



NARRATE

Regenerative Resilient Smart Manufacturing Networks

D3.1 DESIGN & IMPLEMENTATION OF CONTEXTUALIZATION SERVICES (A)

2025/02/25



Grant Agreement No.: 101138094

Topic: HORIZON-CL4-2023-TWIN-TRANSITION-01-07

Call: HORIZON-CL4-2023-TWIN-TRANSITION-01

Type of action: HORIZON-RIA

D3.1 DESIGN & IMPLEMENTATION OF CONTEXTUALIZATION SERVICES

Work package	WP 3
Task	T3.1
Due date	28/02/2025
Submission date	26/02/2025
Deliverable lead	SERVTECH
Version	1.0
Authors	Michael Papazoglou, Amal Elgammal Bernd Kraemer (SERVTECH)
Reviewers	María José Nuñez (AIDIMME) Fernando Gigante (AIDIMME) Jorge García (MEDWOOD)
Abstract	Modern manufacturing struggles with fragmented data across ERP, PLM, MES, and IIoT systems, limiting scalability and

security. The Federated Data Integration Framework (FDIF) offers a hybrid federated data mesh for Smart Manufacturing Networks by integrating lightweight APIs, edge computing, and digital twins for secure, scalable decentralized data ownership and data exchange. Its adoption requires standardized ontologies, Digital Twin and AI integration, real-world validation, and robust access controls. By replacing monolithic data warehouses with federated architectures, manufacturers gain real-time intelligence, resilience, and agility, driving digital transformation in SMNs.

Keywords	Data Homogenization, Semantic Integration, Production Intelligence, Contextualisation, Federated data architecture
-----------------	--

Project co-funded by the European Commission in the H2020 Programme

Nature of the deliverable:	R
Nature of the deliverable:	P

DOCUMENT REVISION HISTORY

Version	Date	Description of change	List of contributor(s)
0.1	2025/01/20	Definition of the Table of Contents	SERV
0.2	2025/01/31	First draft	SERV
0.3	2025/02/11	Second detailed draft	SERV
0.4	2025/02/18	Review of V0.3	AID, MED
0.5	2025/02/24	Final version incorporating reviewers' comments, Use case and activity diagrams added as well as on a practical guy for pilots regarding the choice and use of standards.	SERV
1.0	2025/02/26	Final revision by the coordination	AID

STATEMENT ON MAINSTREAMING GENDER

The NARRATE consortium is committed to including gender and intersectionality as a transversal aspect in the project's activities. In line with EU guidelines and objectives, all partners – including the authors of this deliverable – recognise the importance of advancing gender analysis and sex-disaggregated data collection in the development of scientific research. Therefore, we commit to paying particular attention to including, monitoring, and periodically evaluating the participation of different genders in all activities developed within the project, including workshops, webinars and events but also surveys, interviews and research, in general. While applying a non-binary approach to data collection and promoting the participation of all genders in the activities, the partners will periodically reflect and inform about the limitations of their approach. Through an iterative learning process, they commit to plan and implement strategies that maximise the inclusion of more intersectional perspectives in their activities.

DISCLAIMER

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.

COPYRIGHT NOTICE

© NARRATE Consortium, 2023 - All rights reserved

No part of this publication may be translated, reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the written permission of the publisher or provided the source is acknowledged.



Except otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

How to cite this report: NARRATE (2024) D3.1 Design & Implementation of Contextualization Services.

The NARRATE Consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	INSTITUTO TECNOLOGICO METALMECANICO, MUEBLE, MADERA, EMBALAJE Y AFINES-AIDIMME	AID	ES
2	SCIENTIFIC ACADEMY FOR SERVICE TECHNOLOGY EV	SERV	DE
3	FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV	FhG	DE
4	INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON	INSA	FR
4.1	INSAVALOR SA	INSA-V	FR
5	SOFTWARE AG	SAG	DE
6	F6S NETWORK IRELAND LIMITED	F6S	IE
7	SYNESIS-SOCIETA CONSORTILE A RESPONSABILITA LIMITATA	SYN	IT
8	MICUNA SL	MIC	ES
9	DHL EXEL SUPPLY CHAIN SPAIN SL	DHL	ES
10	NUNSYS SA	NUN	ES
12	BUDATEC GMBH	BUD	DE
13	BRUNEL UNIVERSITY LONDON	BUL	UK

Table of Contents

- DOCUMENT REVISION HISTORY 2**
- List of figures 7**
- List of tables..... 7**
- Abbreviations..... 8**
- 1. Introduction 10**
 - 1.1 THE CHALLENGE OF DATA FRAGMENTATION IN MANUFACTURING 10**
 - 1.1.1. The Need for Data Normalization and Contextualization10
 - 1.1.2. Moving Beyond Centralized Data Warehousing: Enabling a Federated Data Mesh Hybrid10
 - 1.1.3. Federated Data Mesh Hybrid: A Next-Generation Approach 11
 - 1.1.4. Connecting to Deliverable D3.112
- 2. Metadata & contextual information..... 14**
 - 2.1 HOW METADATA AND CONTEXTUAL INFORMATION INTERACT IN SMNs..... 14**
 - 2.2 STEPS TO USE METADATA IN SMNS 15**
- 3. Federated Data Integration Framework17**
 - 3.1 THE FEDERATED DATA MESH ARCHITECTURE FOR THE FDIF.....17**
 - 3.1.1 Data Autonomy & Decentralized Ownership17
 - 3.1.2 Avoidance of Data Replication & Data Silos.....17
 - 3.1.3 Reducing Bandwidth & Avoiding Unnecessary Data Transfers18
 - 3.1.4 Domain-Specific Data Categorization18
 - 3.2 FEDERATED DATA SOURCES LAYER 19**
 - 3.3 METADATA & SEMANTIC LAYER.....20**
 - 3.4 DATA ASSEMBLY LAYER (SUPPORTS NARRATE DIGITAL TWINS & THE IMC)..... 20**
 - 3.5 DATA CATALOGUING AND DISCOVERY 23**
 - 3.5.1 Key Functions of Data Cataloguing & Discovery in FDIF 23
 - 3.6 FDIF’S ROLE IN NARRATE’S SMART MANUFACTURING NETWORK..... 24**
- 4. Standards for contextualizing data 26**
 - 4.1 STANDARDS FOR BOM AND MBOM26**
 - 4.2 STANDARDS FOR HISTORICAL MANUFACTURING AND PROCESS DATA 27**
 - 4.3 INTEROPERABILITY STANDARDS 29**
 - 4.3.1 ISO 15926: The Standard for Interoperability in Industrial Data 29
 - 4.3.2 MTConnect 30
 - 4.4. STANDARDS FOR IOT SENSOR DATA 32**
 - 4.4.1 SensorML (OGC) 32

- 4.4.2 MQTT Metadata Extensions 33
- 4.4.3 Key Takeways 35
- 4.4.4 Industrial Data Standardization: A Practical Guide for Pilots 37
- 5. Putting things together 38
 - 5.1 EXAMPLE: CNC-DRIVEN DATA FLOW WITHIN THE FDIF 38
 - 5.1.1. Three Layers of the FDIF for CNC Machine Data Integration..... 38
- 6. Conclusions & Future Work..... 41
 - 6.1 CONCLUSIONS 41
 - 6.2 FUTURE WORK..... 41
- 7. References 43

LIST OF FIGURES

Figure 1 The Federated Data Integration Framework. 19

Figure 2 Use-case diagram for the IMC-driven adaptive manufacturing scenario. 22

Figure 3 Sequence diagram for the IMC-driven adaptive manufacturing scenario.....23

LIST OF TABLES

Table 1 Metadata categorization..... 15

Table 2 Summary of the FDIF’s role in NARRATE’s Smart Manufacturing Network. 24

Table 3 STEP AP239 Example File (Extract)..... 26

Table 4 CNC_Machine_01 (UPC UA Object).....28

Table 5 Potential Implementation Tools for MTConnect in FDIF 31

Table 6 A CNC Machine Vibration Sensor in sensor-ML.32

Table 7 How MQTT Metadata Works in an SMN (Using NARRATE Blueprint Frames). 34

Table 8 Example MQTT Payload with Metadata Extensions. 34

Table 9 BLUEPRINT functions & MQTT metadata role..... 34

Table 10 FDIF layers & the role of MQTT.35

Table 11 Comparative analysis of data contextualization standards.35

ABBREVIATIONS

AI	Artificial Intelligence
API	Application Programming Interface
BOM	Bill of Materials
CA	Consortium Agreement
CDIF	Cross-Domain Interoperability Framework
CEP	Complex Event Processing
CFX	Connected Factory Exchange
DL	Deliverable leader
DM	Data Manager
DT	Digital Twin
EC	European Commission
ERP	Enterprise Resource Planning
FAIR	Findability, Accessibility, Interoperability, and Reusability
FDIF	Federated Data Integration Framework
IMC	Intelligent Manufacturing Custodian
IoT	Internet of Things
IIoT	Industrial Internet of Things
ISO	International Organization for Standardization
KPI	Key Performance Indicator
MaaS	Manufacturing as a Service
MBOM	Manufacturing Bill of Materials
MES	Manufacturing Execution System
PLM	Product Lifecycle Management
REST	Representational State Transfer
SMN	Smart Manufacturing Network
Tx.x	Task x.x
WP	Work Package

EXECUTIVE SUMMARY

Modern manufacturing generates vast, fragmented data across ERP, PLM, MES, and IIoT systems, which often operate in silos with incompatible formats. Traditional centralized data warehouses struggle with scalability, latency, and security risks, highlighting the need for a more flexible approach.

This deliverable introduces the *Federated Data Integration Framework (FDIF)*, a hybrid federated data mesh for Smart Manufacturing Networks (SMNs). FDIF enables real-time interoperability, decentralized data ownership, and contextualized analytics across diverse manufacturing data sources. Key features include:

- *Decentralized yet interconnected data sources*, reducing redundancy while maintaining autonomy.
- *Contextualized data enrichment* for more insightful analytics.
- *Lightweight, scalable APIs* for secure, real-time data exchange.
- *Edge computing* to process data at the source, minimizing large-scale transfers.
- *Digital twin synchronization* for disruption prediction and workflow optimization.

For full-scale FDIF adoption in SMNs, manufacturers should:

1. implement *standardized ontologies* (e.g., CDIF, ISO 15926) for interoperability,
2. deploy *metadata-driven APIs*,
3. have the ability to *integrate AI and digital twins*, validate FDIF in real-world applications (e.g., furniture supply chains, smart factories), and
4. *establish robust access controls* to ensure data sovereignty and compliance.

By shifting from monolithic data warehouses to *federated data mesh architectures*, manufacturers can unlock real-time intelligence, enhance resilience, and accelerate digital transformation. FDIF paves the way for self-orchestrated Smart Manufacturing Networks, where decentralized data ecosystems drive agility, efficiency, and innovation.

