



MOTIUS

WE R&D.

Data Infrastructure

Motius GmbH

December 08, 2025 14:23 (b2b168d)



Data Infrastructure

Status Quo

Working with companies like Siemens, MTU, BMW, and Bosch, we built several data platforms to address challenges with the typical status quo of ERP and MES systems:

Inflexible Data Sources

Companies manufacturing highly customized products with short delivery times cannot effectively optimize their production, if every new automation application takes weeks or months to implement, because of complex data infrastructure.

Disconnected Data

Business intelligence is difficult, because there is no single source of truth for data. Companies who rely on SAP and its data warehouse solutions have to contend with a complex and expensive data landscape.

No Self Service

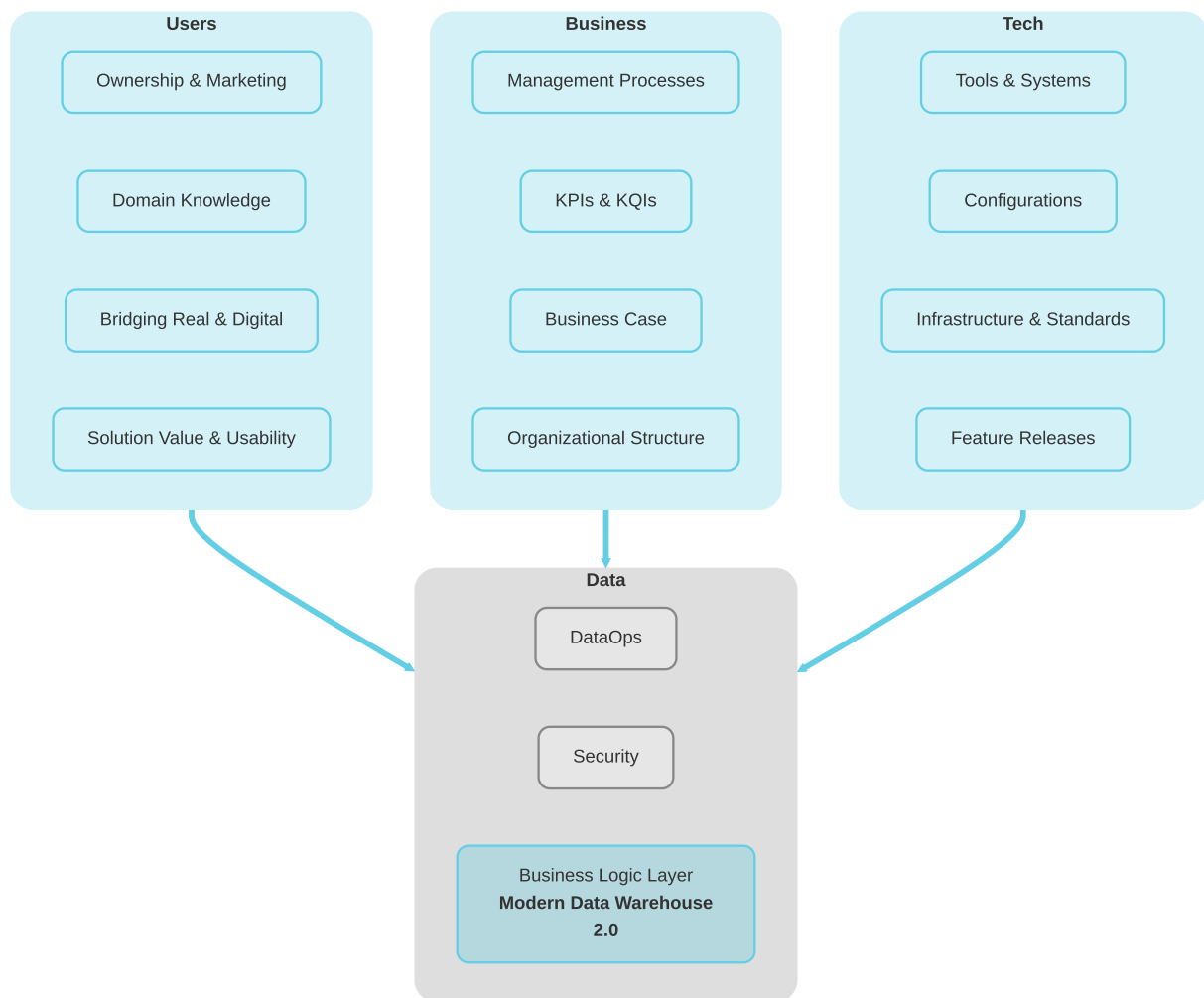
Because of high licensing and compute costs, systems like SAP HANA become locked down to a few experts, who are overwhelmed with requests for data. Business users cannot access the data they need to make decisions.

