Daten Virtualisierung

als Enabler für Data Governance, Self-Service und AI

Data Excellence Konferenz 2025

Wien, 2025-04-10

Drei. Macht's einfach.

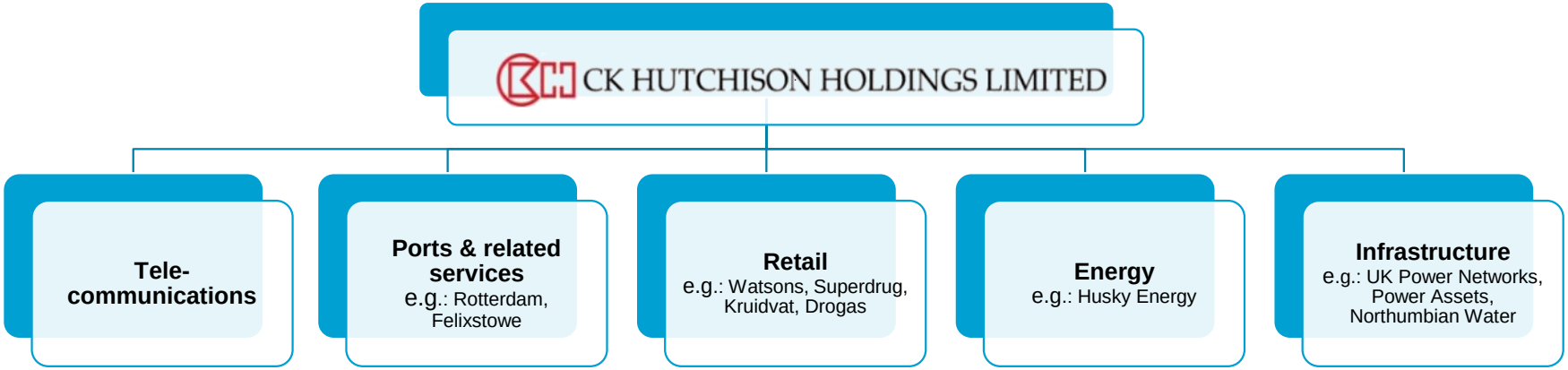
1. Who are we?
2. Before & After Data Virtualization
3. Areas of Improvement
   (with Informatica)
4. Things for the future
   (with Denodo & AVDM)
5. Things we learned

© Hutchison Drei Austria GmbH – Confidential

Who are we?

© Hutchison Drei Austria GmbH – Confidential

CKHH - CK Hutchison Holdings Ltd.



3Group Europe
Italy, **Austria**, Denmark, Sweden, UK, Ireland

Hong Kong & Macau

Indonesia, Vietnam, Sri Lanka

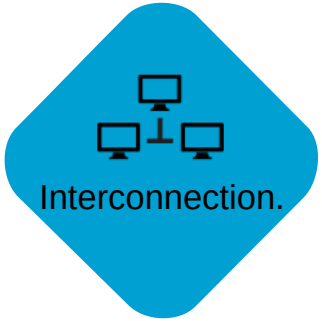
Chairman:	Victor Li
Headquarter:	Hong Kong
Trading centre:	CKHH is listed on the HK stock exchange
Employees:	about 300,000 in over 50 countries (December 31, 2023)
Structure:	Conglomerate of 5 core businesses
customers:	40.2 mio customers within Europe (December 31, 2023)
Revenue Group:	HK$ 77,925 mio (about € 9.36 billion) in Europe (Dec 2022)

Drei:
Everything from a single source.

Drei Business: Ideal Solutions for corporate customers.

Individual tariffs for small, medium and large enterprises incl. Mobile Device Management.

Mobile and fixed line-solutions support efficient work in the office and on the go.



The connection of individual enterprise locations enables efficient and secure cooperation; campus networks for enterprise locations.

M2M/ IoT solutions for new business areas, Digital Signage, **Data Analytics**.

Drei Business: Ideal Solutions for corporate customers.

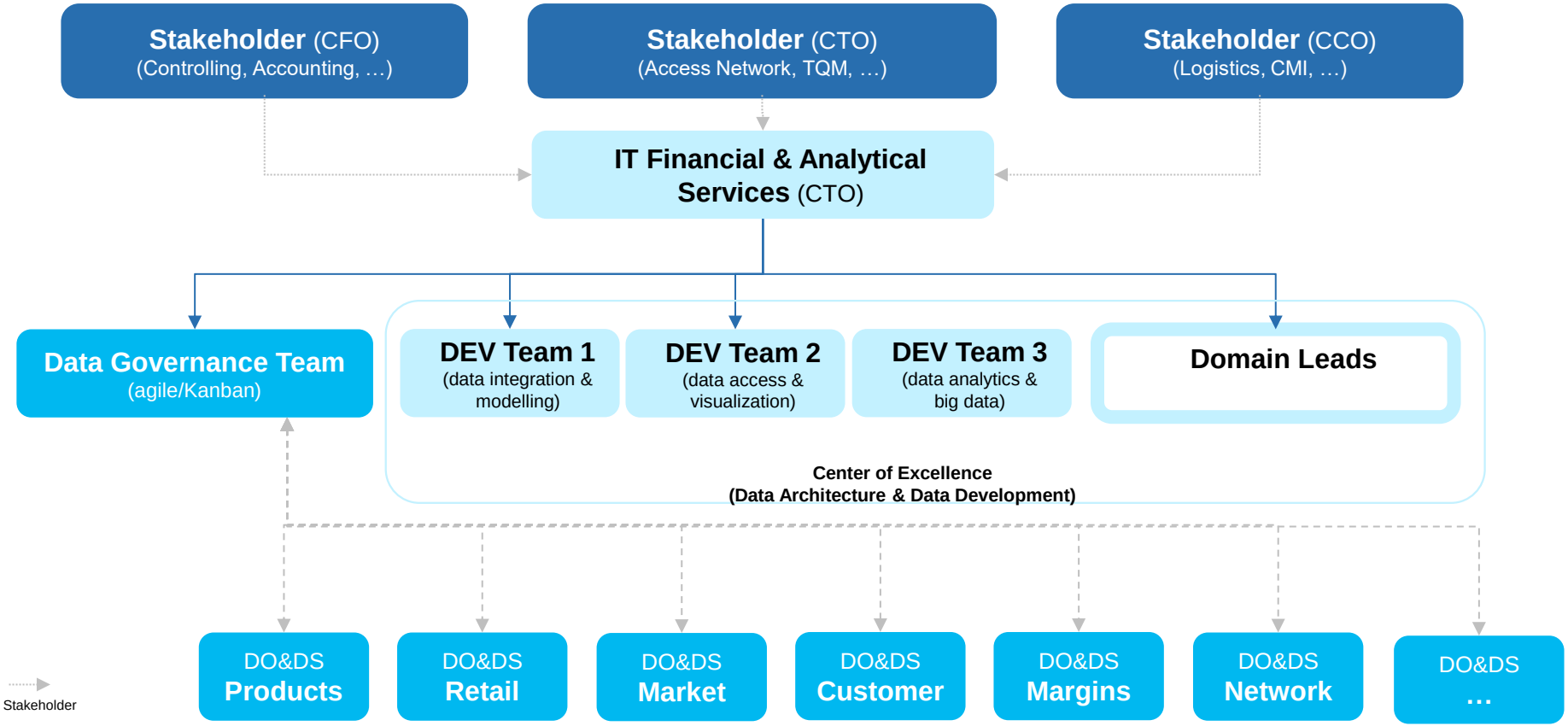
Drei Support of IT-Infrastructure; Cloud and Enterprise Filesharing

E-Card-Connection and complete individual solutions, e.g. for pharmacies and medical doctors.