1. INTRODUCTION

1.1 THE CHALLENGE OF DATA FRAGMENTATION IN MANUFACTURING

Modern manufacturing environments generate vast and diverse datasets, originating from multiple sources, such as:

- Enterprise *Resource* Planning (ERP) systems, managing procurement, inventory, and financials.
- Product Lifecycle *Management* (PLM) platforms, tracking product design, specifications, and engineering changes.
- Manufacturing *Execution* Systems (MES), capturing real-time production data, machine performance, and process execution.
- Industrial IoT (IIoT) sensors and edge devices, monitoring machine conditions, environmental factors, and *operational* anomalies.

Supply chain management systems, providing logistics, supplier performance, and order fulfilment data.

Each of these systems is designed for a specific function, often developed by different vendors, using incompatible data formats, structures, and protocols. As a result, manufacturers struggle with data silos, where valuable information remains isolated within separate platforms, making it difficult to achieve a holistic view of data and operations. This lack of interoperability impairs predictive decision-making, increases inefficiencies, and slows responses to supply chain disruptions.

1.1.1. The Need for Data Normalization and Contextualization

To enable real-time visibility and operational intelligence, manufacturers must move beyond merely collecting data and focus on:

- *Normalization* – Converting heterogeneous data from various systems into a standardized format that can be easily compared and analysed and allows data from disparate sources to be amalgamated.
- *Contextualization* – Enriching raw data by linking it to relevant operational parameters, such as production schedules, quality metrics, and machine conditions.
- *Federated Integration* – Allowing data to remain in its source systems while enabling seamless access and synthesis through a federated data mesh hybrid rather than a traditional centralized data warehouse.

1.1.2. Moving Beyond Centralized Data Warehousing: Enabling a Federated Data Mesh Hybrid

Historically, manufacturers sought to overcome data fragmentation by consolidating vast amounts of information into centralized data warehouses. These repositories were designed to serve as a single

source of truth, integrating data from Enterprise Resource Planning (ERP), Product Lifecycle Management (PLM), Manufacturing Execution Systems (MES), and Industrial IoT (IIoT) devices. However, as modern manufacturing operations expand across geographically distributed factories, multi-tier supplier networks, and automated production systems, these monolithic architectures fail to meet real-time operational demands due to several limitations:

- **Scalability Issues** – As manufacturing operations generate exponentially increasing data streams, centralized repositories become expensive and difficult to maintain. High storage and processing costs, coupled with the need for frequent system upgrades, make traditional data warehouses impractical for large-scale industrial environments.
- **Latency & Access Constraints** – Moving large datasets across geographically dispersed factories and suppliers introduces data synchronization delays and bandwidth limitations. These constraints hinder real-time analytics, making it difficult for manufacturers to react swiftly to operational changes, supply chain disruptions, or equipment failures.
- **Inflexibility** – Traditional warehouses rely on predefined data models that require significant restructuring when new data sources, formats, or operational requirements emerge. This lack of adaptability impedes manufacturers from integrating modern data-driven technologies such as AI-powered analytics, Digital Twins, and predictive maintenance systems.
- **Single Point of Failure** – If the system goes down due to hardware failure, software issues, or cyberattacks, the entire organization’s data operations can be disrupted. In addition, downtime can lead to significant business disruptions, especially for organizations that rely on real-time data for decision-making.
- **Security Risk and Compliance Issues** – Centralized systems are attractive targets for cyberattacks. A breach can compromise sensitive data across the entire organization. Storing all data in one location can make it harder to comply with regional data privacy regulations (e.g., GDPR, CCPA), which may require data to be stored locally.

1.1.3. Federated Data Mesh Hybrid: A Next-Generation Approach

To address these limitations, a federated data mesh hybrid provides a scalable, flexible, and decentralized approach to data aggregation and interoperability in smart manufacturing ecosystems. Unlike traditional centralized approaches, a federated data mesh does not require all data to be physically transferred to a single repository. Instead, it:

- Distributes data ownership across different systems, allowing manufacturing sites, suppliers, and logistics partners to manage their data autonomously while ensuring interoperability.
- Enables real-time data exchange by connecting decentralized datasets through standardized APIs, ontologies, and AI-driven data synthesis tools.
- Facilitates contextualized data analysis, where information is processed at the source, reducing the need for large-scale data transfers while enhancing decision-making capabilities.
- Supports heterogeneous data formats, allowing seamless integration of structured (e.g., relational databases), semi-structured (e.g., JSON, XML), and unstructured (e.g., sensor logs) data.

1.1.4. Connecting to Deliverable D3.1

Objective of D3.1 is to design a context that faithfully represents production data and operational events spread across different production sites within a smart manufacturing network in a reusable structure grounded in open standards using data categorisation and metadata tagging. One of the primary accomplishments of this deliverable is the Federated Data Integration Framework (FDIF) for SMNs, a structured approach to implementing a federated data mesh hybrid in real-world manufacturing environments. Integrating metadata and contextual annotations is essential for effectively implementing the FDIF within Smart Manufacturing Networks. In the following we explain how each component of the FDIF incorporates these elements:

1. Define Standardized Ontologies

Establishing a common semantic data model involves creating standardized ontologies that harmonize data definitions across various systems, ensuring consistency in interpretation and usability. By annotating data with metadata aligned to these ontologies, we facilitate seamless integration and interoperability. For instance, the Cross-Domain Interoperability Framework (CDIF cdif.codata.org) supports FAIR implementation by establishing a 'lingua franca' for information, based on existing standards and technologies to support interoperability in both human- and machine-actionable fashion [1]. CDIF is a set of implementation recommendations, based on profiles of common, domain-neutral metadata standards which are aligned to work together to support core functions that are required by FAIR [2].

2. Develop API-based Interoperability Solutions

Implementing lightweight, scalable APIs enables seamless data exchange between manufacturing partners while maintaining security and access controls. These APIs should be designed to handle metadata effectively, ensuring that contextual information accompanies data as it moves across systems. This approach aligns with metadata-driven data integration frameworks, which emphasize the importance of metadata as the foundation for describing data assets, including their structure, content, quality, lineage, and usage [3].

3. Enable Contextualized Data Aggregation

Utilizing data synthesis tools allows for processing and enriching data at the source, reducing dependence on bulk data transfers. By annotating data with contextual metadata during this process, we enhance its meaning and relevance, facilitating more effective aggregation and analysis [3]. A generic framework for the semantic contextualization of indicators proposes structuring and linking distributed datasets with contextual metadata according to a standard model, which can be applied here to ensure consistency and interoperability.

4. Validate Through Pilot Implementations

Testing and refining the federated data mesh hybrid in real-world manufacturing environments - such as bespoke baby cot production networks, automotive supply chains, and smart factory ecosystems - provides valuable insights. In the pilots in WP-5, metadata and contextual annotations play a critical role in ensuring that data from diverse sources can be integrated and interpreted correctly, leading to more informed decision-making. For instance, in a harvesting/federated

catalogue system, metadata about the metadata (such as its source, format, and update dates) is useful to keep track of where metadata came from and what format/profile it uses, which is essential for processing.

5. Provide Decision-Support Schemes

Delivering industry-specific reference architectures and best practices helps manufacturers adopt federated data strategies effectively. These schemes emphasize the importance of metadata management and contextual annotations, guiding organizations in implementing systems that capture, store, and utilize metadata to enhance data integration and decision-making processes. Metadata management is the critical foundation for data leverage, enabling organizations to effectively manage and utilize their data assets.

By embedding metadata and contextual annotations throughout the FDIF, we ensure that data within SMNs is not only integrated but also enriched with the necessary context to make it meaningful and actionable.

By shifting from monolithic data warehousing to a federated data mesh hybrid, manufacturers can unlock real-time insights, enhance supply chain resilience, and accelerate the adoption of Smart Manufacturing Networks. This deliverable will be a practical foundation for researchers and industries seeking to transition towards more *adaptive, decentralized, and data-driven* manufacturing ecosystems.

2. METADATA & CONTEXTUAL INFORMATION

Metadata and contextual information refer to *higher-level descriptive data* and *situational awareness* that help define and refine manufacturing processes. Metadata types can be categorized as:

- **Descriptive Metadata:** Information to identify and describe data (e.g., part names, machine IDs, timestamps).
- **Structural Metadata:** How data is organized (e.g., hierarchy in BOM, relationships in MBOM).
- **Administrative Metadata:** Details about the data's creation, ownership, versioning, and access rights.
- **Provenance & Data Lineage Metadata:** Records of data origin and processing history (e.g., source of historical manufacturing data).
- **Operational Metadata:** such as machine uptime/downtime data, supplier lead times, and energy consumption patterns.

Contextual types can be categorized as:

- **Production context:** Real-time production environment conditions (temperature, machine status, supplier delays, and workforce availability)
- **Customization requirements:** Customer-driven requirements and personalization needs.
- **Supply Chain context:** such as external supplier conditions, transportation logistics and geopolitical factors, shipping delays triggering, demand fluctuations adjusting procurement strategies, and risk mitigation plans.
- **Environmental Context:** Sustainability goals, energy efficiency, and environmental regulations such as carbon footprint tracking, energy-efficient routing of production steps and sustainable material sourcing.
- **Regulatory & Compliance Context:** Safety standards, and industry regulations that impact production and distribution, such as material safety standards.

2.1 HOW METADATA AND CONTEXTUAL INFORMATION INTERACT IN SMNs

In summary, metadata-driven data integration is better than traditional data integration because it provides a standardized, automated, flexible, and easier-to-maintain approach to integrating data from multiple sources that can be scaled to meet the growing manufacturing data integration needs.

In the NARRATE Digital Twin architecture, metadata and contextual information serve distinct yet complementary roles in ensuring both stability and flexibility within a Smart Manufacturing Network (SMN).

- Metadata provides a structured, repeatable framework for defining archetypes, ensuring consistency, governance, and adherence to industry standards. Archetypes act as templates that encapsulate best practices and standardized processes.
- Contextual information, on the other hand, introduces dynamism by adapting these archetypes in real time based on evolving conditions, allowing for agile decision-making and optimized operations.

Illustrative Example: Baby Cot Assembly Process

Below we discuss how metadata could be used in an illustrative example.

1. Metadata Defines a Process Archetype

- A Process Archetype is established for assembling a baby cot, outlining the standard sequence of operations, required resources, and predefined quality control parameters.
- This ensures repeatability across multiple production cycles and facilities, maintaining process integrity and compliance.

2. Contextual Information Enables Dynamic Adaptation

- Machine Status: If a key assembly machine undergoes maintenance, the system automatically reroutes production to an alternative workstation.
- Supply Chain Delays: If a delayed component shipment affects production, the system adjusts the schedule or suggests an alternative supplier.
- Customer Design Modifications: If a customer submits a last-minute customization request, the system reconfigures assembly parameters while ensuring feasibility.