Isolated Production Data

Factory managers cannot optimize production lines because real-time sensor data never reaches the people making scheduling decisions. Production data is also often siloed in different databases, log files, and separate QMS systems.

Modern Data Warehouse

A modern data-warehouse (DWH) architecture can solve these problems by providing a single source of truth for all data, enabling self-service access to data for application engineers, and even allowing for real-time streaming & processing of sensor data.



This architecture improves upon the status quo in many ways:

- ✓ **Single Source of Truth** for all data, including production data, sales data, and sensor data
- ✓ **Self-Service Data Access** for application engineers, enabling them to build new applications and features quickly
- ✓ **Real-Time Data Processing** for sensor data, enabling real-time monitoring and optimization of production lines
- ✓ **Scalable Architecture** that can handle Petabyte-scale data warehouses, a large numbers of users, and many business units

Reference Project: Serverless Reporting Engine

Our client, a large DAX company, was relying on an SAP-based system for monthly reporting of global procurement KPIs. The system had accumulated technical debt to the degree that maintenance and new feature development cycles were spiraling. We re-built the system based on a serverless architecture around a PySpark-based ETL-pipeline to ingest and enrich data into a Snowflake Data Warehouse.

With the release of Snowpark, AWS Glue jobs could easily be replaced to **cut operating costs by ~60%**.

