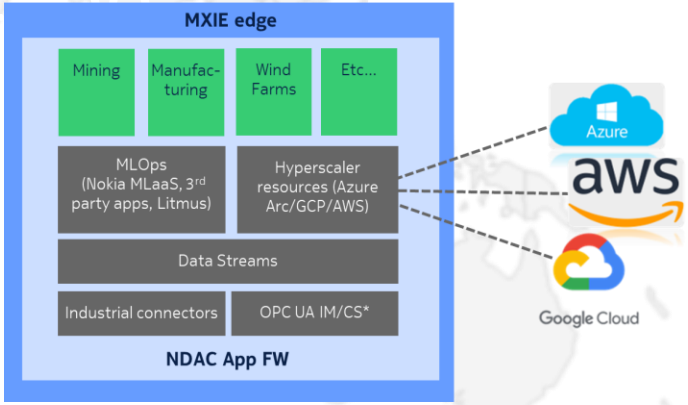
Data-driven enterprise architecture with standardized, distributed data products

Tomi Lahti
Nokia DAC / Enterprise Campus Edge

About presenter

Tomi Lahti, Senior Product Manager, MXIE Apps

- Studied intelligent & learning control, thesis on soft sensors (2004)
- Industrial historian, analytics, ML pipeline and AI for OT digitalization at Finnish refining company
- Member of Finnish Automation Society & OPC Day Finland organizing committee
- LinkedIn
- Smart Refinery is built on Structured data (blog)



Nokia DAC: Edge compute and AI platform powered by private wireless and Wi-Fi

Technology Vision 2030 on the metaverse opportunities

Concepts of ‘Human Augmentation’ and ‘Digital-Physical Fusion’ frame this vision

Metaverse enablers

Metaverse opportunities

~ today

~ 2030

Human Augmentation

Handhelds
VR HMDs
Tethered AR glasses
Haptic-enabled remote control

Connected bio-medical implants Industrial exoskeletons
Ergonomic, untethered XR glasses
XR interoperability

Digital-Physical Fusion

Basic, organization-level digital twins
Smart sensor networks
Persistent virtual worlds & objects

Complex, enterprise-wide digital twins
Ecosystem interoperability
Interactive 3D digital twins

Consumer Metaverse

Enterprise Metaverse
(IT-centric)

Veturi scope of work

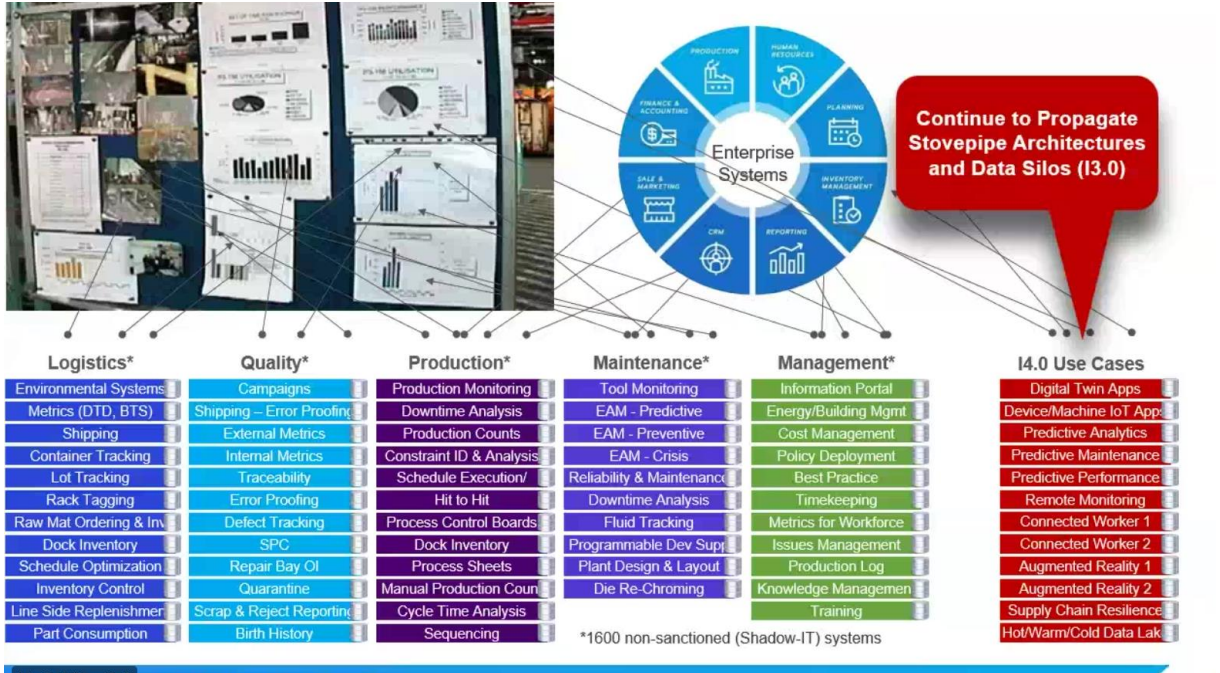
Industrial Metaverse
(OT-centric)

* Virtual Reality Head-Mounted Displays
** Augmented Reality
*** Extended Reality

Reality check: Digital transformation…

…will not happen with current approach.