By integrating stable metadata structures with real-time contextual intelligence, Blueprint Frames—which serve as the foundational realization of Digital Twins in NARRATE (see D3.2) —enable adaptive, resilient, and self-optimizing manufacturing processes. This ensures that Smart Manufacturing Networks remain agile, efficient, and responsive to evolving production demands.

Furthermore, this dynamic interaction empowers NARRATE’s Intelligent Manufacturing Custodian (IMC) to:

- Predict potential disruptions before they impact operations.
- Optimize workflows by dynamically adjusting processes in response to real-time data.
- Enhance resilience by proactively mitigating risks and ensuring seamless coordination across the manufacturing ecosystem.

Through this synergy, NARRATE fosters a smart, data-driven manufacturing environment, where Digital Twins and AI-driven analytics work in unison to achieve operational excellence and supply chain robustness.

2.2 STEPS TO USE METADATA IN SMNS

The following are the steps of using metadata effectively to support Digital Twins in NARRATE.

1. Define Metadata Categories

To ensure a structured approach, metadata must be categorized appropriately as shown in Table 1.

TABLE 1 METADATA CATEGORIZATION.

Metadata Type	Description	Example in SMNs	Implementation Tools
Descriptive Metadata	Identifies and describes manufacturing data.	- Machine IDs, part names, timestamps - Product specifications	- Manufacturing Execution Systems (MES) - ERP Systems - Data catalogs (Apache Atlas, Alation)

Structural Metadata	Defines relationships and data hierarchy.	- Bill of Materials (BOM), MBOM	- PLM Systems
Administrative Metadata	Governs data creation, ownership, versioning, and access.	- Version history - User permissions	- Data governance platforms (Collibra) - Version control (Git, DVC for ML models)
Provenance Metadata	Tracks data origin, lineage, and processing history.	- IoT sensor data timestamps - Supplier data lineage	- Data provenance tools (Databricks Unity Catalog)

2. Annotate Data with Context

Metadata should include *contextual information* that enhances interpretability and traceability.

Implementation Steps:

- Attach metadata to datasets describing their purpose, source, and unit of measurement.

Example: Annotate IoT sensor data with:

- Timestamp (2025-02-08T12:30:00Z)
- Sensor location (Machine A5 - Assembly Line 3)
- Unit of measurement (°C)
- Operational status (Nominal)

Potential Implementation Tools:

- **Metadata Tagging Systems:** Apache Atlas, OpenMetadata

3. Implement Metadata Standards

Adopting industry standards ensures interoperability and data consistency across systems.

The following standards could be used:

Standard	Use Case
ISO 10303 (STEP)	Representing and exchanging product data, including BOM and MBOM.
OPC UA	Machine-to-machine communication and metadata exchange
MTCConnect	Standardized machine data collection

Potential Implementation Tools:

- **Schema Validation Tools:** JSON Schema
- **Interoperability Frameworks:** OPC UA Frameworks (UAExpert, Prosys OPC UA)

4. Store Metadata Alongside Data

Metadata should be *co-located* with manufacturing data for easy retrieval and governance.

Implementation Strategies:

Store metadata in **structured formats** such as:

- JSON (e.g., for REST APIs)
- RDF (for linked data in ontologies)

3. FEDERATED DATA INTEGRATION FRAMEWORK

The cornerstone of Deliverable D3.1 is the Federated Data Integration Framework — a robust, scalable hybrid federated data mesh data management architecture designed to enable seamless, secure, and intelligent data exchange across distributed, heterogeneous data sources in a Smart Manufacturing Network. FDIF integrates heterogeneous data from factories, suppliers, logistics, IoT devices, and enterprise systems, ensuring real-time interoperability, data sovereignty, and resilience within the manufacturing ecosystem. By leveraging event-driven architectures, semantic data models, and federated governance mechanisms, FDIF provides the source data for digital twins in NARRATE (see deliverable D3.2) and a robust data management foundation that supports important IMC functions including automated decision-making, predictive analytics, and adaptive production workflows, empowering manufacturers to dynamically respond to disruptions and optimize supply chain efficiency.

3.1 THE FEDERATED DATA MESH ARCHITECTURE FOR THE FDIF

A Federated Data Mesh Architecture [4] enhances the Federated Data Integration Framework by enabling decentralized data management while ensuring interoperability, data sovereignty, and real-time insights across a Smart Manufacturing Network. A Federated Data Mesh Architecture relies on Edge Computing — a decentralized computing paradigm where data processing happens closer to the source — at the factory floor, IoT devices, or edge servers — rather than relying solely on a centralized cloud or data centre. In an SMN, edge computing enables low-latency decision-making, bandwidth optimization, and operational resilience. Below are the key characteristics and their role in supporting FDIF.

3.1.1 Data Autonomy & Decentralized Ownership

What it Means: Each domain (factory, supplier, logistics provider) *retains full control* over its own data while making it accessible within the federated ecosystem.

How it Supports FDIF:

- Factories and suppliers own and manage their respective production, logistics, and sensor data without a centralized authority.
- Each participant exposes well-defined APIs instead of centralizing all data into a monolithic repository.
- The Digital Twins act as an orchestrator, ensuring compliance with interoperability and access policies.

3.1.2 Avoidance of Data Replication & Data Silos

What it Means: Data remains distributed within its original source systems rather than being duplicated or aggregated into a central database.

How it Supports FDIF:

- Prevents the creation of redundant copies, reducing storage costs and data inconsistencies.

- Ensures real-time access to the most up-to-date operational data instead of working with stale copies.
- Factories and suppliers maintain their own data governance policies while still participating in a shared data ecosystem.

3.1.3 Reducing Bandwidth & Avoiding Unnecessary Data Transfers

What it Means: Data is accessed **on-demand** using event-driven mechanisms rather than being moved in large batches.

- In a Federated Data Mesh, each domain (factories, suppliers, logistics) manages its own data streams and only exposes essential metadata instead of transmitting all raw data.
- Instead of bulk data transfers, only critical events, metadata, or summarized analytics flow through the FDIF.

How it Supports FDIF:

- MQTT-based messaging, OPC UA, and REST APIs allow streaming access to real-time machine status, logistics updates, and production insights.
- Instead of performing bulk ETL (Extract, Transform, Load) jobs, the IMC queries and integrates data dynamically when needed.
- Edge computing processes data locally, reducing the need to transmit large datasets across the SMN.

3.1.4 Domain-Specific Data Categorization

What it Means: In a Federated Data Mesh, each domain (e.g., factories, suppliers, logistics) defines its own data models based on industry-specific standards and ontologies while maintaining interoperability across the Federated Data Integration Framework.

How it Supports FDIF & Data Mesh Principles:

- Decentralized Data Ownership & Standardization
 - Each domain is responsible for structuring its data using manufacturing-specific ontologies such as ISO 15926 (industrial data exchange) and ISA-95 (manufacturing operations hierarchy).
 - Ensures consistent terminology and semantic meaning across distributed systems, preventing data silos.
- Function-Based Data Categorization for AI & Digital Twins
 - Data is categorized by business function (e.g., machine performance data, material availability, order fulfilment [5])
 - Enables Digital Twins and AI-driven analytics to efficiently extract insights by focusing on relevant data subsets rather than processing bulk datasets.
- Semantic Interoperability Across Domains
 - By using standardized ontologies, different stakeholders (factories, suppliers, logistics providers) can interpret and integrate shared data correctly without transformation overhead [6].
 - Supports event-driven data exchange where only relevant metadata and contextualized data are exposed rather than raw data dumps.

- Self-Serve Data Discovery & Federated Governance
 - Domains expose their categorized datasets as discoverable and queryable data products in the Federated Data Mesh rather than centralizing all data in one system.
 - Maintains data sovereignty while ensuring access via standardized APIs and metadata layers in FDIF.

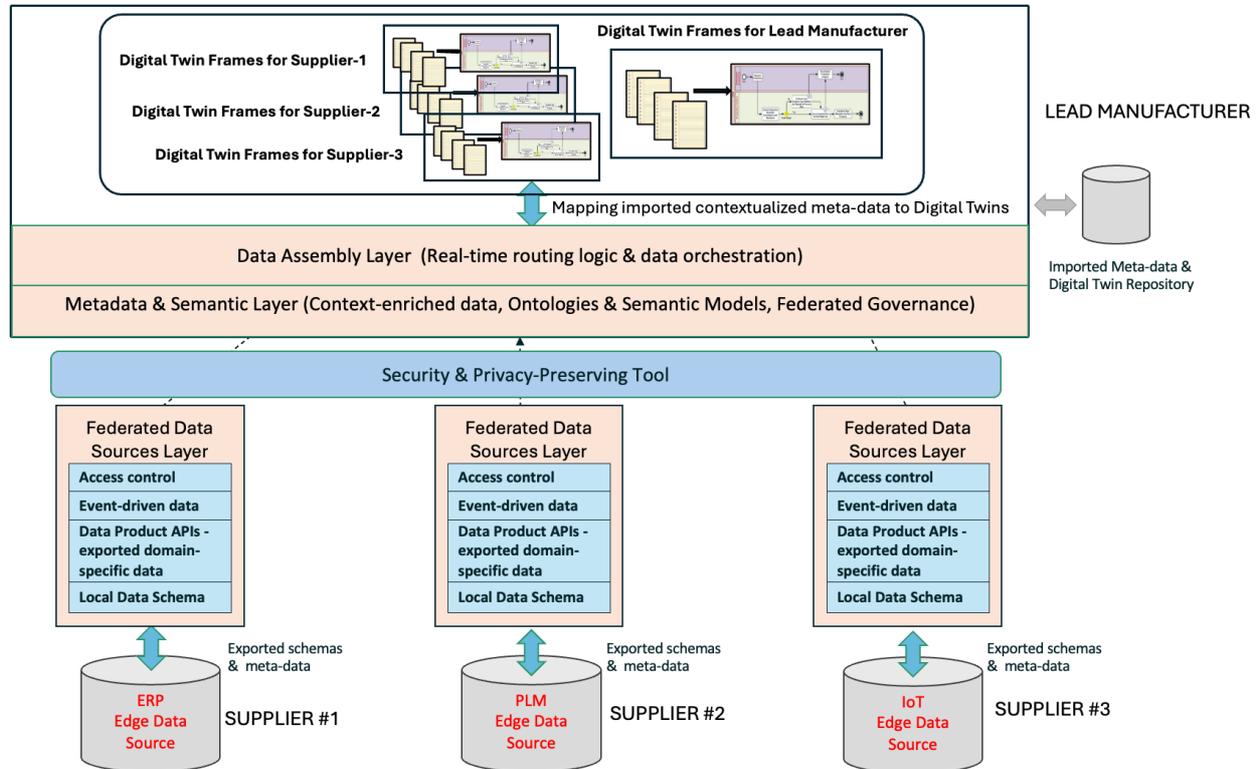


Figure 1 The Federated Data Integration Framework.