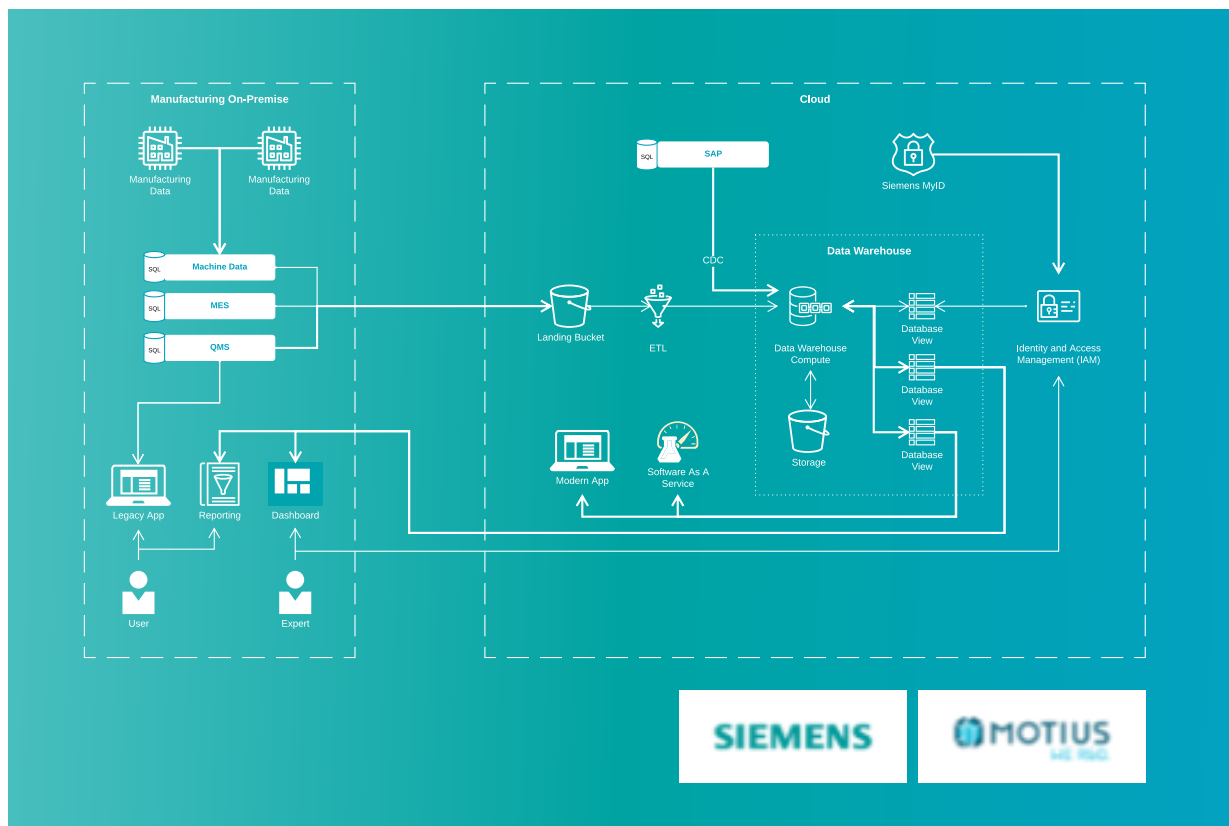
Results

- ✓ Operating costs **cut by ~60%** via Snowpark migration
- ✓ Processing 40k (and growing) procurement files per month (10M records) for over 1k users
- ✓ Technological shift enables much quicker development cycles, with new features being implemented weekly

Technological Shift

Technology Shift	Before	After	Effect
Infrastructure	Self-hosted	Serverless	Operating costs cut by ~60%
Data Processing	Sequential	Parallel	Much faster jobs, much closer to real-time reports
Architecture	Monolith	Modular	Implementing new usecases and ingesting new data sources became a weekly occurrence
Deployment/Updates	Infrequent updates	CI/CD	Continuous integration & deployment allows these frequent changes, while avoiding downtime & errors in data sources

Example Architecture



At Siemens, we transformed a SAP HANA based data lake into a cloud-based data warehouse. We worked with over 20 data sources, including SAP, Salesforce, and various manufacturing data sources ranging from OPCUA, to one-off SQL databases, and Excel reports.

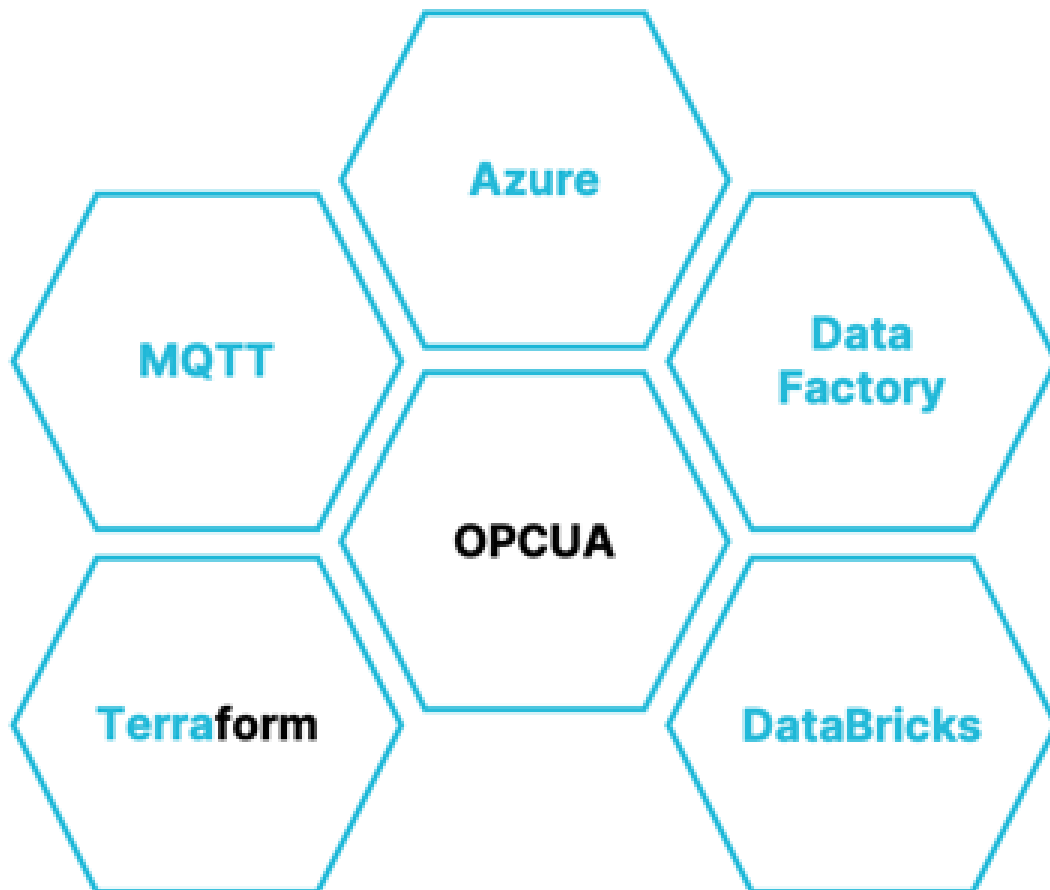
Reference Project: Manufacturing Data Platform

