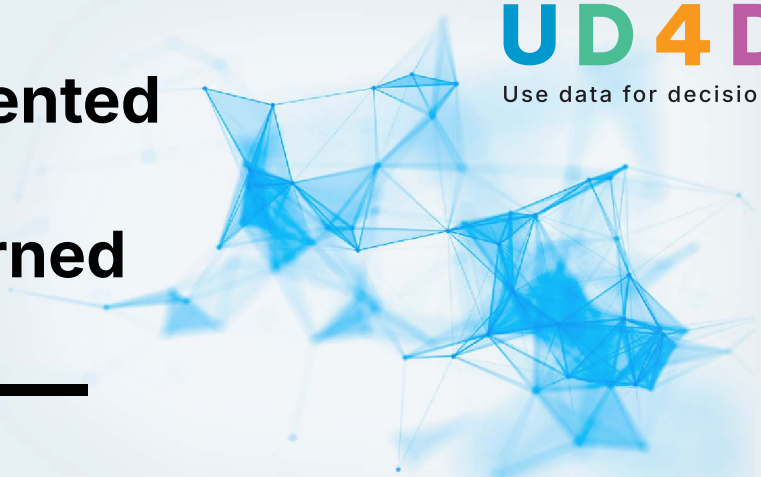
CASE STUDY

Replacing a Fragmented Legacy Landscape with a Single Governed Cloud Data Hub

UD4D
Use data for decisions



CLIENT

Global biopharmaceutical company operating across multiple divisions.

THE CHALLENGE

The client needed a unified, cloud-based data hub for its Global Market Access (GMAX) domain to replace a fragmented legacy ecosystem.

- Inconsistent data structures and naming conventions across SAP, Oracle, BPC, TM1, VEEVA, IQVIA, and multiple manual sources
- No central data catalog, metadata repository, or governance model for ownership, stewardship, or quality
- Manual ingestion and inconsistent transformation logic on non-scalable on-premise pipelines (Informatica, Teradata, Oracle)
- No automated deployment or standardized modeling, limiting analytics across GMAX product lines (NGP, Tender1, ACADS, Forecasting)

SOLUTION

UD4D led the architecture, migration, and governance enablement of the GMAX Data Hub.

- Design of an AWS data hub consolidating structured, semi-structured, and manual datasets, with ingestion patterns (Airflow, Glue, AWS Transfer Family) and a Data Validation Wall enforcing format, completeness, and structural consistency
- Migration of legacy pipelines from Informatica, Teradata, and Oracle to cloud-native processing and automated orchestration, preserving existing business logic
- Introduction of WhereScape 3D and RED for metadata-first, model-driven design and automated, version-controlled deployment across all GMAX data products
- Co-design of the GMAX Data Governance Framework covering ownership, stewardship, approval workflows, and data quality issue management
- Standardization of core data structures with ontology-based harmonization, and integration of business and technical metadata, glossary, and lineage into Collibra

BUSINESS IMPACT

1st **unified cloud data hub for GMAX:** Consolidated structured, semi-structured, and manual datasets into a single governed platform on AWS, with structural consistency enforced by the Data Validation Wall, replacing a fragmented multi-system landscape.

3 **legacy on-premise platforms migrated to AWS:** Informatica, Teradata, and Oracle pipelines were replaced with cloud-native processing and automated orchestration, while preserving existing business logic.

6 **core data domains standardized for consistent analytics:** Standard structures for Product, Geography, Customer, Competitors, Pricing, and Tenders were harmonized through an ontology-based model, with WhereScape automating model-driven design and version-controlled deployment.

4 **GMAX product lines enabled for cross-product analytics:** Consistent, governed data with business and technical metadata and lineage in Collibra now supports NGP, Tender1, ACADS, and Forecasting, with capacity for future GMAX initiatives.

TECHNOLOGY STACK AWS (S3, Glue, IAM, Redshift, Transfer Family) | Apache Airflow | WhereScape RED & 3D | Collibra | Alteryx | TIBCO EBX | Dataiku | SAP HANA | Oracle | BPC | TM1 | Power BI

UD4D s.r.o. | Rohanské nábřeží 678/25 | 186 00 Prague | Czech Republic | +420 776 660 450 | **www.ud4d.com**