Individual consultation from the first SIM-Card.

Competent, friendly, flexible and in proximity.

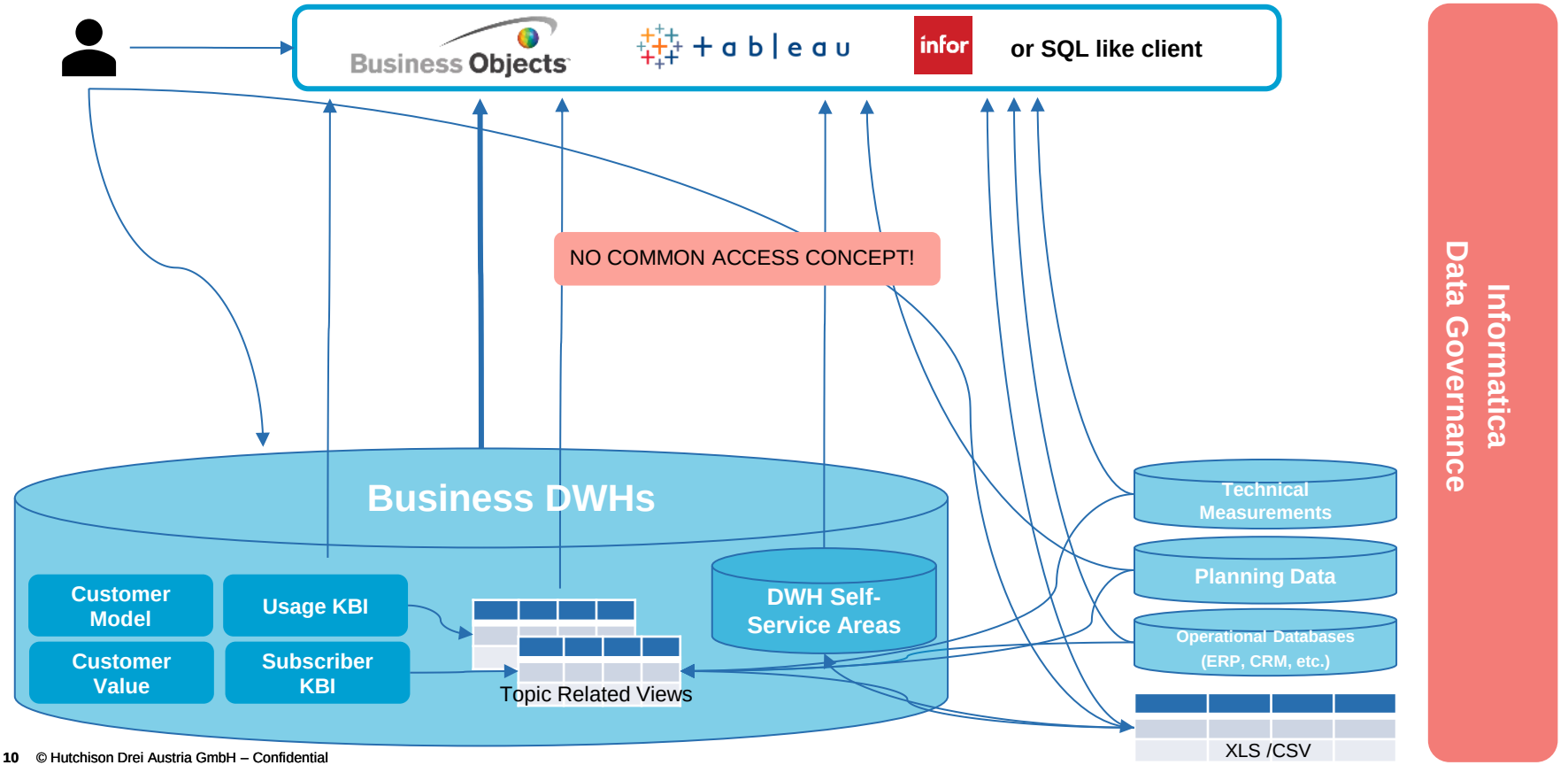
Organizational embedding for data analytics.



Before & After Data Virtualization

© Hutchison Drei Austria GmbH – Confidential

Before (2021): Data access handled individually



 © Hutchison Drei Austria GmbH – Confidential

Virtual Data Access Layer: Why ?