The aim of this project was to build a secure data platform  for factories to support future expansion plans. We built a template data platform that is ready to be used for factory data and applications and doesn't require an on-premise installation. The platform can not only ingest data from different sources, but also includes real time data streaming.

The benefit of this solution is near-real-time notification and tracking of any data and visualization of the results for continuous process optimization.

Results

- ✓ Additional platform features provide ability for **continuous process optimization**
- ✓ We built a scalable solution able to be used as template for **future factory expansions**
- ✓ Solution was implemented in a **secure environment** adhering to client's cyber security standards



Success Factors

There are some clear success factors for data transformation projects:

One-Size-Doesn't-Fit-All

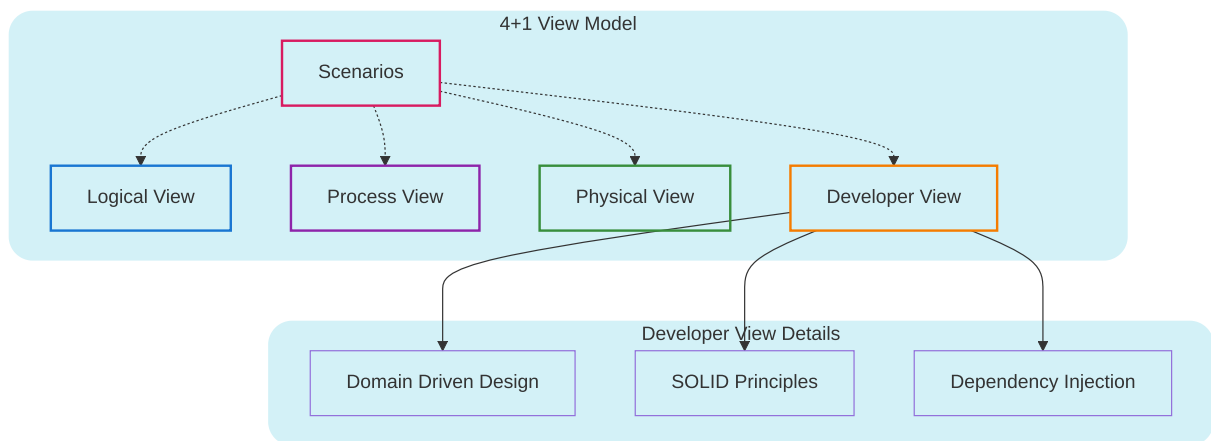
Every company has its own data landscape, and there is no one-size-fits-all solution. We work with our customers to understand their specific needs, and build a data platform that fits their requirements.