FIGURE 1 illustrates how a hybrid federated data mesh enables the integration of multiple supply chain datasets at the edge including supplier data, material flows, production output, and logistics flow, into a unified virtual federated database. The Federated Data Integration Framework in NARRATE is structured into three interconnected layers, -as shown in FIGURE 1- each playing a critical role in ingesting, standardizing, integrating and sourcing distributed supply chain and manufacturing data into the Blueprint Frames (Digital Twins). This layered approach ensures seamless data federation, real-time processing, and AI-driven decision support for resilient Smart Manufacturing Networks.

3.2 FEDERATED DATA SOURCES LAYER

The foundation of the FDIF, this layer consists of distributed, heterogeneous data sources that generate and publish real-time operational, production, and logistics data.

Characteristics:

- **Decentralized & Heterogeneous:** Data is sourced from independent, geographically distributed manufacturing sites, suppliers, logistics partners, and enterprise applications.
- **Multi-Modal Data Streams:** Includes structured, semi-structured, and unstructured data from IoT sensors, machines, RFID systems, ERP/MES platforms, and external supply chain partners.

- **Real-Time Event-Driven Data:** Uses lightweight protocols like MQTT and OPC UA to push live updates for instant decision-making.
- **Edge & Cloud Integration:** Data is processed both locally (edge devices) and in cloud environments, optimizing latency and resource utilization.

Key Functions:

- **Data Collection & Publication:** Sensors, IoT devices, and enterprise systems publish data using standardized protocols like MQTT, OPC UA, or REST APIs, see section-4.
- **Event-Based Streaming:** Supports real-time anomaly detection, machine performance monitoring, and production tracking.
- **Secure Data Transmission:** Implements encryption (TLS), access control, and federated authentication to ensure data integrity and sovereignty.

This layer feeds raw data into the Metadata & Semantic Layer, where it is structured and enriched before integration into the Blueprint Frames.

3.3 METADATA & SEMANTIC LAYER

This intermediate layer ensures data consistency, traceability, and semantic interoperability before it is incorporated into Digital Twins. It transforms raw data into context-aware, enriched datasets by applying metadata annotations, ontologies, and governance policies.

Characteristics:

- **Context-Enriched Data:** Raw data from federated sources is annotated with metadata to improve traceability, provenance, and interoperability.
- **Ontologies & Semantic Models:** Uses manufacturing-specific ontologies (e.g., ISO 15926, ISA-95) to create a shared vocabulary across the ecosystem.
- **Federated Governance & Policies:** Ensures role-based access control, compliance, and versioning of integrated data streams.

Key Functions:

- **Metadata Annotation:** Adds descriptive (e.g., machine type), structural (e.g., part relationships), and provenance (e.g., calibration history) metadata to each dataset. Ensures traceability and data lineage across distributed manufacturing nodes.
- **Semantic Interoperability:** Standardizes terms across factories (e.g., “temp” vs. “temperature”) to harmonize terminology across different systems ensuring cross-system compatibility.
- Ensures compatibility between independent supply chain and production data sources.

This layer acts as a bridge between raw data sources and the Blueprint Frames, ensuring that all integrated data is standardized, interpretable, and actionable.

3.4 DATA ASSEMBLY LAYER (SUPPORTS NARRATE DIGITAL TWINS & THE IMC)

The top layer of the FDIF assembles, synchronizes, and continuously sources data into Blueprint Frames, enabling real-time decision support for the Intelligent Manufacturing Custodian.

Characteristics:

- **Real-Time Data Aggregation:** Consolidates live data from multiple federated sources, while preserving their autonomy. Transforms federated data into high-level operational insights for continuous monitoring, simulation, and optimization.

Key Functions:

- **Digital Twin Synchronization:** Real-world manufacturing data is mirrored into Digital Twins for real-time tracking, anomaly detection, and performance analysis.
- **MaaS Enablement:** Supports on-demand production capabilities by allowing factories and suppliers to dynamically source data and allocate appropriate resources.
- **AI-Powered Decision Support for the IMC:** Provides support for important IMC functions through the digital twins, such as predictive maintenance, dynamic scheduling, and automated supply chain rerouting, intelligent logistics, and optimization of production workflows. This is explained briefly in the example below.

By interconnecting federated data sources, metadata standardization, and AI-driven decision support, the FDIF provides a comprehensive data backbone for NARRATE’s Blueprint Frames. This ensures that Digital Twins operate with real-time insights, empowering the Intelligent Manufacturing Custodian to drive resilient, adaptive, and self-optimizing Smart Manufacturing Networks.

Example Scenario: IMC-Driven Adaptive Manufacturing

1. Machine Failure Detected (Real-Time Event)

- A CNC machine in Factory A detects spindle overheating and publishes an MQTT message to the Resource Blueprint Frame. A CNC (Computer Numerical Control) machine is an automated manufacturing tool that uses computer programming to control machining operations with high precision. It processes various materials such as metal, wood, plastic, and composites by following a pre-programmed set of instructions (G-code) to cut, shape, drill, or engrave components.
- The MQTT Broker routes the message, ensuring it reaches relevant subscribers, including the NARRATE Digital Twins and Intelligent Manufacturing Custodian.

2. Metadata & Context Enrichment (Resource Frame Analysis)

- The IMC queries the blueprint’s Resource Frame, which integrates contextual data from the Metadata & Semantic Layer.
- Using historical calibration records, machine failure patterns, and operational status, the IMC assesses:
 - Time since last maintenance
 - Severity of overheating
 - Availability of alternate CNC machines
 - Impact on ongoing production schedules

3. Automated Decision via IMC (Dynamic Production Rerouting)

- The IMC synthesizes information from the blueprint frames (Process, Resource, and Product Frames) to make a real-time decision.
- It reroutes production tasks to a backup CNC machine in Factory B, which is:
 - Capable of the same machining operations
 - Available with minimal setup time

- Operationally compatible based on metadata constraints (e.g., tool availability, calibration status, material compatibility)
- This minimizes downtime and prevents production bottlenecks.

4. Supply Chain Notification (Proactive Resource Allocation)

- The IMC triggers an automated request via the SMN Blueprint Frame, notifying suppliers of the required replacement spindle parts.
- Just-in-Time (JIT) logistics are activated, ensuring that:
 - The new spindle is ordered, shipped, and installed before the backup machine also reaches a failure threshold.
 - Downtime remains minimal, with production resuming seamlessly.
- The IMC continuously monitors the production flow, ensuring SMN-wide resilience.

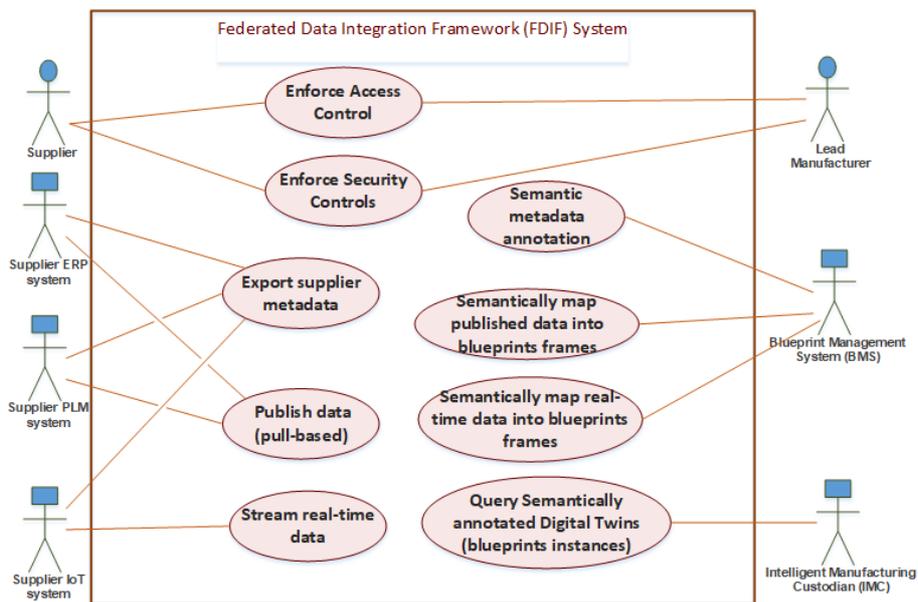


Figure 2 Use-case diagram for the IMC-driven adaptive manufacturing scenario.

FIGURE 2 presents a UML use-case diagram that illustrates the interactions among the lead manufacturer, suppliers, and supplier-supported systems—including ERP, PLM, and IoT platforms—alongside the Blueprint Management System and the Intelligent Manufacturing Custodian. It highlights how these entities collaborate within the IMC-driven adaptive scenario, showcasing the seamless exchange of data and coordinated decision-making necessary for dynamic and resilient manufacturing operations. The figure shows specific functionalities or services (e.g., "Semantic Mapping" or "Exporting MetaData" and relationships, i.e., connections between actors and use cases, showing interactions, (e.g., an actor initiating a use case).