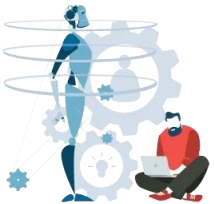
Faster time to access our data

Reducing data replication

Support Data Governance

Centralized control of data query executions

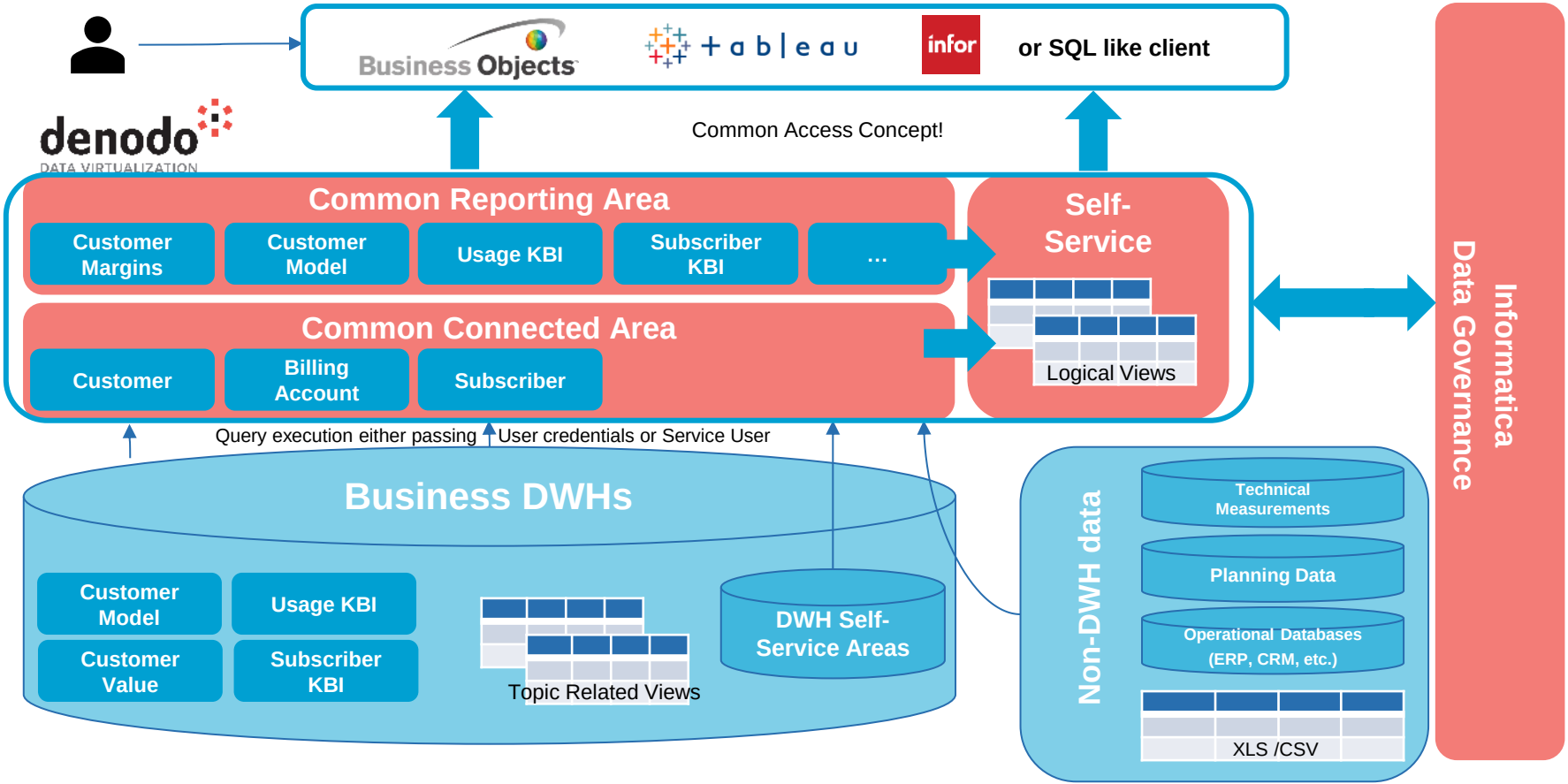
Meet security, Privacy & Audit requirements



Enables to access and understand new data sources

 © Hutchison Drei Austria GmbH – Confidential

Now: Centralized Access to Data



Illustrative Example

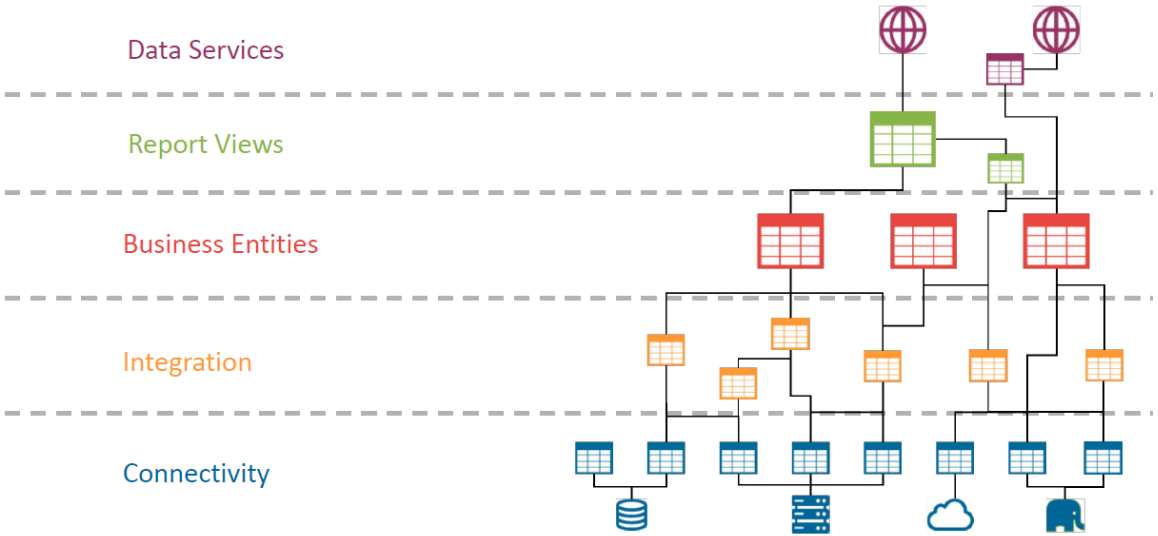
 © Hutchison Drei Austria GmbH – Confidential

Areas of Improvement

 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

- **Repository** of data sources
  - Re-usability once connected
- **Layered Data Architecture**
  - From Source to Consumption
- **Last mile transformation** to match data consumption / reporting tool needs
- **Rapid prototyping** with interface views