These are some common factors, that influence the architecture and design of a data platform:

Requirement	Effect on Data Infrastructure
Early adopters of digitalization in the form of ERP and MES systems	Need to integrate with legacy and modern systems; complex data landscapes
Small lot sizes	Requires flexible, scalable data processing to handle frequent changes and variability
High product variety	Demands adaptable data models and pipelines to support diverse data sources and structures
Short delivery times	Necessitates real-time or near-real-time data ingestion and processing for timely decision-making
Mix of min/max warehousing and made-to-order production	Sourcing and manufacturing need completely different views on the data, usually not possible in SAP
High degree of automation, but a lot of pressure to automate more	Infrastructure must be extensible and support quick iterations for new automation and analytics capabilities

Architecture & Systems Engineering

We start all of our projects with a systems engineering phase, to create a clear architecture and understand the domain our customers operate in.



We deliver an architecture design document, and first tangible results with a technical proof of concept (PoC) in only a few weeks.

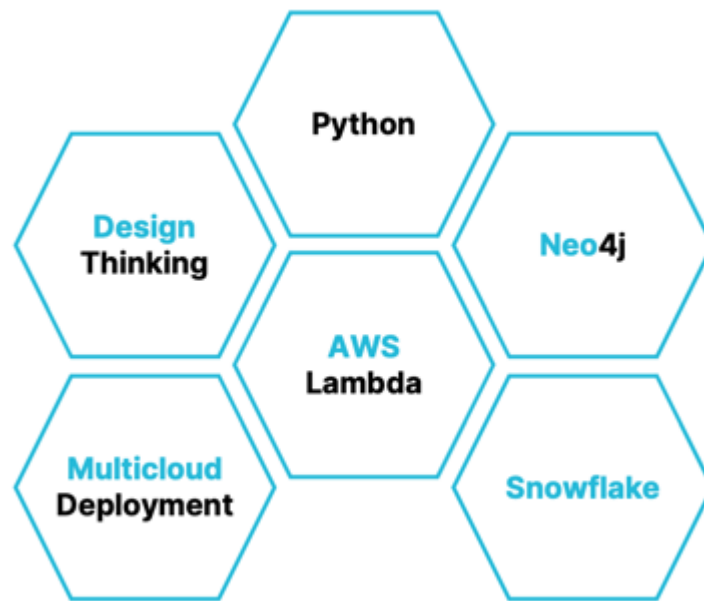
Change Management plays a crucial role in the success of data projects. We usually work in parallel to the existing data infrastructure, and begin implementing new use cases that would have been impossible with legacy systems.

Reference Project: AI-based Selling Recommendations

The objective was to increase the effectiveness of sales representatives (reps) and managers on a global level by creating individual digital sales assistants. The system generates selling recommendations on an individual level based on a huge amount of internal sales data.

The modular structure allows high flexibility in technical environment, broadness and depth. As a result, the digital assistants now help sales reps make better decisions, prioritize leads, find opportunities and ultimately: save time.

- ✓ **Tremendous time savings** based on process automation
- ✓ Digital assistants create **individual recommendations**
- ✓ Architecture can deal with huge amounts of data (Petabytes) & ~30.000 users globally
- ✓ **Integrated into the tools** that sales reps use every day



Iterative Development

Even transforming Petabyte-scale data warehouses should be done iteratively. We work in bi-weekly sprints, talk to important stakeholders daily or weekly, depending on the phase of the project. Modern development workflows with Infrastructure as Code (IaC) and Continuous Integration/Continuous Deployment (CI/CD) allow us to deliver new features quickly, and with high quality.

Modern APIs

Snowflake, Databricks, and all other modern data warehouses have APIs that let our customer's application engineers build better production automation, integrate new data sources, and combine data with only a quick review and approval from the data engineering team.

Application at Resideo

With our experience in building data platforms, and your specific domain knowledge, we believe that we can help create a new foundation for more flexible production planning, and automation.



- 1 Our DataOps team can set up a new cloud-based architecture using Microsoft Fabric and Azure Data Factory as a basis for a modern data warehouse.

② We would prioritize usecases, that are currently hard to implement with the existing data infrastructure.

③ In a matter of weeks, we can deliver a first PoC that shows the potential of the new architecture.