and linked to structured systems.

This stage is about creating trust. Reports and submissions become faster to prepare, errors are reduced, and compliance evidence is generated automatically rather than pieced together. Governance doesn't just improve reporting it lays the foundation for what comes next. At this stage, companies can start diagnosing why issues happen for example, linking deviations in MES back to lab results in LIMS because data is cleaner and better connected.

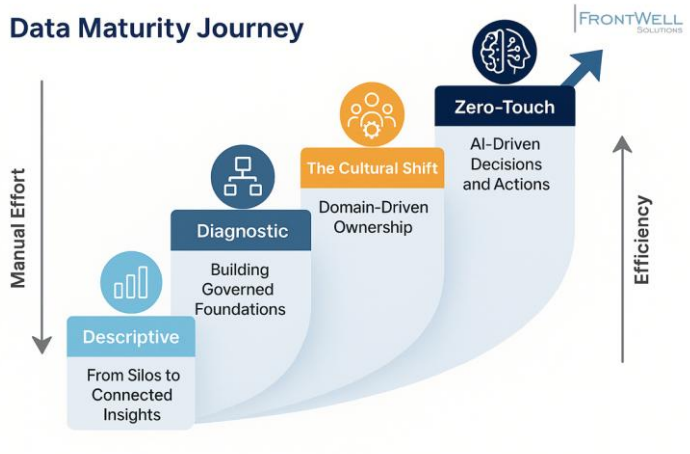

Figure 3 - Data Maturity Stages

[5] Quality Management System

C. The Cultural Shift: Domain-Driven Ownership

The real transformation begins when ownership of data moves closer to the business. Instead of IT or engineers controlling access, domain experts, clinical researchers, QA specialists, manufacturing leads take responsibility for their data and publish it as **data products**. Source data products (lab results, trial visit data, batch records) are governed and standardized. Analytical data products (dashboards, forecasts, risk scores) are derived from these sources.

This change in culture empowers teams to own and trust their data, creates a strong foundation for agility, compliance, and collaboration. Governance becomes federated, so regulatory rules like GxP, Part 11, and GDPR are applied consistently, but without slowing down innovation [4].

D. Zero-Touch: AI-Driven Decisions and Actions

Data in this stage is prepared, cleaned and managed to provide not only predictions or recommendations but also triggering actions. With governed data products and cultural ownership in place, the stage is set for automation. AI can now move beyond pilots and become part of daily operations. In manufacturing, models can detect anomalies in real time and trigger corrective workflows automatically. In clinical trials, patient eligibility models can pre-screen participants and feed results directly into trial systems.

This is where organizations achieve zero-touch or augmented decision-making: data flows directly into action, with little or no human input. But because every AI-driven action is backed by governed data products, decisions remain explainable, transparent, and compliant. The organization shifts from being data-rich but insight-poor, to being data-driven and action-oriented.

4. Domain Driven Architecture and Data Mesh

In a traditional centralized model, data is owned by IT or a single platform team. This approach leads often to the above-described bottlenecks as the organization grows. Data Mesh on the other hand is a modern approach for management of data in large organizations across all domains. Instead of relying on one central data team that collects, cleans, and distributes all the data, Data Mesh shifts responsibility to the business domains themselves (such as Production, Operations, Sales, ...). Each domain becomes the owner of its own data and treats it as a reusable product for herself and other domains.

At the basis, a unified platform team provides shared tools and governance. While domains operate independently, the entire organization follows consistent principles for quality, security, and interoperability.

Types of Data Products:

1. **Source Data Products**

Each source data product within a domain is a dataset published directly from a business domain's operational systems (ERP, CRM, MES, etc.). It represents the raw or lightly processed data that originates in a specific domain. For example, data that is coming from LIMS is stored in "Quality Assurance Domain" in a "Source Data Product" called e.g. "sdp_lims". Similarly, manufacturing, HR, ... data are stored in their corresponding domains and source data products. And while all data is in a unified platform access to it can be easily managed on a data product level.

2. **Analytical Data Products**

On top of the Source Data Products, which are essentially closer to source systems, Analytical Data Products are providing various aggregations, KPIs, etc. from one or more Source Data Products of their domain (or Analytical Data Products of other Domains), to support analytics, reporting, or AI/ML.

5. How FIDAS Supports Its Client in Data Maturity Journey

Pharma organizations need more than just technology to move forward. In fact, they need a clear path. FIDAS provides that path by guiding customers step by step along the data maturity curve, while making sure compliance is never compromised.

A. Assessment: Finding the starting point

It starts with an assessment by mapping where the organization stands today in the Maturity Model presented in section 4. The outcome of such an assessment is a clear roadmap that shows current gaps, achievable milestones and identifies opportunities to unlock the full value of the data captured in day-to-day operations.

B. Governance: Building trust and compliance into data

Right from the start governance is built in as the backbone into the data platform. The Data Platform can be further enhanced with AI modules curated for the organization. This platform is vendor agnostic and is ready to be used on any hyper-scale cloud provider or on-prem.

C. Enablement: Scaling across the enterprises

The Data Mesh approach enables each domain to take ownership and managing its own data products while it establishes shared compliance rules. This combination of autonomy and governance allows the organization to scale the data platform gradually and confidently and avoid bottlenecks.

Designing and creating source and analytical data products for easy usage and in a concessive way is the key to success. These come directly from domains like clinical trials, manufacturing, regulatory, HR, and finance.

6. Conclusion

Data maturity in Pharma, Manufacturing and Life sciences is not just about systems, it's about turning data into a trusted, agile, and compliant asset. By our domain-driven design, organizations can transform source data into insights, insights into action, and action into faster science and better outcomes.

Our Company

FrontWell Solutions is the global front-runners in shaping the digital future of global Life Sciences ecosystems, generating value beyond.We are setting industry standards and shaping the digital future by intelligently connecting data, processes, systems, and people business partner and in the life sciences sector - powered by deep expertise and leading technologies.

FrontWell Solutions is an expert in the digital transformation of the pharmaceutical manufacturing process. Our team of experts is engaged in providing digital solutions to 12 of the 20 leading pharmaceutical, biotechnology, chemical, and medical device companies and suppliers spanning Europe, the United States, and Asia.

Our domain expertise lies across deep and differential specialized consulting services, primarily centered around Manufacturing Execution Systems (MES), Laboratory Information Management Systems (LIMS), Data and AI, Compliance seamlessly integrating these Level 3 systems with Enterprise Resource Planning (ERP) platforms and driving Manufacturing Intelligence initiatives such as Overall Equipment Effectiveness (OEE) reporting.

Moreover, we have partnered with prominent digital solutions platforms in the market, showcasing our proficiency in leveraging cutting-edge technology.

Get in Touch with us

Thinking about taking your next steps towards the digitalization journey? Our experts are ready to support you! Contact us under ReachUs@frontwell-solutions.com or via +49 (6101) 595 89 85.

References

[1] Evans, Eric. 2004. *Domain-Driven Design: Tackling Complexity in the Heart of Software*. Boston: Addison-Wesley.

[2] Machado, Inês Araújo, Carlos Costa, and Maribel Yasmina Santos. 2022. *Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures. Procedia Computer Science* 196: 106–115.

[3] Bitterer, Andreas. 2007. “Gartner’s Data Quality Maturity Model.” Stamford, CT: Gartner Inc.

[4] Petrie, Kevin, Dan Potter, and Itamar Ankorion. 2018. Streaming Change Data Capture: A Foundation for Modern Data Architectures. Sebastopol, CA: O’Reilly Media.