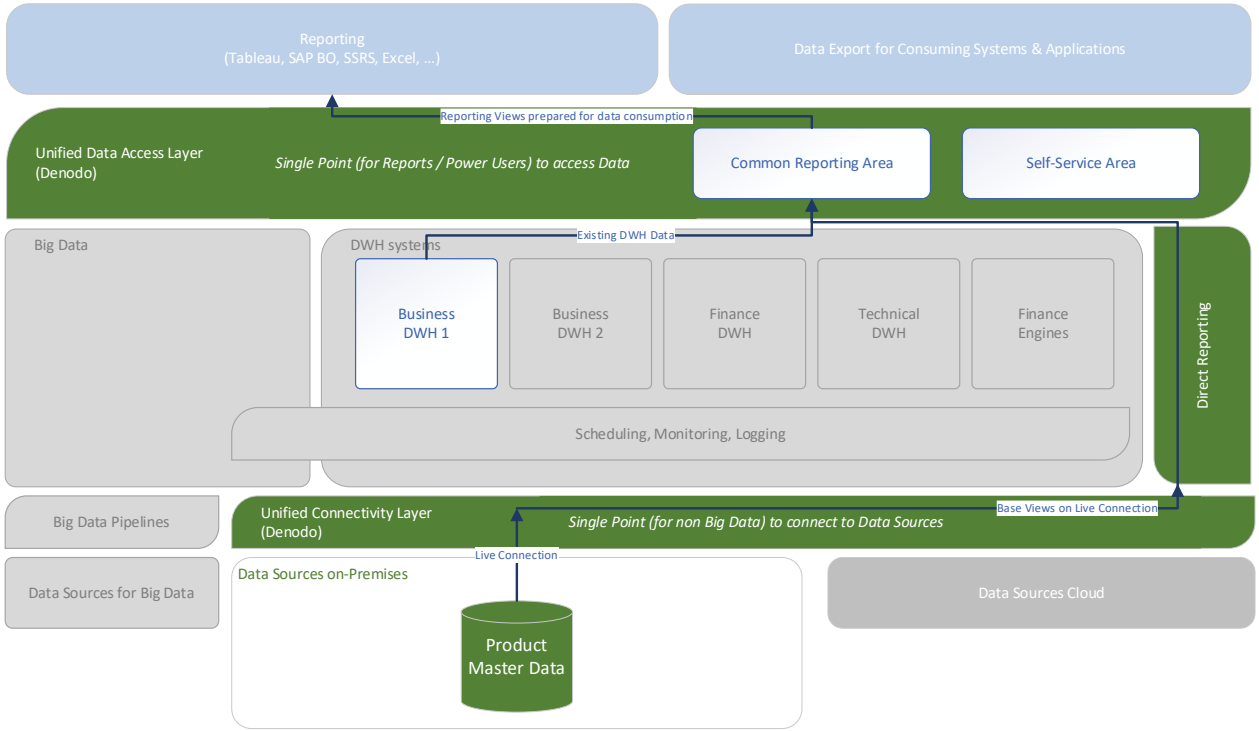


 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

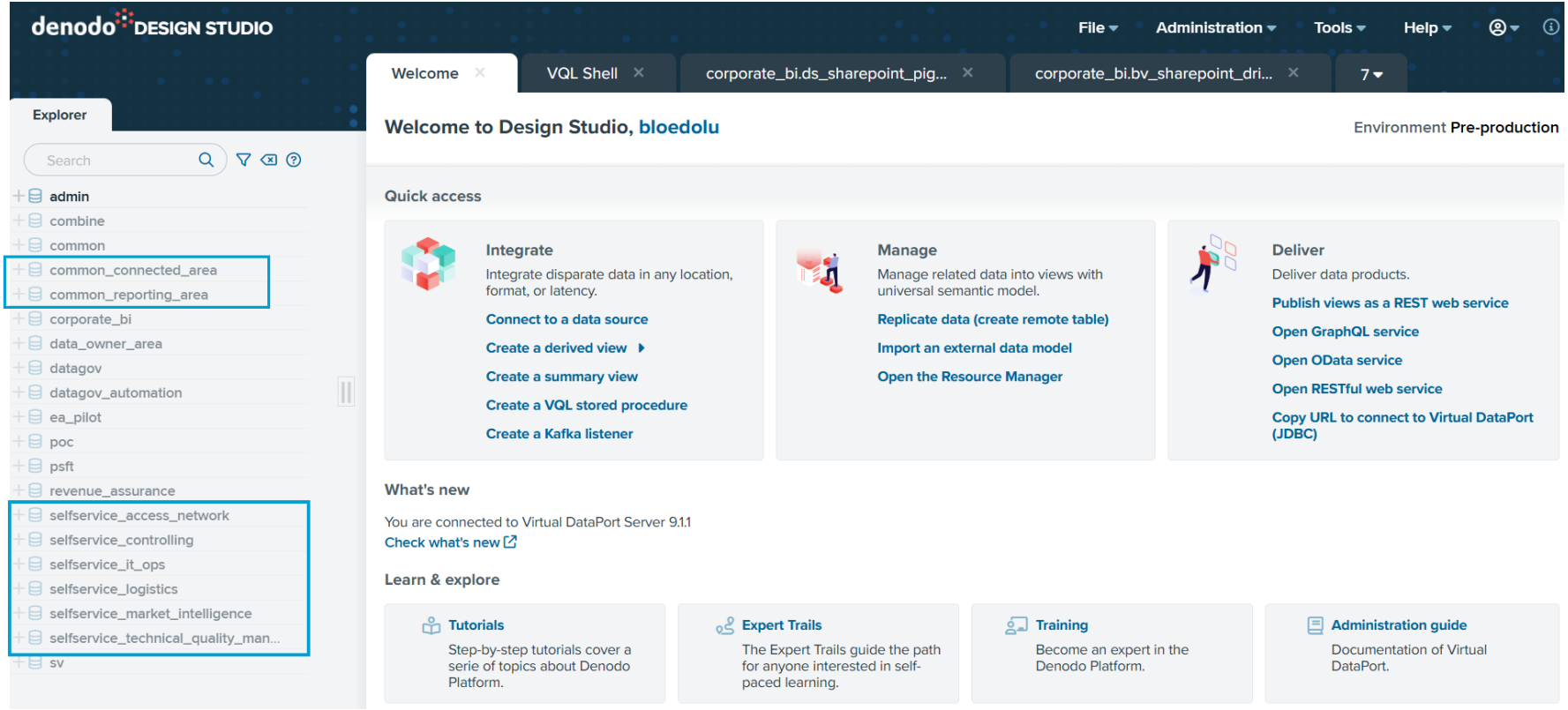
“New tool in our toolbox”

- Mix **DWH** & **Live** Data
- Improve **File Integration**
- Improve **combination of different data sources**
- Use Virtualization on “**both ends**”
- Slowly **remove direct access** to DWH



 © Hutchison Drei Austria GmbH – Confidential

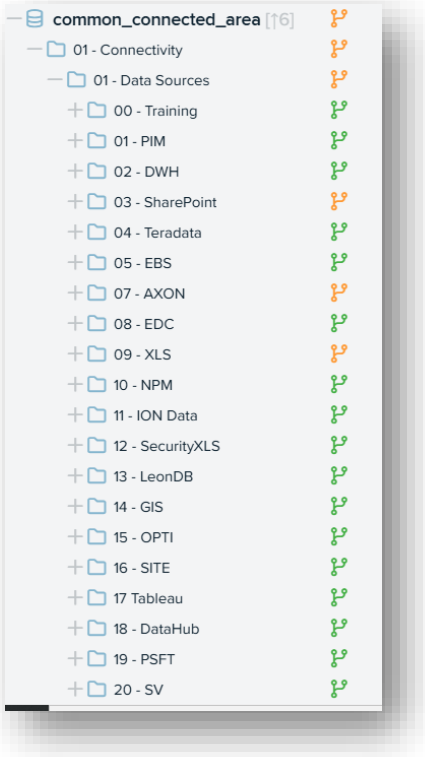
1 - Improve central data & report development



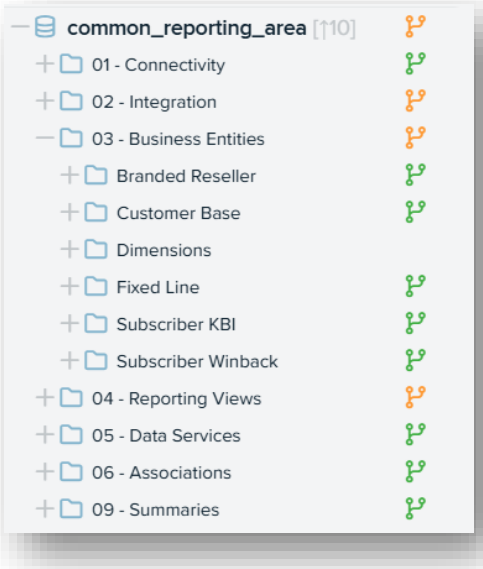
 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