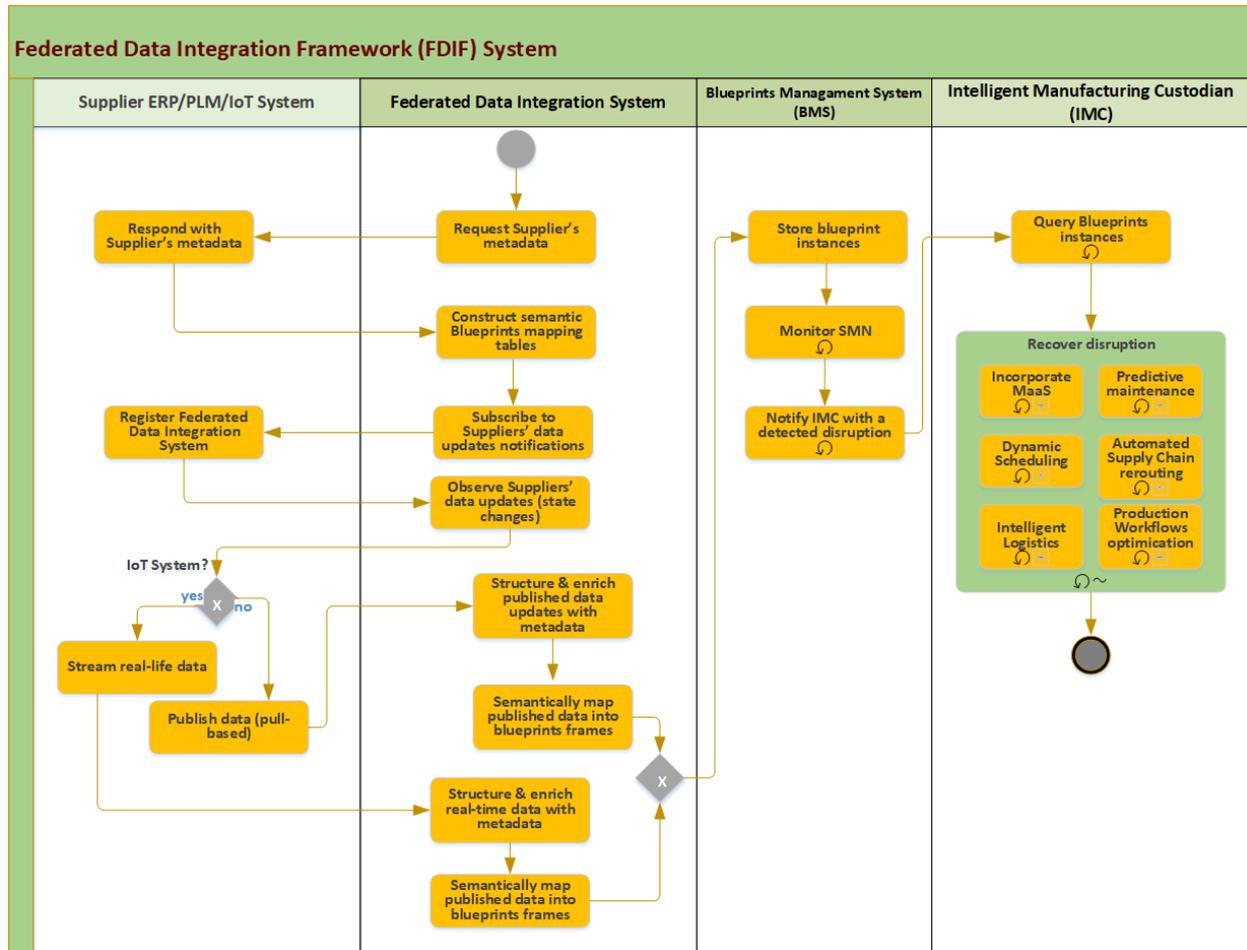


Figure 3 Sequence diagram for the IMC-driven adaptive manufacturing scenario.

The activity diagram in FIGURE 3 provides a detailed visualization of the interactions among existing data sources, the Federated Data Integration Framework, the Blueprint (i.e., Digital Twin) Management, and the Intelligent Manufacturing Custodian. It illustrates how the FDIF contextualises pre-existing data sources, and how data flows seamlessly across these components, enabling real-time interoperability, contextualized analytics, and optimized decision-making within a Smart Manufacturing Network.

3.5 DATA CATALOGUING AND DISCOVERY

The Federated Data Integration Framework enables efficient data cataloguing, discovery, and contextualization within a Smart Manufacturing Network. By leveraging metadata enrichment, ontologies, and federated governance, FDIF ensures that distributed data sources (factories, suppliers, logistics, IoT devices, etc) can be seamlessly discovered, queried, and integrated into NARRATE Blueprint Frames (Digital Twins) as shown in FIGURE 1.

3.5.1 Key Functions of Data Cataloguing & Discovery in FDIF

As shown in FIGURE 3 the FDIF contains a data repository that can be used for cataloguing and discovering all sorts of data including, metadata, contextualized data and digital twin frames. The Key Functions of Data Cataloguing & Discovery in FDIF include the following.

1. Metadata Enrichment & Contextualization

- **What It Does:**
 - FDIF automatically tags and annotates data sources with metadata (e.g., timestamps, machine IDs, process parameters).
 - Uses domain-specific ontologies (e.g., ISO 15926, ISA-95, MASON) to ensure semantic consistency across factories and supply chains.
 - Enriches data with contextual information (e.g., sensor readings linked to specific machine operations) [7].
- **Why It's Important:** Enables fast search and retrieval of relevant datasets & ensures data remains interpretable across different domains.

2. Data Discovery & Federated Search

- **What It Does:**
 - Provides self-serve access to distributed data sources through a federated search engine.
 - Users (e.g., AI models, engineers, and Digital Twin systems) can query the metadata layer instead of accessing raw data directly.
 - Supports event-driven discovery, allowing the IMC (Intelligent Manufacturing Custodian) to detect and pull only the necessary data for real-time decision-making.
- **Why It's Important:**
 - Eliminates bulk data transfers by exposing only relevant metadata.
 - Supports real-time, API-driven data access rather than static data dumps.

3. Exporting Contextualized Data to NARRATE Blueprint Repository

- **What It Does:**
 - Once relevant metadata and contextualized data are identified, FDIF exports structured datasets into NARRATE Blueprint Frames (Digital Twins), see FIGURE 1.
 - The blueprint repository stores high-level resource, process, and event frames, ensuring that AI models and Digital Twins operate on clean, contextualized data.
 - Uses standardized formats (JSON-LD, RDF) to ensure compatibility with federated systems.
- **Why It's Important:**
 - Ensures Digital Twins receive real-time, structured data for decision-making.
 - Prevents data duplication by using linked metadata instead of raw datasets.

3.6 FDIF'S ROLE IN NARRATE'S SMART MANUFACTURING NETWORK

The Table 2 below provides a summary of the FDIF's role in NARRATE's Smart Manufacturing Network. It shows that FDIF ensures that NARRATE's Intelligent Manufacturing Custodian has access to trusted, contextualized data, and provides live federated data streams to digital twins — providing the foundation enabling self-orchestrated, resilient Smart Manufacturing Networks.

TABLE 2 SUMMARY OF THE FDIF'S ROLE IN NARRATE'S SMART MANUFACTURING NETWORK.

FDIF Layer	Purpose	Key Technologies	Connection to NARRATE
Federated Data Sources	Captures real-time manufacturing, logistics, and supplier data.	IoT, MQTT, OPC UA, Edge Computing.	Provides live data streams to NARRATE Digital Twins.

Metadata & Semantic Layer	Enriches data with context, standardization, and governance.	RDF, SPARQL, JSON-LD, Manufacturing Ontologies.	Ensures data interoperability and traceability.
Data Assembly Layer	Aggregates, analyzes, and provides support for transforming data into decisions.	Digital Twins.	Powers NARRATE’s IMC for adaptive manufacturing.

4. STANDARDS FOR CONTEXTUALIZING DATA

When implementing large-scale industrial projects like NARRATE and FDIF, standardizing contextualized data is critical for achieving interoperability, data integration, and AI-driven analytics. However, not all standards will be relevant to every deployment. Some standards may overlap, may not be applicable for all pilot sectors. In contrast, others require alignment with specific industry supply chain models before adoption or compatibility mapping with MQTT and federated data sources—pilot testing is needed to select the most efficient approach. FDIF & NARRATE must prioritize standards that enhance interoperability without excessive data transformation. MQTT Metadata Extensions & Semantic Data Models will be key for real-time federated data exchange. Industry and technical partners will critically evaluate which standards best align with their pilot use cases before integrating them into the NARRATE Blueprint Repository.

4.1 STANDARDS FOR BOM AND MBOM

ISO 10303 (STEP): A standard for representing and exchanging product data, including BOM and MBOM. ISO 10303 (STEP—Standard for the Exchange of Product model data) is widely used in manufacturing industries to facilitate product data exchange between different CAD, CAM, and PLM systems. In the context of bespoke baby cot manufacturing, where Smart Manufacturing concepts are applied, STEP can play a crucial role in ensuring seamless data interoperability across the supply chain. Here are some practical use cases:

Use Case: Bill of Materials (BOM) & Supplier Integration

Scenario: A cot manufacturer orders materials (wood, fasteners, safety rails) from multiple suppliers. A standardized STEP AP239 file ensures each supplier receives the correct order specifications.

TABLE 3 STEP AP239 EXAMPLE FILE (EXTRACT)

```
# Product Life Cycle Support (PLCS) File
# Format: STEP AP239
# Baby Cot Smart Manufacturing Network

HEADER;
FILE_DESCRIPTION(('ISO 10303-239 PLCS'), '1;0');
FILE_NAME('Cot_BOM_AP239.stp', '2025-02-07', ('ERP System'), ('Supply Chain Manager'), 'Procurement');

DATA;
#10 = PRODUCT('Baby Cot - Custom Model 2025', '', '#15');
#15 = PRODUCT_CONFIGURATION('Cot_001_Custom', 'Parent cot structure', #20);
#20 = PRODUCT_VERSION('Cot_001_Custom_v1', 'Version 1', #30);
#30 = BILL_OF_MATERIALS('Complete Cot Assembly', (#40, #50, #60, #70));
#40 = BOM_ITEM('Side Panel', 'Wooden panel (Oak)', 'Supplier: WoodWorks Inc.', 'QTY: 2', 'Tolerance: 0.05 mm');
#50 = BOM_ITEM('Front Panel', 'Wooden panel (Oak)', 'Supplier: WoodWorks Inc.', 'QTY: 1', 'Tolerance: 0.05 mm');
#60 = BOM_ITEM('Fasteners', 'Screws (Stainless Steel)', 'Supplier: HardwarePro', 'QTY: 12', 'Thread Size: M6');
#70 = BOM_ITEM('Paint', 'Non-toxic white paint', 'Supplier: EcoCoatings', 'Volume: 500ml');
#80 = SUPPLY_CHAIN_EVENT('Wood delivery', 'Expected: 2025-02-10', 'Status: In Transit');
#90 = SUPPLY_CHAIN_EVENT('Paint delivery', 'Expected: 2025-02-12', 'Status: Confirmed');
#100 = QUALITY_COMPLIANCE_CHECK('ASTM F1169', 'Pending Inspection', 'QC Approval Needed');
```

ENDSEC;
 END-ISO-10303;

Explanation:

- Product Definition & Versioning: Tracks different baby cot models and revisions.
- Bill of Materials (BOM): Lists all components with supplier details, material specifications, and quantities.
- Supply Chain Events: Tracks material deliveries and expected dates.
- Quality Compliance Check: Ensures materials meet safety standards (ASTM F1169) before use.

4.2 STANDARDS FOR HISTORICAL MANUFACTURING AND PROCESS DATA

OPC UA (Unified Architecture): OPC UA (Unified Architecture) is a widely adopted industrial communication protocol for secure, reliable, and platform-independent data exchange in industrial automation. It provides a standardized framework for contextualizing process data, historical data, alarms, and events, making it essential for Smart Manufacturing and Industrial IoT (IIoT) applications. OPC UA represents data using a hierarchical, object-oriented model. Key elements include:

- Objects: Represent machines, sensors, materials, or processes.
- Variables: Store real-time data (e.g., temperature, speed, quality metrics).
- Methods: Define actions (e.g., start/stop machine, adjust settings).
- Events & Alarms: Trigger notifications when conditions change.

Key Features of OPC UA

- *Interoperability* – Enables seamless communication between devices, machines, and enterprise systems across different vendors and platforms.
- *Scalability* – Supports everything from small embedded devices to large cloud-based architectures.
- *Security* – Includes encryption, authentication, and access control mechanisms.
- *Contextualization* – Provides semantic data models for understanding real-time and historical process data.
- *Extensibility* – Allows the integration of custom information models for specific industries.