"With many industrial facilities still operating at Industry 3.0, the need for structured and unstructured data orchestration, AI-driven automation, and scalable DataOps solutions has never been greater" – RTInsights/Verdantix



Tech debt (cumulating)

- Point-to-point integrations
- Vendor lock-ins
- Outdated connectivity
- No possibility to run ML on edge
- No shared understanding of data meaning and context

Path to Vision 2030

Current focus is on AI – applies to Digital Twins as well

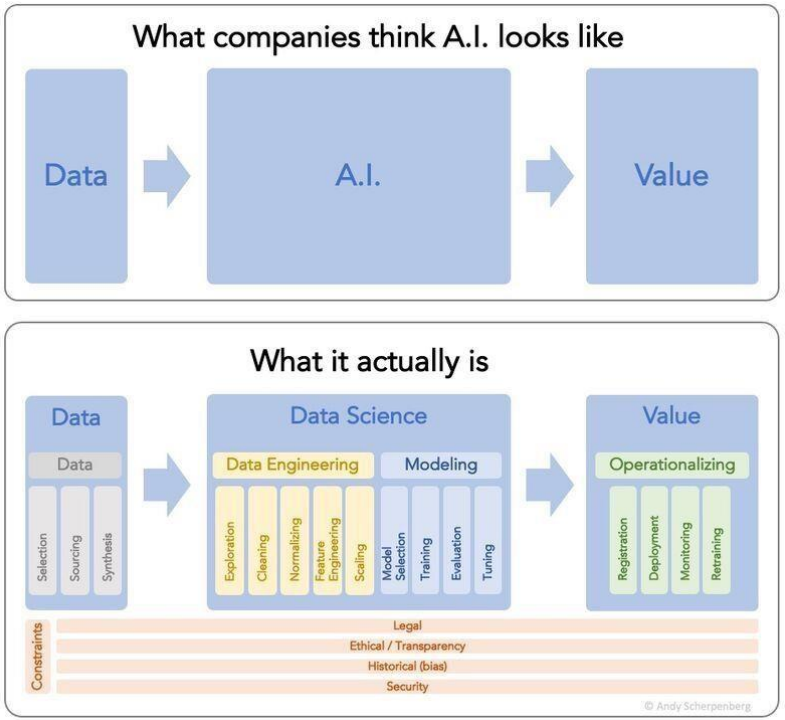
There are plenty of use cases for AI/ML and Digital Twins

- Those aren't' magic, and it most definitely is not a simple "bolt-on" or "plug-in"
- The foundation step is getting the data right. Sourcing, standardizing and contextualizing it

World is mostly brownfield

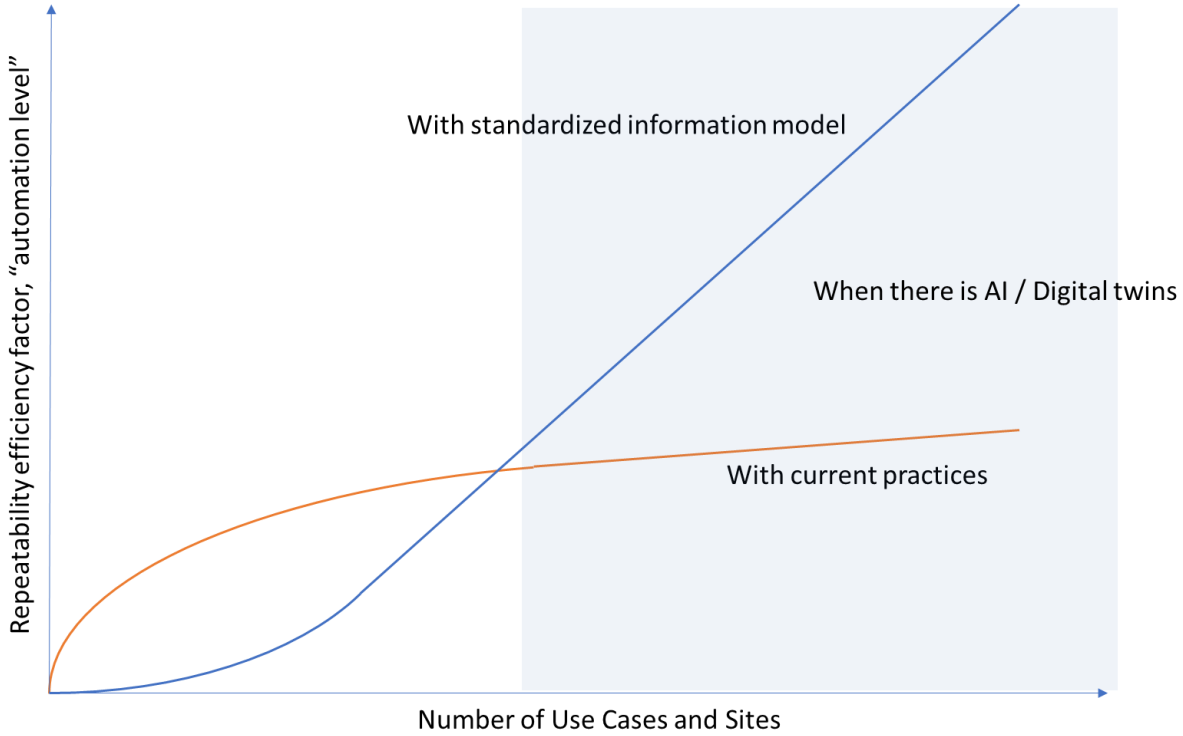
- How to bolt Digital Twins and AI to existing data infrastructure (ALM, ERP, MES/SCADA)? Add Industrial Data Spaces, with pressure to modernize enterprise data infrastructure and future SDA, vPLC, …

Paradigm shift and digitalization acceleration needed, from centralized solutions to decentralized and distributed computing architectures



Acceleration requires standardized interfaces (= data models)

Industries are struggling with replicability and scaling of I4.0



Emerging data architectures

- Data Fabric, Industrial DataOps
- Distributed execution, centralized management (Data Mesh)
- Semantic layer
- Swarm Intelligence
- Worker self-service engagement

Emerging OT architectures

- Open Industry 4.0, O-PAS, NAMUR NOA

Emerging requirements

- Carbon reduction, asset and production changes, resiliency, cyber security

Business must pay effort to find clarity how to solve their tangles. Other ways productivity keeps stalling, and the prominent AI scaling remains undone.