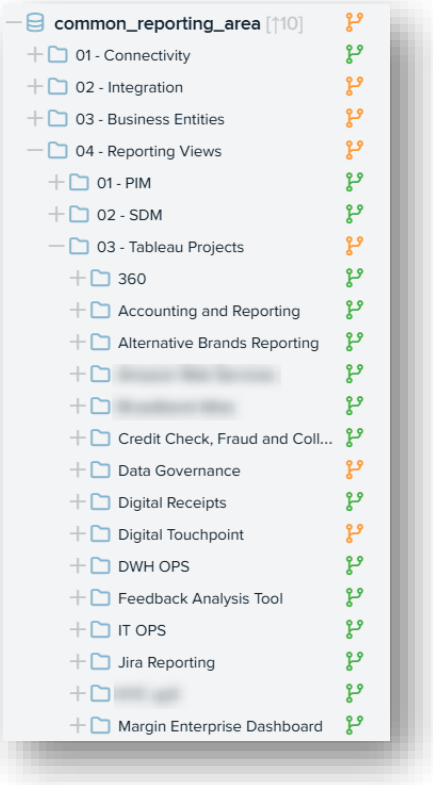
Repository of data sources



Layered data architecture



Last mile transformation



© Hutchison Drei Austria GmbH – Confidential

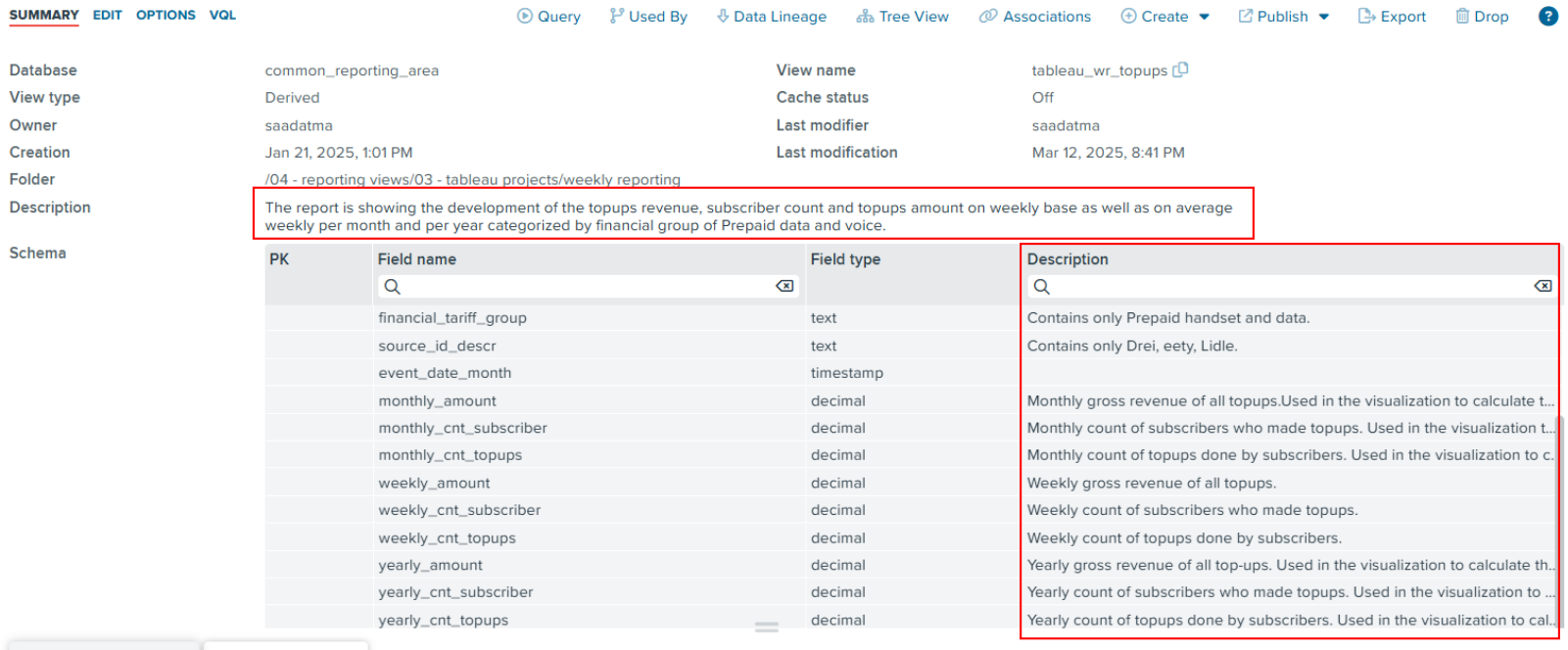
1 - Improve central data & report development

Transparent data lineage & execution trace

© Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

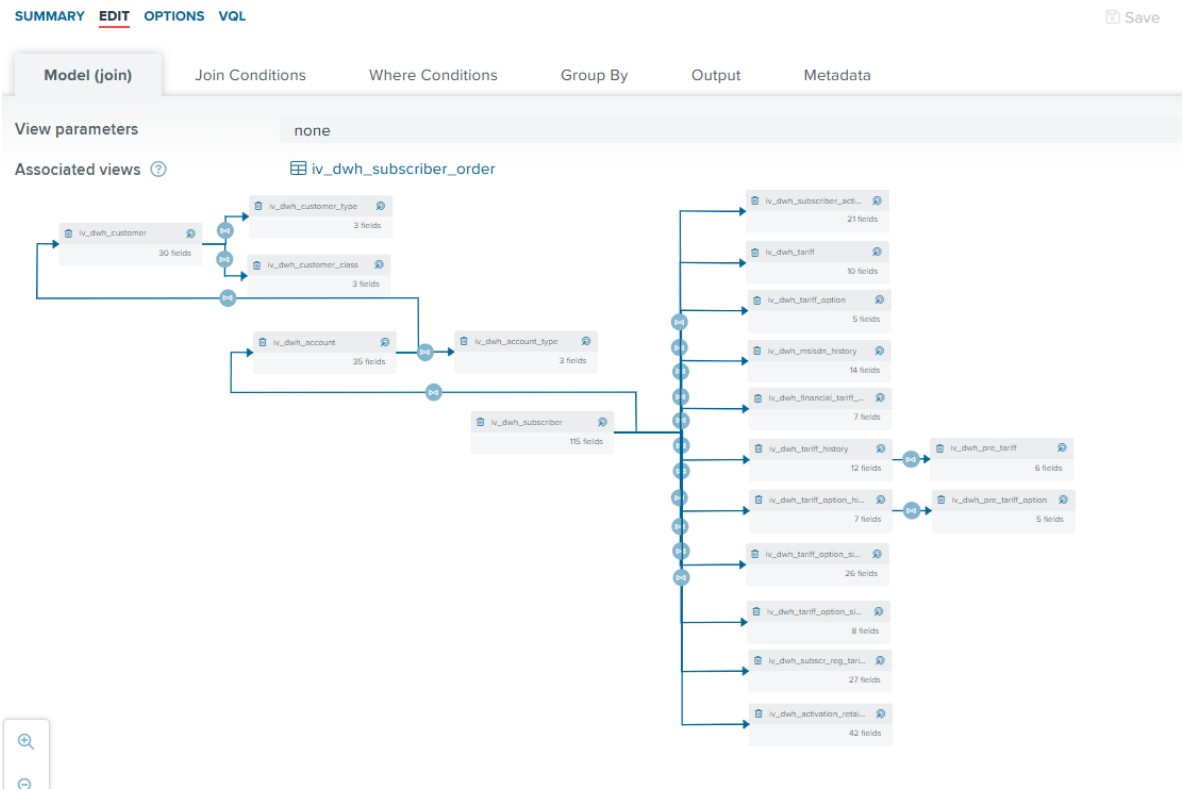
Central (and cascading) definition of object & column **descriptions**