Role of OPC UA in Smart Manufacturing & Digital Twins

- Real-time Data Exchange: Enables factories to collect and share live sensor and machine data.
- Historical Data Analysis: Helps in predictive maintenance and quality control by storing and analysing past performance metrics.
- Integration with Digital Twins: Feeds accurate, real-time production data into Digital Twin models for simulation, optimization, and decision-making.

Supply Chain Coordination: Enhances visibility by providing structured, standardized data across different stakeholders in a Smart Manufacturing Network.

In the context of *bespoke baby cot manufacturing*, OPC UA can be used to:

- Connect and monitor CNC machines, robotic arms, and finishing stations.
- Gather quality control data for customized orders.

- o Enable suppliers and manufacturers to share real-time production data securely.

OPC UA Data Model for the CNC Cutting Machine: CNC cutting is the first step in baby cot production. The CNC cutting machine is a key part of the bespoke baby cot manufacturing process, responsible for precisely cutting wooden panels according to custom designs. To ensure high-quality cuts, real-time monitoring and historical analysis of cutting parameters are essential. The OPC UA data model for the CNC machine provides a structured representation of its operational state, process variables, alarms, and control methods.

Example: Structure of the OPC UA Object for CNC_Machine_01

The following is an example of how the OPC UA can model a CNC machine. A CNC (Computer Numerical Control) machine is an automated manufacturing tool that uses pre-programmed software and code to control the movement of machinery and tools. CNC machines can perform complex operations such as cutting, drilling, milling, and turning with high precision, repeatability, and efficiency. The CNC_Machine_01 object consists of:

- o Machine status variables (e.g., operational state, temperature, tool wear)
- o Real-time process parameters (e.g., spindle speed, feed rate)
- o Alarms & events (e.g., tool breakage, feed rate drops)
- o Methods for machine control (e.g., start, stop, change tool)

TABLE 4 CNC_MACHINE_01 (UPC UA OBJECT)

└ MachineStatus (Variable, Enum: {Idle, Running, Error, Maintenance})
└ SpindleSpeed (Variable, Float, RPM)
└ FeedRate (Variable, Float, mm/min)
└ ToolWearLevel (Variable, Float, %)
└ MaterialType (Variable, String)
└ CuttingDepth (Variable, Float, mm)
└ VibrationLevel (Variable, Float, mm/s ²)
└ Temperature (Variable, Float, °C)
└ StartCutting() (Method)
└ StopCutting() (Method)
└ ChangeTool(ToolID) (Method)
└ Alarm_ToolBreakage (Event)
└ Alarm_FeedRateDrop (Event)
└ Alarm_VibrationTooHigh (Event)
└ Alarm_CoolantLow (Event)
└ HistoricalData (Object)
└ PastCuttingSpeed (Array, Float, RPM)
└ PastFeedRates (Array, Float, mm/min)
└ ToolUsageHistory (Array, ToolID, TimeUsed)
└ MaintenanceLogs (Array, String, DateTime)

A Detailed Explanation of Key Data Elements follows.

Machine Status (MachineStatus)

This variable represents the overall state of the CNC machine. It helps determine if the machine is operating correctly or needs maintenance. Possible values:

- o Idle – Machine is powered on but not cutting.
- o Running – Machine is actively cutting a workpiece.
- o Error – A problem has occurred, requiring human intervention.

- Maintenance – The machine is undergoing maintenance or repairs.

Real-Time Process Variables

- SpindleSpeed (RPM) – The rotational speed of the cutting tool, measured in revolutions per minute (RPM). Possible values:
 - Too high → May cause excessive tool wear.
 - Too low → May lead to inefficient cutting.

The Resources blueprint (see deliverable D3.2) monitors this and suggests adjustments based on the material being cut.

- FeedRate (mm/min) – The speed at which the cutting tool moves through the material.
 - If the feed rate is too fast, it may cause rough edges.
 - If the feed rate is too slow cutting time and costs increase.
- ToolWearLevel (%) – Indicates how much the cutting tool has worn down.
 - When ToolWearLevel > 80%, the Alarm_ToolBreakage event is triggered.
- MaterialType (String) – Specifies the type of wood being processed (e.g., Oak, Maple, Birch).
 - Different materials require different spindle speeds and feed rates, and OPC UA allows dynamic adjustment.
- CuttingDepth (mm) – The depth of the cut into the wood.
 - If cutting depth exceeds the limit for a given tool, an alarm can be triggered.
- VibrationLevel (mm/s²) – Measures machine vibrations.
 - High vibration levels can indicate tool imbalance or spindle misalignment.
 - If VibrationLevel > threshold, the Alarm_VibrationTooHigh event is triggered.
- Temperature (°C) – Monitors the temperature of the spindle and tool.
 - If the temperature exceeds safe limits, OPC UA alerts operators to prevent tool overheating.

4.3 INTEROPERABILITY STANDARDS

4.3.1 ISO 15926: The Standard for Interoperability in Industrial Data

ISO 15926 is an international standard designed for the integration, sharing, exchange, and lifecycle management of industrial plant data. It is widely used in process industries such as oil and gas, petrochemicals, power plants, and large-scale manufacturing where long-term data interoperability is essential.

Key Characteristics of ISO 15926

Designed for Data Interoperability

- Ensures that data from different systems, vendors, and formats can be seamlessly integrated and understood across the entire lifecycle of industrial assets.
- Supports federated data exchange without requiring a single, centralized database.

Semantic Data Modelling & Ontologies

- Uses an ontology-based approach to define the meaning of data so that it remains understandable even when systems or formats change.
- Provides a common reference data library (RDL) with standardized terms, ensuring that different organizations interpret data consistently.

Lifecycle Integration

- Supports end-to-end data management, from design and engineering to operation and maintenance of industrial plants and manufacturing systems.
- Example: Data from an equipment supplier can be structured so that it remains useful for operation, maintenance.

Supports Linked Data & Knowledge Graphs

- Uses a linked-data approach, allowing distributed databases to refer to each other using semantic web technologies (RDF, OWL, SPARQL) [8].
- Enables federated queries across multiple industrial data sources without data duplication.

Avoids Vendor Lock-in:

- By standardizing how data is structured and exchanged, ISO 15926 ensures that industries are not locked into proprietary software systems.
- This allows interoperability between legacy systems and new digital solutions such as Digital Twins, Smart Manufacturing Networks (SMNs), and AI-driven analytics.

How ISO 15926 Supports Smart Manufacturing & FDIF

Seamless Data Exchange: ISO 15926 can be used in the Federated Data Integration Framework to enable secure and standardized data flow between factories, suppliers, and logistics providers.

Metadata & Context Enrichment: The ontology-driven structure of ISO 15926 helps enhance metadata within Digital Twins (NARRATE Blueprint Frames), ensuring that all data has proper context and meaning.

Event-Driven Interoperability: Using ISO 15926 alongside MQTT, OPC UA, and REST APIs, manufacturing systems can dynamically integrate with real-time machine data, supplier information, and logistics updates.

Long-Term Industrial Data Sustainability: Ensures that historical production data remains accessible and understandable decades later, supporting predictive analytics and maintenance in Smart Manufacturing Networks.

4.3.2 MTConnect

MTConnect is an open, royalty-free interoperability standard that enables real-time data exchange between manufacturing equipment. It provides a common vocabulary and standardized data format (XML/JSON) for machine tool data, allowing seamless communication between devices from different manufacturers.

Key Features of MTConnect:

- Standardized Data Model – Defines a structured representation of machine status, tool conditions, and operational parameters.
- Hierarchical Data Representation – Organizes data into logical categories (e.g., devices, components, and streams).
- Real-Time Data Collection – Captures and streams machine states, alarms, temperatures, spindle speeds, etc.
- Lightweight & Scalable – Uses RESTful APIs over HTTP for efficient data transfer.

How MTConnect Can Be Used in the Federated Data Integration Framework

A Federated Data Integration Framework (FDIF) in Smart Manufacturing Networks (SMNs) aims to integrate heterogeneous data sources without centralizing them. MTConnect plays a crucial role in this by enabling standardized machine data exchange across distributed factories.

Integration Steps for MTConnect in FDIF

1. Standardizing Data from Heterogeneous Machines
 - MTConnect adapters extract raw data from machines and agents convert it into a standardized format (XML/JSON).
 - This ensures that all machine tools in a Federated Smart Manufacturing Network provide uniform data, enabling easy aggregation and analysis.
 - Example: Different CNC machines (Siemens, Fanuc) output varying proprietary data formats. MTConnect normalizes these into a common format.
2. Federated Data Querying and Streaming
 - MTConnect streams real-time data to a federated data platform where different nodes (factories, suppliers) can subscribe to relevant information.
 - Example Use Case: A manufacturer integrates multiple suppliers’ production status into a Resilience Blueprint for predictive scheduling.
3. Enhancing Digital Twins with Real-Time Machine Data
 - MTConnect feeds real-time machine states into Digital Twin models, enabling accurate virtual simulation of production processes.
 - Example Use Case: In bespoke cot manufacturing, a digital twin uses MTConnect data to simulate assembly conditions across distributed manufacturing nodes.
4. Improving Predictive Maintenance and Resilience
 - MTConnect provides a continuous stream of sensor data, which is used for anomaly detection and predictive maintenance within the FDIF.
 - Example Use Case: A Resilience Blueprint adjusts supplier allocations based on real-time machine health data from different factories.
5. Integrating with Other Data Standards
 - MTConnect works alongside OPC UA, ISA-95, and IEC 62264 to provide a comprehensive federated data exchange.
 - It bridges the gap between low-level machine data and high-level enterprise applications (ERP, MES, AI analytics).

TABLE 5 POTENTIAL IMPLEMENTATION TOOLS FOR MTCONNECT IN FDIF

Component	Purpose	Example Potential Tools
MTConnect Adapter	Extracts raw machine data	MTConnect SDK, Custom Adapters
MTConnect Agent	Converts raw data to XML/JSON	MTConnect Agent (Open-Source)
Data Exchange Middleware	Enables real-time streaming	Apache NiFi, MQTT Brokers

4.4. STANDARDS FOR IOT SENSOR DATA

4.4.1 SensorML (OGC)

Sensor ML is a standard for encoding metadata for IoT sensors, including location, accuracy, and data format. Sensor Model Language (SensorML) is an Open Geospatial Consortium (OGC) standard that encodes metadata for IoT sensors. In the Smart Manufacturing Network SensorML can describe:

- Sensor metadata – Sensor type, manufacturer, accuracy, and operating range.
- Sensor observations – Real-time measurements such as temperature, vibration, and humidity.
- Sensor processes – How sensor data is collected, transformed, and used for decision-making.

By using SensorML, the Intelligent Manufacturing Custodian and Digital Twin can access standardized, machine-readable descriptions of all IoT sensors involved in production. This allows for seamless integration, interoperability, and automation.