"To change performance, change thinking"
> Start OT data modelling

Industry 5.0 & Industrial Metaverse

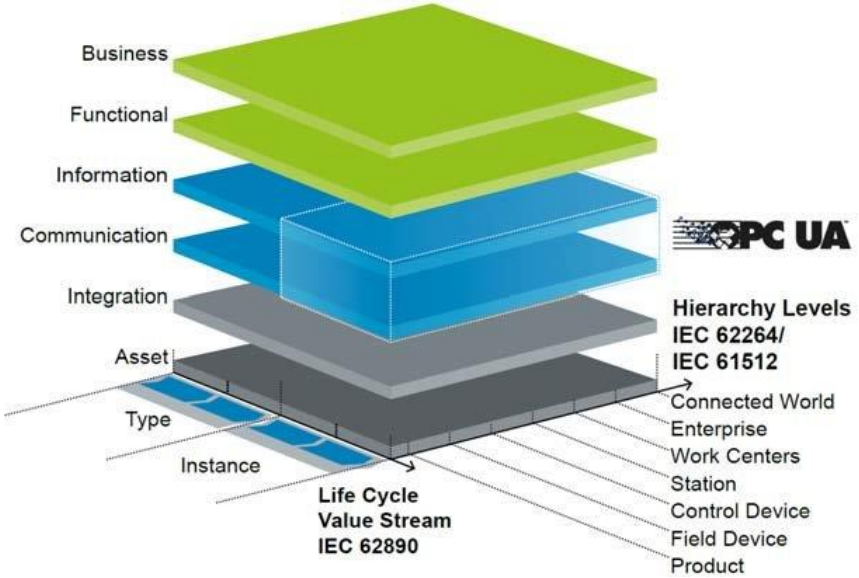
Industry 4.0 as enabler – focus on semantic information for modularity

Digital transformation is a communication challenge

- Human dimension & system dimension; human/machine readable information for decision-making
- Interoperability beats integration and separates success from stagnation

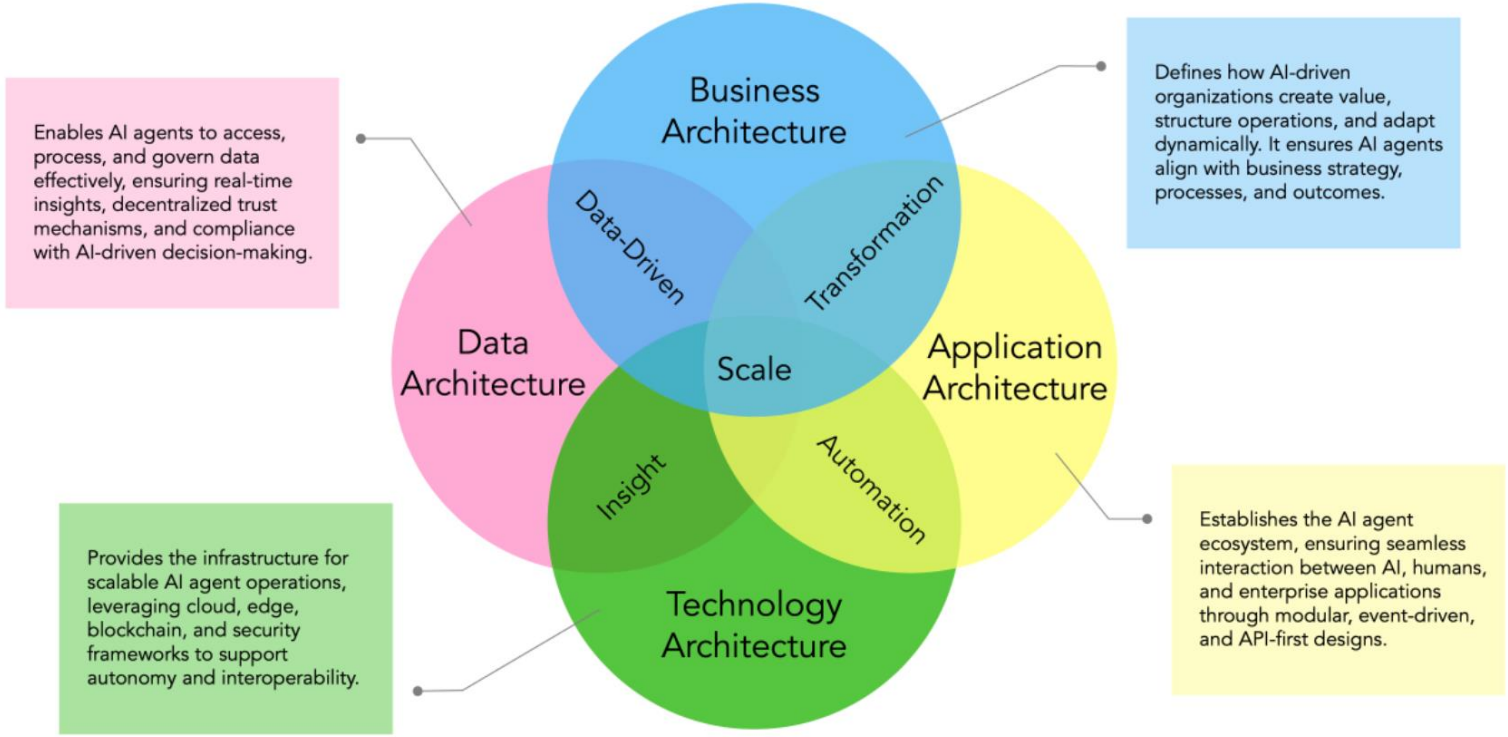
I4.0 benefits will realize by unlocking seamless semantics

- Intelligence/cognitive features
- IT/OT/ET convergence for modern data-driven enterprise architecture (with Data Spaces/AAS/DPP)
- Industrial Data Fabric AND Data Mesh features with distributed data products
  - "Central management, distributed compute" and Edge-cloud continuum
  - Industrial/Edge DataOps for consumable information and seamless data lifecycle interoperability, security, flexibility/resilience and maintainability