 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

Low-code options for building objects



 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

- SQL-like syntax
- Improve **File Integration**
- Improve **combination of different data sources**

SUMMARY EDIT OPTIONS VQL

VQL options

☐ Export dependencies ☐ Include view statistics

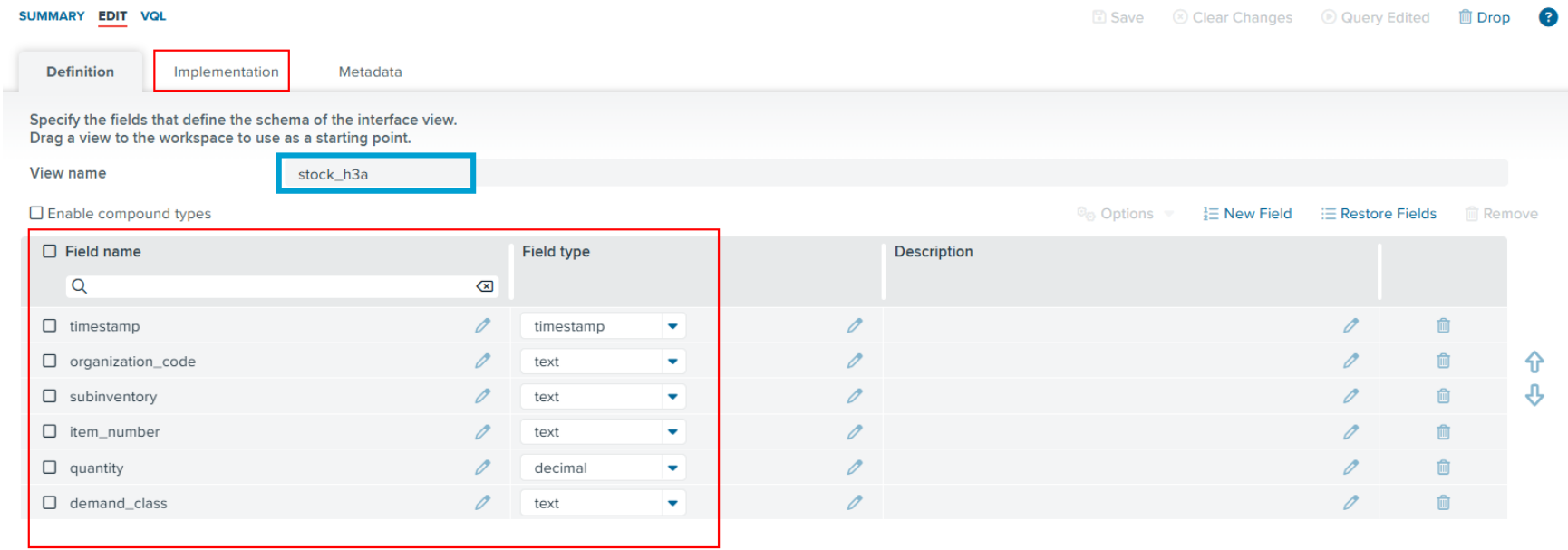
◉ Drop elements before creating them ○ Replace existing elements ○ Do not replace existing elements

```
# Generated with Denodo Platform 9.1.1.

DROP VIEW IF EXISTS tableau_dt_cs_contacts CASCADE;

CREATE VIEW tableau_dt_cs_contacts
    FOLDER = '/04 - reporting views/03 - tableau projects/digital touchpoint'(
        date(description = 'Date when interaction with a customer occured.'),
        item_type(description = 'Type of interaction (call, written item, ...)'),
        channel(description = 'Channel through which the interaction happened (email, letter,...)'),
        customer_class(description = 'Customer class (Consumer, SE, Internal, ...)', sourcetypeid = '12'),
        number_of_contacts(description = 'Number of interactions')
    ) AS
    SELECT common_connected_area.bv_xls_cs_contacts.date AS date,
           CASE WHEN (common_connected_area.bv_xls_cs_contacts.itemtype = 'schriftliche Kundenanfrage') THEN
               'Written Inquiries'
           ELSE
               common_connected_area.bv_xls_cs_contacts.itemtype
           END AS item_type,
           common_connected_area.bv_xls_cs_contacts.kanal AS channel,
           CASE WHEN (common_connected_area.bv_xls_cs_contacts.segment <> 'Consumer') THEN
               upper(common_connected_area.bv_xls_cs_contacts.segment)
           ELSE
               'Consumer'
           END AS customer_class,
           common_connected_area.bv_xls_cs_contacts.cnt AS number_of_contacts
    FROM common_connected_area.bv_xls_cs_contacts
    WHERE date >= trunc(addyear(CURRENT_DATE (), - 2), 'YEAR');

ALTER VIEW tableau_dt_cs_contacts
 LAYOUT (bv_xls_cs_contacts = [199, 80, 200, 200]);
```