Example: SensorML Encoding for a CNC Machine Vibration Sensor

Scenario: A vibration sensor is attached to the CNC cutting machine to monitor abnormal vibrations that could indicate tool wear, misalignment, or material inconsistencies.

A SensorML document as shown in Table 6 describes:

- Sensor ID and manufacturer details
- Measurement units and accuracy
- Physical properties (location, power source, connectivity)
- Observation process (how data is recorded and processed)

TABLE 6 A CNC MACHINE VIBRATION SENSOR IN SENSOR-ML.

```
<SensorML xmlns="http://www.opengis.net/sensorML/1.0">
  <member>
    <Sensor>
      <identification>
        <identifier name="SensorID">VIBR001</identifier>
        <identifier name="Manufacturer">Acme Sensors Ltd.</identifier>
        <identifier name="Model">VIBRO-1000</identifier>
      </identification>
      <classification>
        <type>Vibration Sensor</type>
        <applicationDomain>Manufacturing</applicationDomain>
      </classification>
      <capabilities>
        <capability name="MeasurementRange">0 - 100 mm/s²</capability>
        <capability name="Accuracy">±0.1 mm/s²</capability>
        <capability name="SamplingRate">100 Hz</capability>
        <capability name="Connectivity">OPC UA</capability>
      </capabilities>
      <location>
        <srsName="EPSG:4326">
          <coordinates>52.1234, 5.6789</coordinates>
        </srsName>
      </location>
      <outputs>
```

```

<Output name="VibrationLevel">
  <Quantity>
    <uom code="mm/s2" />
  </Quantity>
</Output>
</outputs>
<processes>
  <Process name="FilterNoise">
    <description>Applies noise reduction to raw vibration data.</description>
  </Process>
  <Process name="AnalyzeThreshold">
    <description>Triggers an alarm if vibration exceeds 50 mm/s2.</description>
  </Process>
</processes>
</Sensor>
</member>
</SensorML>

```

Explanation of Key SensorML Elements in the SensorML Encoding for a CNC Machine Vibration Sensor Example:

- Identification Section
 - Defines the unique sensor ID, manufacturer, and model number.
 - Allows sensor traceability across the Smart Manufacturing Network.
- Classification Section
 - Specifies that this is a vibration sensor used in manufacturing.
- Capabilities Section
 - Describes the operating range (0 - 100 mm/s²), accuracy (± 0.1 mm/s²), & sampling rate (100 Hz).
 - Indicates the sensor supports OPC UA connectivity, ensuring compatibility with the IMC.
- Location Section
 - Specifies the geographic position of the sensor, which is useful for large-scale factories with multiple CNC machines.
- Outputs Section
 - Defines the measured variable (vibration level) and its unit (mm/s²).
- Processes Section
 - FilterNoise: Applies noise reduction to prevent false alarms.
 - AnalyzeThreshold: If vibration exceeds 50 mm/s², the system triggers an alert to the Digital Twin.

4.4.2 MQTT Metadata Extensions

MQTT Metadata Extensions is lightweight metadata in MQTT-based IoT systems. In an MQTT-based IoT system, lightweight metadata extensions enhance data context, traceability, and interoperability without increasing payload size significantly. In an SMN (Smart Manufacturing Network) aligned with the NARRATE blueprint frames, metadata allows seamless integration, decision-making, and resilience across distributed factories, suppliers, and logistics partners.

How MQTT Metadata Works in an SMN (Using NARRATE Blueprint Frames)

MQTT metadata extends standard messages by embedding additional contextual information at the protocol level. This allows Digital Twins, AI models, and supply chain partners to make more informed, autonomous decisions. Metadata categories, their purpose, and examples are shown in Table 7.

TABLE 7 HOW MQTT METADATA WORKS IN AN SMN (USING NARRATE BLUEPRINT FRAMES).

Metadata Category	Purpose in SMN	Example in MQTT Message
Descriptive Metadata	Identifies & describes machine-generated data.	Machine ID, sensor type, timestamp.
Structural Metadata	Defines relationships between manufacturing entities.	Links between BOM components, supplier sources.
Administrative Metadata	Controls access, ownership, and data lifecycle.	User authentication, versioning, role-based access.
Provenance Metadata	Tracks the origin and processing history of data.	Sensor calibration, past failures, process logs.

TABLE 8 EXAMPLE MQTT PAYLOAD WITH METADATA EXTENSIONS.

```

{"topic": "factory1/CNC1/temperature",
  "message": {
    "value": 75.4,
    "units": "Celsius",
    "timestamp": "2025-02-08T12:30:00Z",
    "metadata": {
      "source": "CNC1",
      "machine_type": "5-axis mill",
      "operator": "autonomous",
      "QoS": 2,
      "provenance": "sensor-calibrated-2025-02-01"
    }
  }
}
    
```

Table 8 illustrates an example MQTT payload with metadata extensions. In the MQTT example:

- Standard MQTT clients receive only the payload (value: 75.4°C).
- Advanced SMN systems (IMC, Digital Twins) parse metadata for additional context.

MQTT metadata plays a critical role in enabling the NARRATE blueprint frames, ensuring that data from various data sources is structured, traceable, and actionable this is shown in Table 9.

TABLE 9 BLUEPRINT FUNCTIONS & MQTT METADATA ROLE.

NARRATE Blueprint Function	MQTT Metadata Role
Real-Time Data Aggregation	Metadata-rich MQTT topics allow seamless integration of machine/sensor data into Digital Twins.
Predictive Analytics & Resilience	Provenance metadata ensures AI models can track sensor calibration, machine history, and predictive maintenance logs.
Supply Chain Visibility	Structural metadata enables the IMC to track raw material flow, order fulfillment, and disruptions.
Manufacturing-as-a-Service (MaaS)	Lightweight metadata enables interoperability between federated factories and suppliers.

What is an MQTT Broker?

An MQTT broker is the central server in an MQTT (Message Queuing Telemetry Transport) communication system. It facilitates message exchange between devices (clients) in a publish-subscribe model, ensuring reliable and efficient data transfer in IoT, IIoT (Industrial IoT), and Smart Manufacturing applications.

The following describe how an MQTT Broker Works.

1. Clients (Publishers) Send Data
 - o Devices (e.g., IoT sensors, machines, CNC tools) publish messages to a specific topic (e.g., factory/machine1/temperature).
2. Broker Manages Communication
 - o The MQTT broker receives the message and routes it to the right subscribers based on topic filtering.
3. Clients (Subscribers) Receive Data
 - o Devices or applications that subscribe to the topic receive the message in real time.

Example:

- o A CNC machine publishes temperature sensor data to the broker.
- o A predictive maintenance system subscribes to the topic and receives updates in real-time.

FDIF relies on federated, loosely coupled data sources, making MQTT an ideal messaging backbone. The integration occurs at three key levels as shown in Table 10.

TABLE 10 FDIF LAYERS & THE ROLE OF MQTT.

FDIF Component	MQTT Role
Federated Data Sources	MQTT enables real-time data publishing from distributed sensors, CNC machines, robots, and suppliers.
Metadata & Semantic Layer	MQTT metadata (structured JSON/XML) enriches data streams with context, provenance, and QoS information.
Data Integration Layer (IMC in NARRATE)	MQTT provides an event-driven backbone for streaming data into Digital Twins, AI-driven analytics, and ERP/MES systems.

Example FDIF Architecture with MQTT

- o Factories & Suppliers publish real-time production, inventory, and logistics data to the MQTT broker.
- o The MQTT broker ensures data interoperability, filtering, and security before distributing it to subscribers (analytics, AI models, Digital Twins).
- o Decision-making systems (IMC in NARRATE) consume MQTT streams to detect disruptions and optimize production planning.

4.4.3 Key Takeways

Table 11 serves as a comparative table of data contextualization standards for FDIF.

TABLE 11 COMPARATIVE ANALYSIS OF DATA CONTEXTUALIZATION STANDARDS.

Category	Standard	Purpose	Relevance to FDIF	Considerations
BOM & MBOM (Product & Manufacturing Data)	ISO 10303 (STEP)	Standardized exchange of BOM and CAD data.	Enables structured product data for NARRATE blueprints.	Needs integration with ERP/MES for real-time

				manufacturing adaptability.
	ISO 23386 & ISO 23387	Metadata structures for BOM in manufacturing & construction.	Supports metadata-driven product categorization for supply chain integration.	May require adaptation for industry-specific needs.
Historical Manufacturing & Process Data	ISA-95 (Enterprise-Control Integration)	Defines hierarchical models for factories & processes.	Essential for mapping factory operations to federated data models.	Needs alignment with IoT & Digital Twin implementations.
	OPC UA	Secure, semantic industrial data exchange protocol.	Provides interoperable M2M communication within FDIF.	Requires companion specs for industry-specific applications.
Interoperability & Data Federation	ISO 15926	Semantic data integration for industrial ecosystems.	Supports cross-domain data interoperability.	Most relevant to process industries; needs assessment for discrete manufacturing.
	B2MML (ISA-95 XML Schema)	Enables seamless ERP-MES-PLM integration.	Helps unify IT & OT layers in FDIF-based SMNs.	XML-based; may need adaptation for modern data streaming architectures.
IoT & Sensor Data	MTConnect	Standardized machine tool data exchange.	Provides real-time shop-floor insights for Digital Twins.	Limited outside CNC machines; integration with MQTT or OPC UA needed.
	SensorML (OGC)	Metadata model for IoT sensor data.	Enhances sensor data contextualization for Digital Twins.	May require custom extensions for manufacturing contexts.
	OPC UA for IIoT	Secure, standardized M2M messaging for IIoT devices.	Supports real-time factory monitoring via FDIF.	Needs mapping to MQTT-based federated data layers.
Real-Time Data Streaming & Metadata Management	MQTT Metadata Extensions	Adds lightweight metadata to MQTT payloads.	Enables real-time data contextualization at the Edge in FDIF.	Still evolving—requires pilot implementations.

- Multiple standards exist, but not all will be used—critical evaluation is needed. FDIF and NARRATE require standards that support interoperability, metadata management, and real-time data streaming.
- ISA-95, OPC UA, and MQTT Metadata Extensions are highly relevant for federated smart manufacturing networks.
- Pilot testing by industry & technical partners will determine which standards are best suited for NARRATE.

4.4.4 Industrial Data Standardization: A Practical Guide for Pilots

Industrial pilot partners and technical experts in NARRATE will take several key steps to ensure their pre-existing data sources and systems re contextualised and align with industry-wide standards such as OPC-UA and MTConnect. This process will be carried out with close collaboration from technical experts and will involve in-depth discussions to identify the most pertinent and useful standards for the specific needs of the pilots and the overall project. These steps include:

1. Assess and Map Existing Systems

- Identify all data sources, including ERP, PLM, MES, IoT, and sensor networks.
- Analyze their data formats, protocols, and communication methods to determine compatibility gaps with OPC-UA and MTConnect.