Enterprise Architecture 4.0

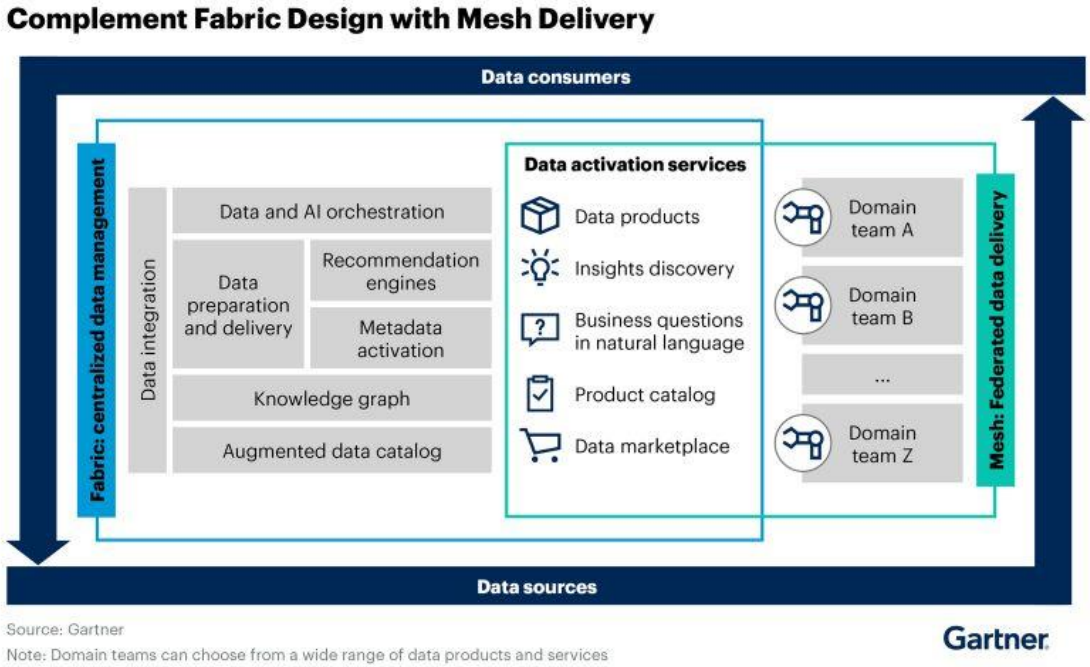
Agentic AI view – Jesper Lowgren, DXC Chief Enterprise Architect



Journey towards Enterprise Architecture 4.0 – OT view

I4.0 networked CPS still core requirement for Industrial Metaverse / Industry 4/5.0

- Data product
  - "Delivers a high-quality, ready-to-use data set that people across an organization can easily access and apply to different business challenges" (McKinsey)
- Support hybrid Data Fabric/Data Mesh Enterprise Data Architecture
  - Standardized data products are perfectly suited for a hybrid Data Fabric/Data Mesh approach. They can act as the building blocks for both centralized data services (Data Fabric) and decentralized data ownership (Data Mesh), promoting interoperability and consistency across the entire data landscape



Complement Data Fabric Design with Mesh Delivery

Semantics – mandatory for effective AI, DT, I4.0…

“Secret” ingredient for seamless collaboration

Enterprise ontology vs. subdomain (segment) ontologies
OPC UA Companion Specifications are "real-world operational technology ontology implementations" for specific industrial verticals like the process industry - practical, standardized implementation of semantic data models

Jérémy Ravenel:
- In AI, ontologies allow large language models (LLMs) to ground their reasoning in structured, factual data.
- In knowledge graphs, ontologies define the backbone that connects nodes and edges, ensuring the graph isn’t just a web of disconnected information but a meaningful network.
- In healthcare, ontologies like SNOMED CT standardize medical terminology, enabling interoperability between systems.
- In business, ontologies align siloed data, making it easier to retrieve insights for decision-making.

Semantic Integration Patterns for Industry 4.0

“Semantic information enrichment for existing software solutions will be necessary.

It will be a mandatory “feature” to properly prepare the data locally before the data exchange to “higher level” application scenarios”

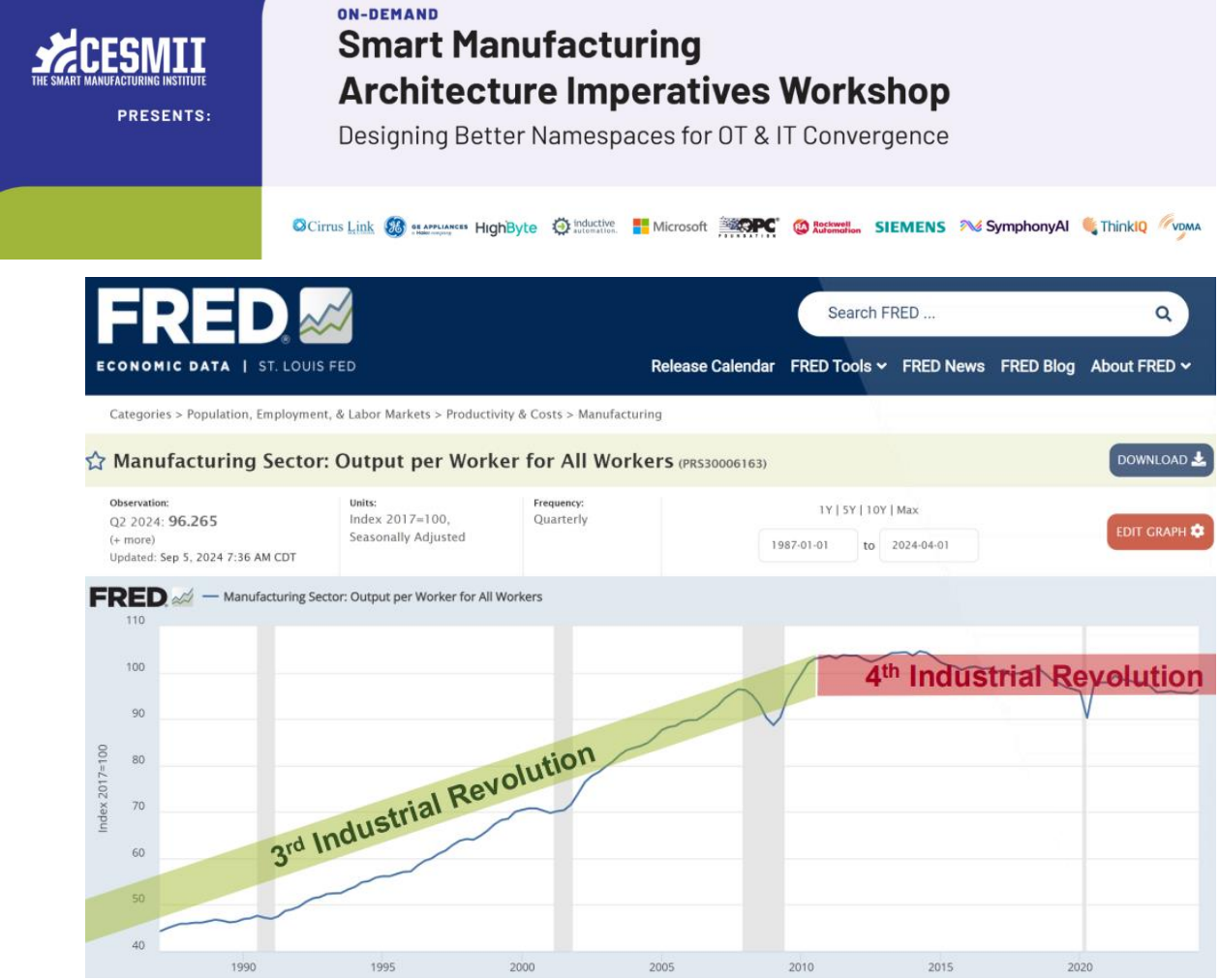
OPC UA and Data Products – benefits for Enterprise