 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

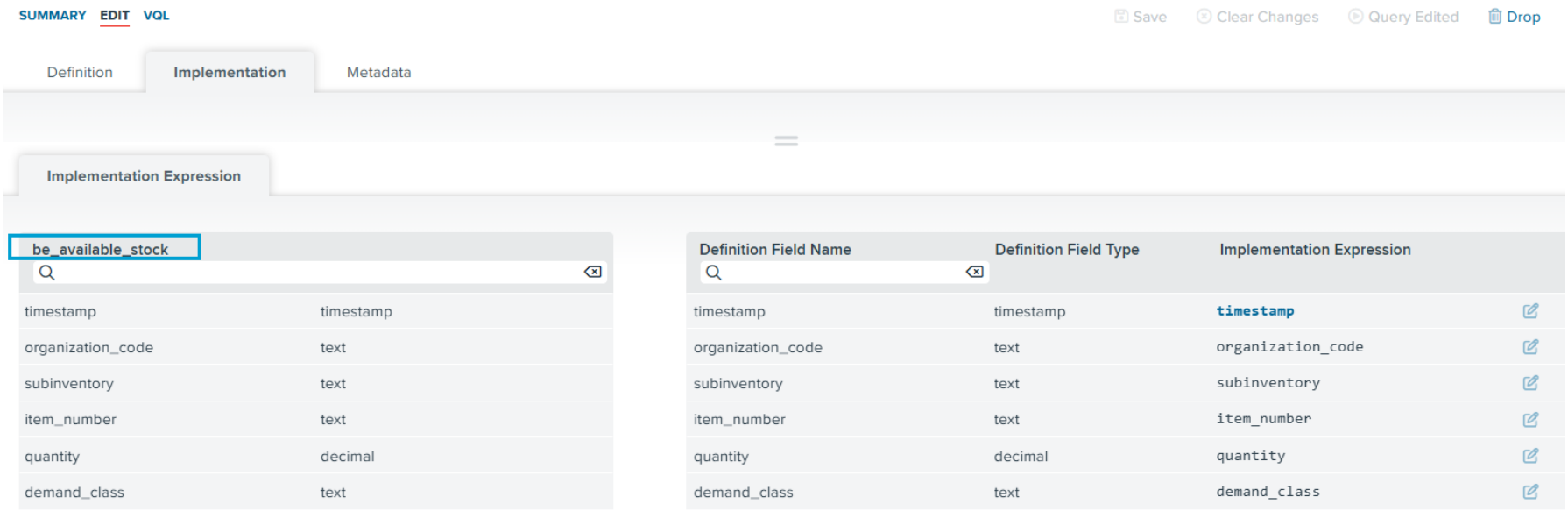
Rapid prototyping with interface views



 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

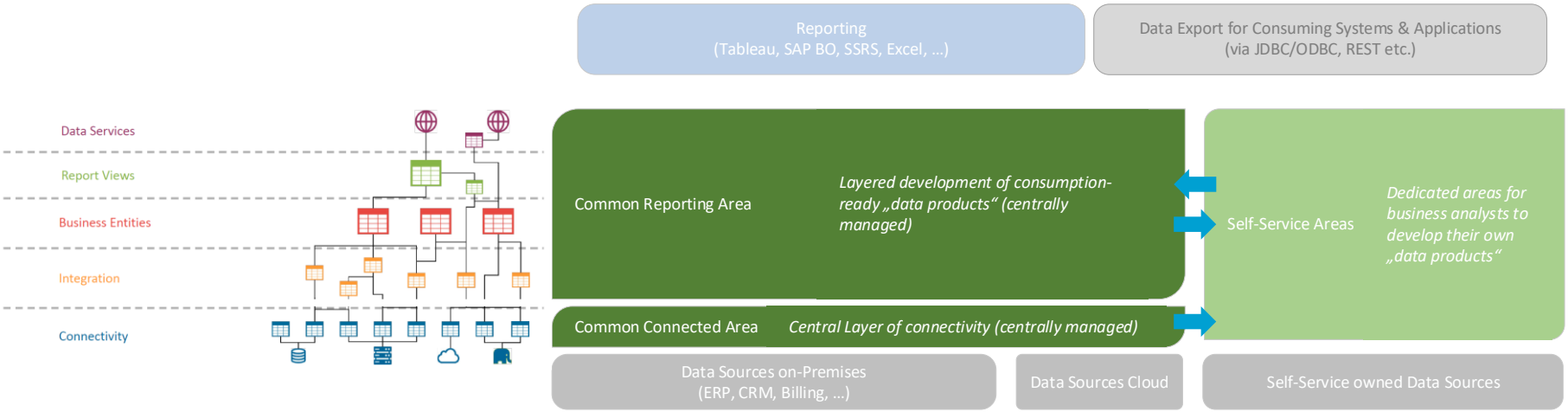
Rapid prototyping with interface views



 © Hutchison Drei Austria GmbH – Confidential

1 - Improve central data & report development

- **Summary** of translating our central & self-service development to the layer architecture



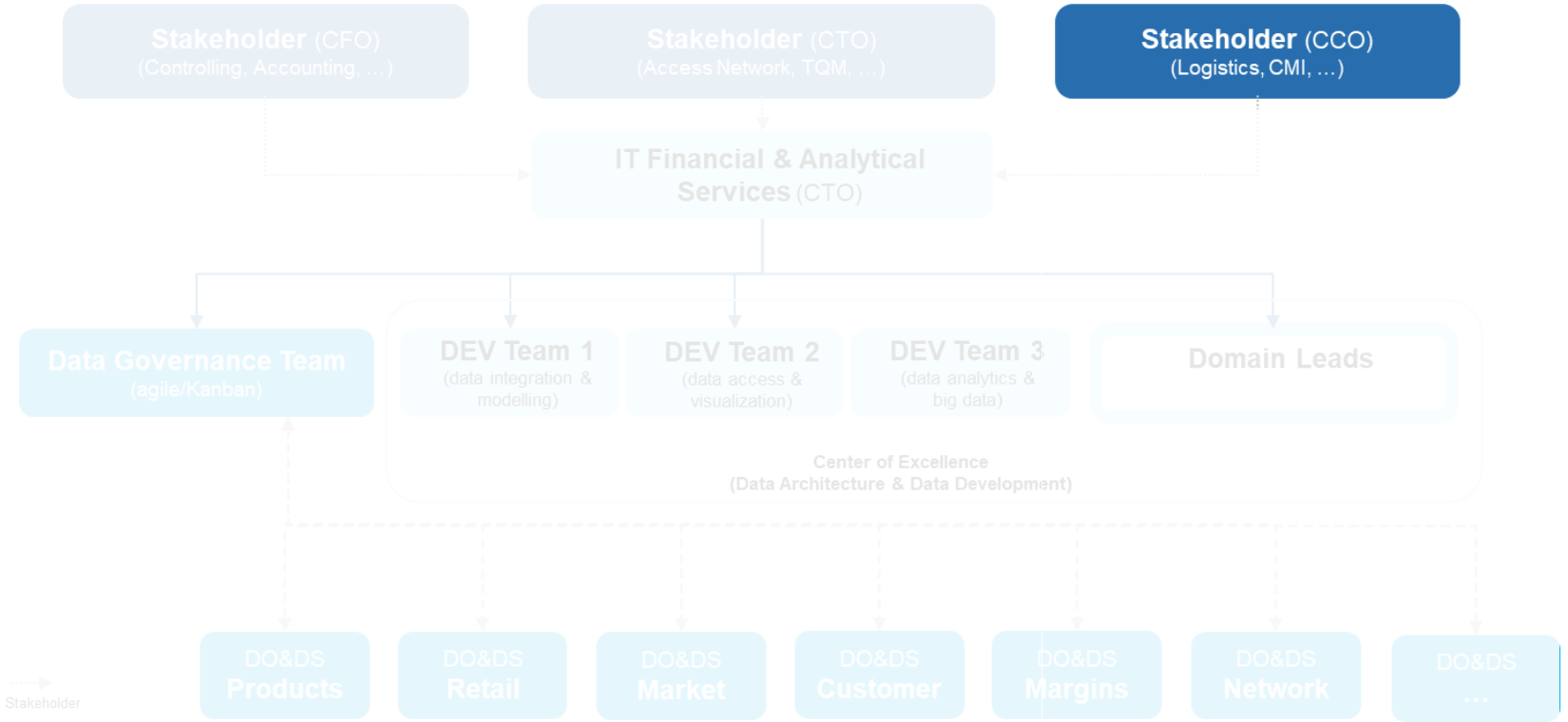
 © Hutchison Drei Austria GmbH – Confidential