2. Implement Standardized Data Models

- Convert proprietary data structures into standardized ontologies to ensure interoperability and adopt semantic models aligned with OPC-UA and MTConnect.

3. Deploy Data Gateways and Adapters

- Use protocol adapters to translate existing system outputs into OPC-UA or MTConnect-compliant formats.

4. Integrate Secure, Scalable APIs

- Develop metadata-driven APIs for standardized data exchange.
- Ensure secure, real-time connectivity between local systems and the broader Smart Manufacturing Network.

5. Validate and Iterate in Real-World Pilots

- Conduct pilot testing with OPC-UA and MTConnect implementations.
- Collaborate with standardization bodies and refine data models based on feedback.

By following these steps, industrial partners can *seamlessly integrate* their existing infrastructure into a federated, standards-compliant Smart Manufacturing Network, enhancing interoperability, automation, and resilience within NARRATE.

5. PUTTING THINGS TOGETHER

This section presents a comprehensive example that demonstrates how the three FDIF layers, as shown in FIGURE 1, interact in the context of CNC machine data integration.

5.1 EXAMPLE: CNC-DRIVEN DATA FLOW WITHIN THE FDIF

The FDIF plays a critical role in integrating CNC machine data into the NARRATE Smart Manufacturing Network. CNC machines generate vast amounts of real-time production data, which, when effectively utilized, can enhance operational efficiency, predictive maintenance, and automated decision-making. The FDIF ensures that this data is collected, standardized, and integrated into Blueprint Frames—the foundational digital twin representations in NARRATE.

By structuring CNC machine data through the three-layered FDIF approach, manufacturers gain a unified and intelligent data flow that supports real-time monitoring, adaptive manufacturing, and process optimization. Each layer of the FDIF contributes uniquely to this goal, ensuring that raw CNC machine data is transformed into actionable insights for supply chain orchestration, predictive maintenance, and dynamic production scheduling.

5.1.1. Three Layers of the FDIF for CNC Machine Data Integration

In the following we describe how the three FDIF layers in FIGURE 1 interact in the case of a comprehensive example involving CRC data integration.

1. Federated Data Sources Layer (CNC Machine-Level)

The Federated Data Sources Layer forms the foundation of the FDIF by capturing real-time data from CNC machines operating across various geographically distributed production facilities. This layer is responsible for collecting, publishing, and securely transmitting CNC data for further processing.

Key Functions & Characteristics

- Real-Time CNC Data Acquisition:
 - Captures critical machining parameters such as spindle speed, tool wear, tool paths, material cutting rates, and power consumption.
 - Monitors real-time machine health indicators like temperature, vibrations, and lubrication levels.
- IoT & Edge Computing for Data Transmission:
 - CNC data is streamed via lightweight IoT protocols such as MQTT (Message Queuing Telemetry Transport) and OPC UA (Open Platform Communications Unified Architecture) to ensure low-latency communication.
 - Edge devices preprocess CNC data, reducing bandwidth requirements and enabling real-time local decision-making.
- Event-Based Data Streaming & Anomaly Detection:
 - Automated alerts trigger when predefined thresholds for tool wear, temperature deviations, or mechanical inconsistencies are detected.

- Supports proactive maintenance strategies by detecting patterns that indicate potential machine failures before they occur.
- Integration with Enterprise Systems:
 - CNC data is fed into ERP (Enterprise Resource Planning) and MES (Manufacturing Execution Systems) to align real-time machining conditions with broader business and supply chain decisions.

2. Metadata & Semantic Layer (Standardization & Governance)

The Metadata & Semantic Layer acts as the governance and standardization hub for CNC machine data within the FDIF. This layer ensures that raw CNC data from multiple factories and machines conforms to common data models, ontologies, and industrial standards, making it easier to integrate, interpret, and utilize within NARRATE's digital twin ecosystem.

Key Functions & Characteristics

- Metadata Annotation for CNC Data:
 - Adds descriptive metadata (e.g., machine type, manufacturing process step), structural metadata (e.g., relationships between machine components), and provenance metadata (e.g., maintenance history, sensor calibration details).
 - Ensures traceability and provenance tracking, allowing manufacturers to maintain detailed records of machine performance.
- Ontology-Based Data Standardization:
 - Utilizes industry-recognized ontologies (e.g., ISA-95, ISO 15926) to harmonize data from different CNC machines and manufacturing plants.
 - Standardizes machine parameters (e.g., ensuring that "cutting speed" and "spindle speed" have consistent definitions across datasets).
- Contextual Integration with External Factors:
 - Incorporates real-time contextual information such as supply chain disruptions, material availability fluctuations, last-minute design changes, and production rescheduling.
 - Enables CNC machines to dynamically adjust tool paths and machining strategies based on external conditions.
- Federated Governance & Security Policies:
 - Implements role-based access control (RBAC) and data versioning policies to ensure data integrity and security.
 - Ensures compliance with data sovereignty regulations, particularly in cross-border manufacturing operations.

3. Data Assembly Layer (Digital Twin & IMC Integration)

The Data Assembly Layer is responsible for transforming CNC machine data into actionable insights. It aggregates, analyses, and synchronizes CNC data into Digital Twins (Blueprint Frames) while also feeding decision-support intelligence into the Intelligent Manufacturing Custodian. This enables adaptive, predictive, and resilient manufacturing workflows.

Key Functions & Characteristics

- Real-Time Data Aggregation & Digital Twin Synchronization:
 - Continuously mirrors CNC machine operations within Blueprint Frames, ensuring real-time tracking of manufacturing performance.

- Provides a unified representation of machine health, process efficiency, and production status.
- AI-Powered Decision Support for Adaptive Machining:
 - AI-driven analytics optimize machining parameters by considering historical performance data, real-time sensor feedback, and predictive maintenance insights.
 - CNC machines can automatically adapt cutting strategies to maximize tool life and minimize defects.
- MaaS Enablement:
 - Allows CNC machines to operate in a flexible, demand-driven environment, where production jobs are dynamically allocated based on machine availability, material constraints, and supply chain conditions.
- Intelligent Manufacturing Custodian (IMC) Integration:
 - CNC machine data feeds directly into the IMC, enabling:
 - Predictive maintenance alerts (reducing downtime and extending machine lifespan)
 - Dynamic production rescheduling (optimizing machining sequences based on real-time factory conditions)
 - Automated supply chain adjustments (intelligent rerouting of manufacturing tasks when disruptions occur)
 - Ensures CNC operations align with the broader NARRATE Smart Manufacturing Network.

6. CONCLUSIONS & FUTURE WORK

6.1 CONCLUSIONS

The Federated Data Integration Framework (FDIF) represents a significant advancement in enabling real-time, decentralized, and contextualized data management within Smart Manufacturing Networks. By leveraging a hybrid federated data mesh, FDIF ensures seamless data interoperability while preserving the autonomy of distributed data sources. This approach enhances data availability, analytics quality, and system resilience, making it well-suited for dynamic and adaptive manufacturing environments.

Key contributions of the FDIF framework include:

- **Decentralized Data Interoperability:** Facilitates *seamless data exchange* across manufacturing systems without centralized dependency.
- **Contextualized Analytics:** Enriches raw data with *real-time contextual insights*, improving decision-making accuracy.
- **Lightweight & Scalable Integration:** Employs *metadata-driven APIs* and *edge computing* to optimize performance and reduce unnecessary data transfers.
- **Digital Twin Synchronization:** Supports predictive analytics and proactive decision-making, enhancing supply chain agility and manufacturing efficiency.

By transitioning from traditional monolithic data architectures to a federated data mesh, manufacturers can improve operational intelligence, strengthen resilience, and accelerate digital transformation. The FDIF approach lays the groundwork for self-orchestrated, intelligent manufacturing ecosystems, where decentralized data ecosystems drive agility, efficiency, and innovation.

6.2 FUTURE WORK

While the FDIF framework demonstrates significant potential, its full-scale adoption and optimization require further research and industrial validation. Key areas for future development include:

1. Standardization & Interoperability Enhancements

- Further refinement of semantic models and ontologies (e.g., CDIF, ISO 15926, ISA-95) to ensure wider interoperability across different industrial domains.
- Development of automated ontology mapping techniques to integrate legacy systems with modern FDIF-based architectures.

2. Advanced AI-Driven Contextualization Services

- Implementation of AI-powered data enrichment techniques to contextualize raw manufacturing data based on real-time operational conditions.

3. Real-World Validation & Pilot Deployments

- Deployment of FDIF in diverse manufacturing environments (e.g., WP-5 pilots) to assess scalability, adaptability, and robustness.

4. Digital Twin & Intelligent Manufacturing Custodian Integration

- Enhancing digital twin synchronization with real-time AI-driven decision support for predictive maintenance, workflow optimization, and supply chain resilience.

5. Security & Data Sovereignty

- Strengthening access control mechanisms to ensure data sovereignty, privacy, and regulatory compliance.
- Implementation of decentralized identity management and zero-trust security architectures to protect sensitive manufacturing data.

By addressing these future challenges, FDIF will continue to evolve as a foundational technology for next-generation smart manufacturing, enabling adaptive, resilient, and data-driven production ecosystems.

7. REFERENCES

- [1] **Schema.org** implementation of CDIF metadata, <https://cross-domain-interopability-framework.github.io/cdifbook/metadata/schemaorgimplementation.html>.
- [2] Cox SJD, Gonzalez-Beltran AN, Magagna B, Marinescu M-C (2021) Ten simple rules for making a vocabulary FAIR. PLoS Comput Biol 17(6): e1009041. <https://doi.org/10.1371/journal.pcbi.1009041>.
- [3] Nicholson N and Stotl I (2024) A generic framework for the semantic contextualization of indicators. Front. Comput. Sci. 6:1463989, doi: 10.3389/fcomp.2024.1463989, vol. 6 – Oct. 2024
- [4] XX Papazoglou M., Krämer B., Raheem M., & Elgammal A. (2025). Medical Digital Twins for Personalized Chronic Care, 2025, 10.21203/rs.3.rs-5826330/v1 and <https://www.researchsquare.com/article/rs-5826330/v1>
- [5] Dehghani, Z. (2022) Data Mesh: Delivering Data-Driven Value at Scale, O'Reilly.
- [6] Mcguinness D. Ontologies Come of Age. In: Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential [Internet]. Wadern, Germany; 2003. p. 171–94. Available from: https://www.researchgate.net/publication/221024668_Ontologies_Come_of_Age
- [7] Bernstein W.J., Hedberg Jr. T. D., Helu M., Barnard Feeney A. (2017), Contextualising manufacturing data for lifecycle decision-making. Int. J. Product Lifecycle Management, Vol. 10, No. 4, 2017
- [8] Berrueta D, Phipps J. Best Practice Recipes for Publishing RDF Vocabularies [Internet]. Cambridge, Mass. USA: World Wide Web Consortium; 2008. Available from: <http://www.w3.org/TR/swbp-vocab-pub/>