Multiple enterprise-level benefits beyond interoperable, modular production

- OPC UA information models create **living documentation** of all production, with structure, metadata and context
- OPC UA standard communication with information models (Companion Specifications) is essential for efficient data (pipeline) automation and utilization trough **data lifecycle from OT to IT, and back**
- OPC UA has **built-in cyber security** for data-in-move
- OPC UA is perfect for real-time (**TSN**) control automation (Core Process Control/Industrial Control System) and is key data modelling tech for future wireless communication (**5G-ACIA**)
- **OPC UA over MQTT** (within Unified Namespace, **UNS**) brings flexible, scalable way to connect OT to IT securely,
  - Needed data structure, context and metadata for automatable, scalable, replicable and maintainable data-driven use case development

With OPC UA CS-based standardized data products:

- Enables Enterprise Data Architecture 4.0
- Unified Namespace with OPC UA over MQTT (Edge DataOps)
- Semantic interoperability throughout data lifecycle
- Supports Digital Thread (asset lifecycle management) with AAS & Industrial Data Spaces
- Supports Semantic layer implementation – edge & cloud
- Supports Knowledge Graph – based on OPC UA IM/CS
- …and finally delivers consumable real-time data for composable solutioning – True Industry 4.0
- … and makes Agentic AI and Cognitive Digital Twins efficient and smooth to deploy and maintain



[Link to WS]

The event highlighted 3 separate panels centered around the following Smart Manufacturing Imperatives:

- ✅ Open, Standards-Based Information-Modeling Strategy for Manufacturing Ops
- ✅ A Clear Set of Requirements for SM Interoperability Platform Compliance
- ✅ An open, common API for Manufacturing Systems

The Importance of Open Specifications & Interoperability for OT

Accelerate Innovation, Foster Competition, Drive Down Cost & Complexity

IT

1980s: Device Drivers were purpose-built (proprietary) for each device, and installed as a 'project'

OT

2020s: Every OT data source (sensor to machine to software) is non-standard and requires a 'project' to extract the right data for every use case & application (vendor lock-in)

Historic Example:

From Device Installation Projects to Plug 'n Play

40 Years of standards USB, Wi-Fi, TCP/IP...

- ☑ Software and hardware from 1000's of vendors can be installed by anyone
- ☑ Standard way to get data through a browser

Urgent Need:

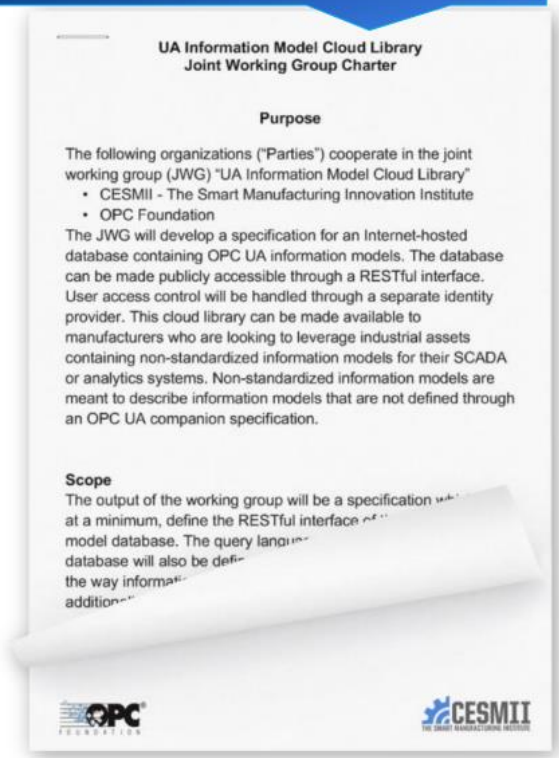
Open Specifications & Model-Driven Architectures

Secure, Interoperable, Plug & Play Infrastructure

- ☑ **Sensors, equipment and processes** can be discovered as objects with graph relationships
- ☑ Applications can be developed against an open API
- ☑ Facilitate the more effective creation and reuse of trained AI models

A Global Community Investing in Manufacturing Information Modelling

An Open Standard for Information Model Storage, Search & Access (Human/Machine Readable)



UA Information Model Cloud Library
Joint Working Group Charter

Purpose

The following organizations ("Parties") cooperate in the joint working group (JWG) "UA Information Model Cloud Library"

- CESMII - The Smart Manufacturing Innovation Institute
- OPC Foundation

The JWG will develop a specification for an Internet-hosted database containing OPC UA information models. The database can be made publicly accessible through a RESTful interface. User access control will be handled through a separate identity provider. This cloud library can be made available to manufacturers who are looking to leverage industrial assets containing non-standardized information models for their SCADA or analytics systems. Non-standardized information models are meant to describe information models that are not defined through an OPC UA companion specification.

Scope

The output of the working group will be a specification wh
at a minimum, define the RESTful interface of
model database. The query langu
database will also be defi
the way informati
additio

Joint Working Group

- ABB
- Ascolab
- Atlas Copco
- amazon
- Beeond
- Bitctrl
- Bosch
- Capgemini
- CESMII
- Endress + Hauser
- Equinor
- Google
- Hilscher
- IBM
- Idata
- Inray
- Microsoft
- Prediktor
- Renault
- SAP
- Siemens
- Softing
- VDMA
- Wago
- 4CE Industry

More than 800 companies contributing
More than 40 active working groups
More than 25 sectors in active development

Standard OPC UA Companion Specs
- VDMA
- MTConnect
- Industry Associations

CESMII Community & Project Teams
- Machine Builders
- System Integrators
- Subject Matter Experts

CESMII Ecosystem
- ThinkIQ
- Inductive Automation
- Aveva OSI AssetFramework
- PTC ThingWorx Templates
- Rockwell Automation
- SymphonyAI Templates
- Siemens

Cloud Library (Open Specification)

Smart Manufacturing Marketplace

SM Apps

SM Profiles

CESMII Requirements for SM Compliance

The Journey to Interoperability

Open, Standardized Access

Open API

SM Imperative #3: Create an open, common API for Manufacturing Systems, Rapid App Dev, Scaling AI Deployments, EAI, Supply Chain Optimization…

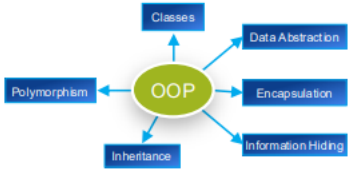
- ✓ Establish a standard API, consisting of a base set of server primitives that a wide array of platforms can implement to commoditize access to manufacturing data

Graph Aware, Graph Preserving

Knowledge Graph

Object Oriented & Type Safe

Namespace

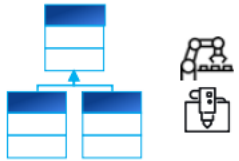


SM Imperative #2: A Clear Set of Requirements Enabling Manufacturing Platform and Application Interoperability (Compliance = SM Interoperability Platform)

- ✓ Instantiate SM Profiles as Type Safe Objects from the SM Marketplace
- ✓ Persist Instance Objects & Relationships (present and past) in a Knowledge Graph

Standardized Information Models

SM Profiles

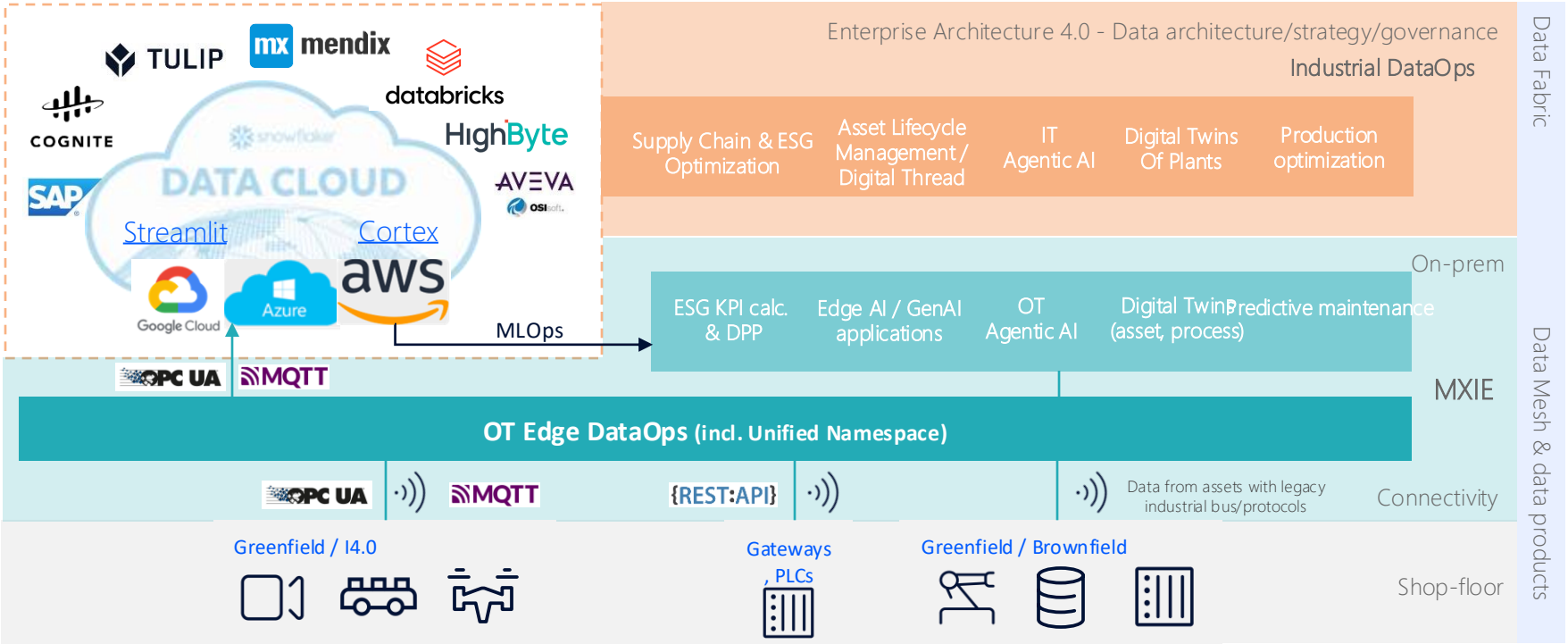


SM Imperative #1: Open, Standards-Based Information-Modeling Strategy for Manufacturing (& related Supply Chain) Devices, Assets & Processes

- ✓ Leverage and contribute to a global community building information models (SM Profiles)

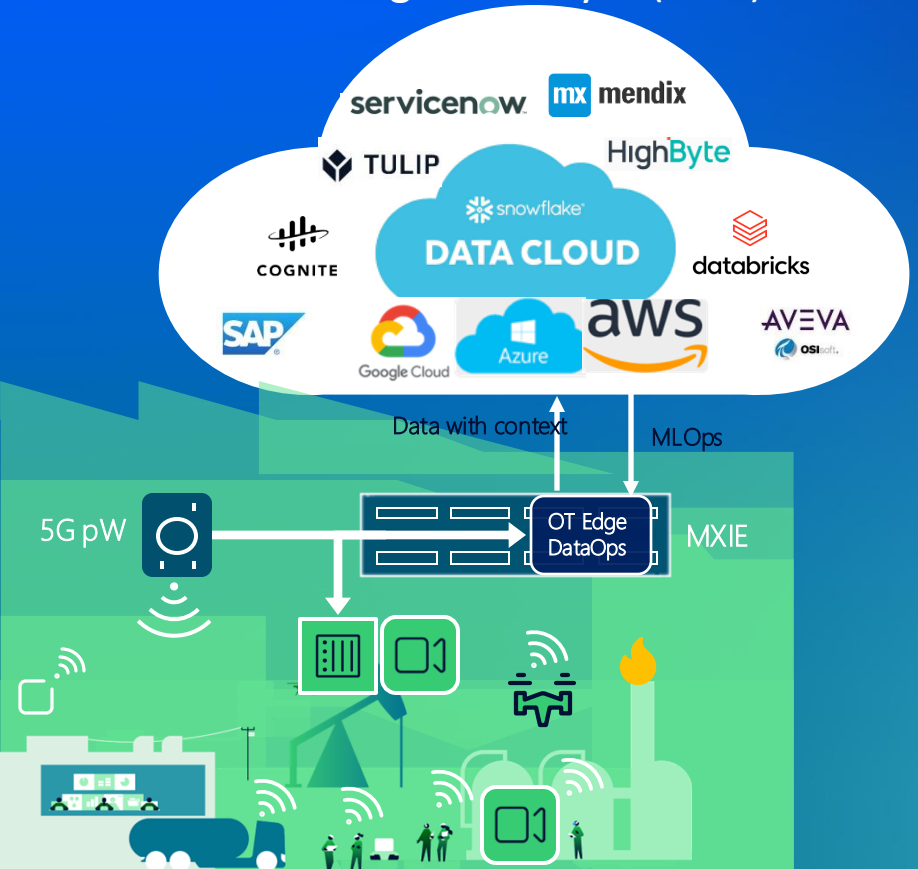
Edge DataOps with Industrial Data Fabric

Edge DataOps with standards enables Industry 4.0 promises at scale



Edge DataOps – Edge AI, Digital Twins and modular integrations

Consumable edge data layer (UNS) enables event- and data-driven use cases



Consumable real-time edge data layer with Unified Namespace (UNS)

Connect, collect, aggregate and harmonize data at edge for single source of truth for different edge use cases

Handle events in real-time and make automated analytics / control to integrate suitable enterprise workflows

Event-driven, real-time harmonized edge data layer provides agile, flexible integration point for edge AI / inferencing and Digital Twins

Centrally managed, distributed execution to combine data fabric & mesh benefits

Seamless integration to hyperscalers (Azure, AWS, Google, Snowflake) and Industrial Data Platforms

Harmonize data between sites/plants for scalability and flexibility

High cyber security and compliance with central governance and data security

Productivity gains throughout supply chain and data ecosystems

Simplifies multiple use case deployment for edge AI: Downtime Prediction, Anomaly Detection or Energy Optimization, for example

Enable collaboration and support Intelligent Supply Networks for Smarter Manufacturing with readily shareable data

Address top challenges Enterprises face to scale edge AI and reduce tech debt across the enterprise

OT Edge DataOps = Nokia MXIE Data Lake + LITMUS + crosser + OPC UA Forge

Read More: OT Edge DataOps

NOKIA

Shop-floor decision-making - demands for Edge Platform

Operational intelligence sets clear demand for edge & smart IIoT app capabilities

1

AI and DT on-prem

Proliferation of artificial intelligence, machine learning and networked digital twins at scale - on-prem

2

Hybrid edge-cloud

Processing edge data needs to be performed at the edge and in the cloud or enterprise.

3

OPC UA (over MQTT)

IEC 62541, framework for industrial interoperability based on data models, context and metadata – data modelling is key for all intelligent operations!

4

Semantic interoperability

Interoperable API's (data products) which serve standardized information modelling are key for successful intelligent use cases

5

Open standards - tested

Supporting open Industry 4.0 standards enables interoperability and secure data sharing to partners – best served via collaboration like Open Industry 4.0 Alliance

6

OT Edge DataOps

Edge must support enterprise data governance and Industrial DataOps (semantics, knowledge databases)

Summary - Edge enables Modern Data Governance at any Industry

Modern data approach starts from edge and reach trough data lifecycle

1

Modern Edge Data Stack: Smart IIoT at edge abstracts protocol and asset diversity and enables OT digitalization acceleration

2

Data Models: Smart IIoT at edge delivers structured OT data with metadata to IT tools to scale up OT digitalization, and fosters OT teams' data ownership

3

Refine IT/OT convergence: Smart enterprise data governance understand OT data nature and specialty

Our solution

Building the core capabilities for sustainable industrial metaverse

Main focus areas

1	2	3	4
Technology enablers for industrial metaverse	Architecture and platforms for sustainable industrial metaverse	AI/ML-enabled industrial applications	Industrial Metaverse ecosystem, new business opportunities and commercialization
Define and build new hardware and software enablers for the industrial metaverse architecture, platforms and use cases	Build essential capabilities to enable networks and platform's security, energy efficiency and optimal performance	Create and validate future capabilities for real-time XR media communication and industrial applications	Lead activities to build the Finnish Industrial Metaverse Ecosystem to accelerate global business opportunity development and commercialization

Edge DataOps

Operational AI, Digital Twin & Industrial Metaverse enabler

"How to make OT data consumable for any use case?"

Core Implementation Principles

Building blocks for Edge and Industrial AI success

Data Consumability

- Harmonize data at the edge
- Standardized information model
- Abstract field-level complexity

Solution Independence

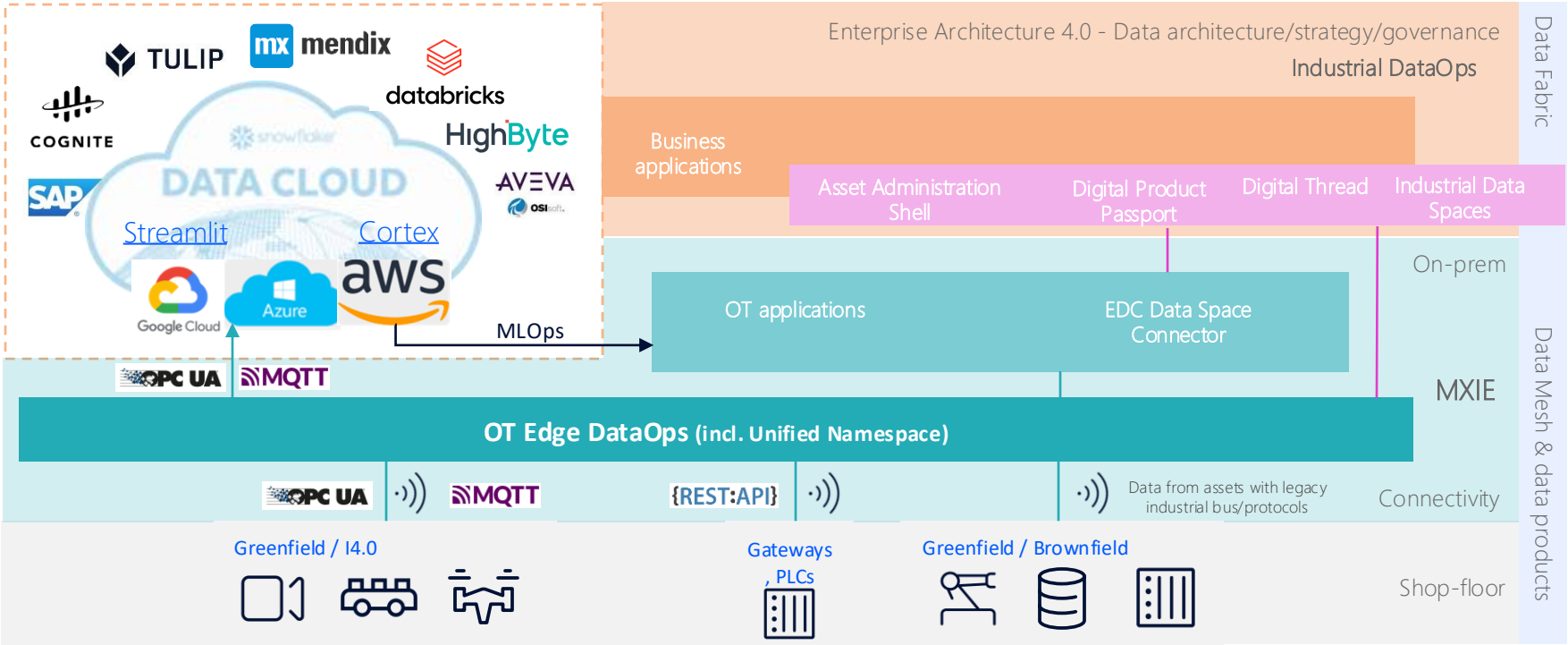
- Decouple data generation from usage
- Reduce integration complexity
- Flexible tools for use-case development

Future-proof AI Stack

- Wide ranged support – GenAI, Causal, Agents, Multi-modal
- On-prem inference and fine-tuning

Unified Namespace with Industrial Data Fabric

UNS and AAS: A Winning Combination for Industrial Metaverse & DPP/AAS



Industrial Information Models & Interoperability standards

Coexistent international development examples

Digital Twins

- ISO 23247(US), Automation systems and integration – Digital Twin framework for manufacturing
- IDTA - Industrial Digital Twin Association – The Standard for the Digital Twin "Asset Administration Shell"
- DTC (Digital Twin Consortium)

Digital Factory

- IEC 62832 (EU), Digital Factory Framework
- DIN SPEC 91345 (DE), RAMI 4.0 (incl. AAS)

Data dictionaries (master data management)

- ECLASS (EU), Cross-sector standard for classification of product and services
- IEC 61360, Common Data Dictionary

Importance of Industry Standards with Regards to a Comprehensive Automation Strategy – ARC Advisory Group

Industry standards with (partial) semantic mappings

- OPC UA (OPC Foundation), OPC UA Companion Specifications – CESMII SM Profiles (US)
- ISA-95, ISA-88, UMATI, …

"European" approach

- Based on I4 standards: AAS & OPC UA CS
  - AAS as standardized digital twin
- Easy migration of brownfield operations – Edge DataOps maps OT data to CS/AAS
- Vertical Integration: CS throughout data lifecycle (Edge DataOps)
- Horizontal integration: AAS throughout asset lifecycle
  - Intersection: Operations / Production


NOKIA