2 - Improve Self-Service

- Offer platform to connect to a **central repository of data sources**
  - If infrastructure changes are needed, this is a one-time effort (e.g. Cluster, Database or Network changes)
- Move away from combining sources in reporting tools
  - or exporting data excerpts only to combine it in tools users have “control over” (e.g. Excel, Access etc.)
- Prepare once – use with **whatever interface** most convenient
  - REST, ODBC, JDBC etc.
- **Empower friendly (power) users** to build their own data products
  - While getting best practices & architecture support from a central team

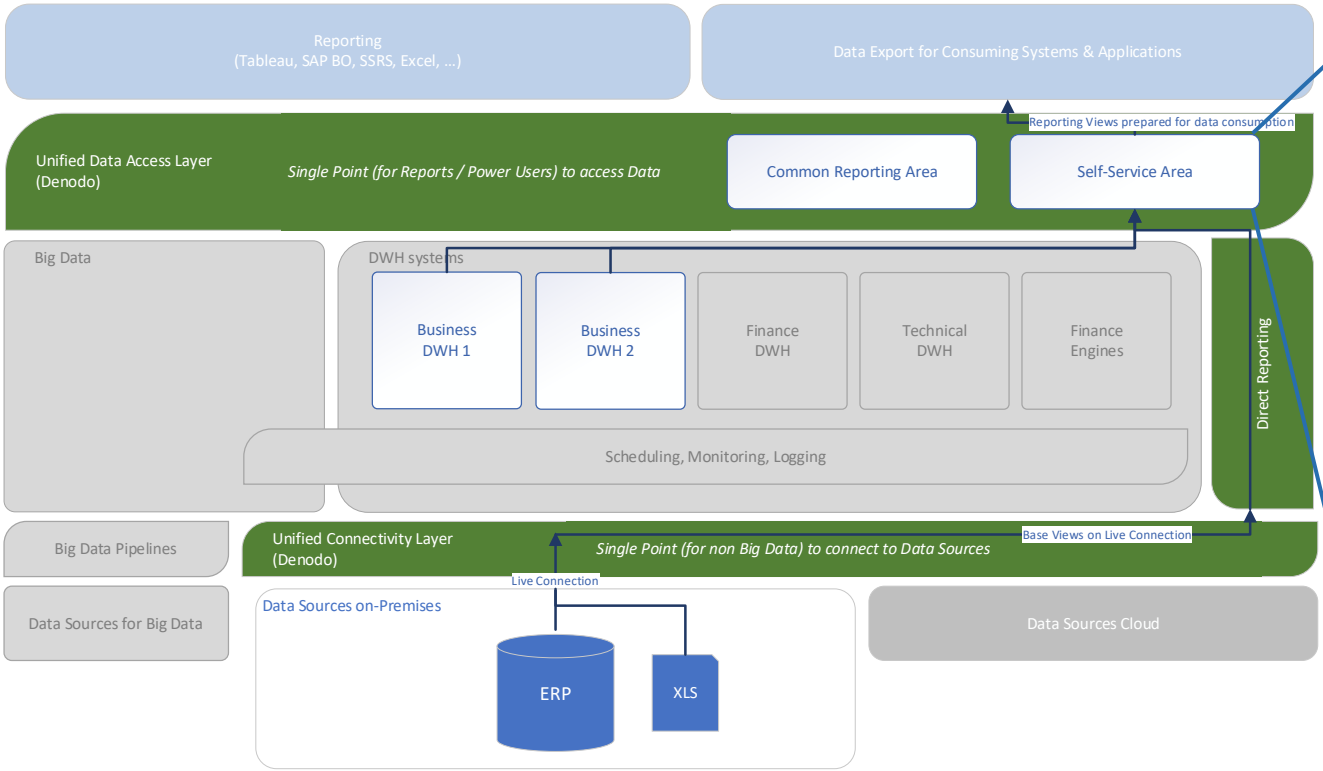
 © Hutchison Drei Austria GmbH – Confidential

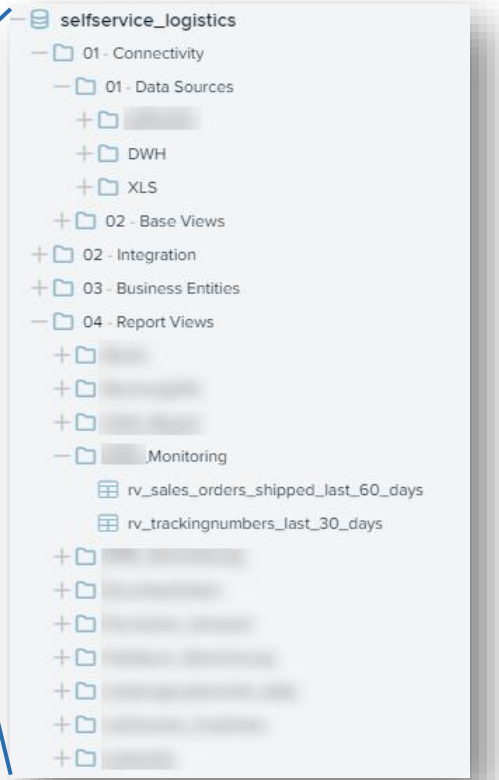
2 - Improve Self-Service



 © Hutchison Drei Austria GmbH – Confidential

2 - Improve Self-Service

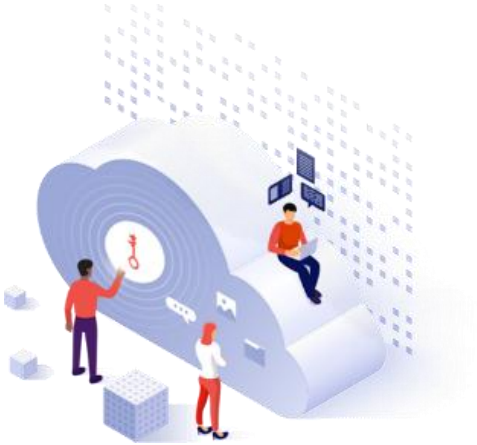




 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

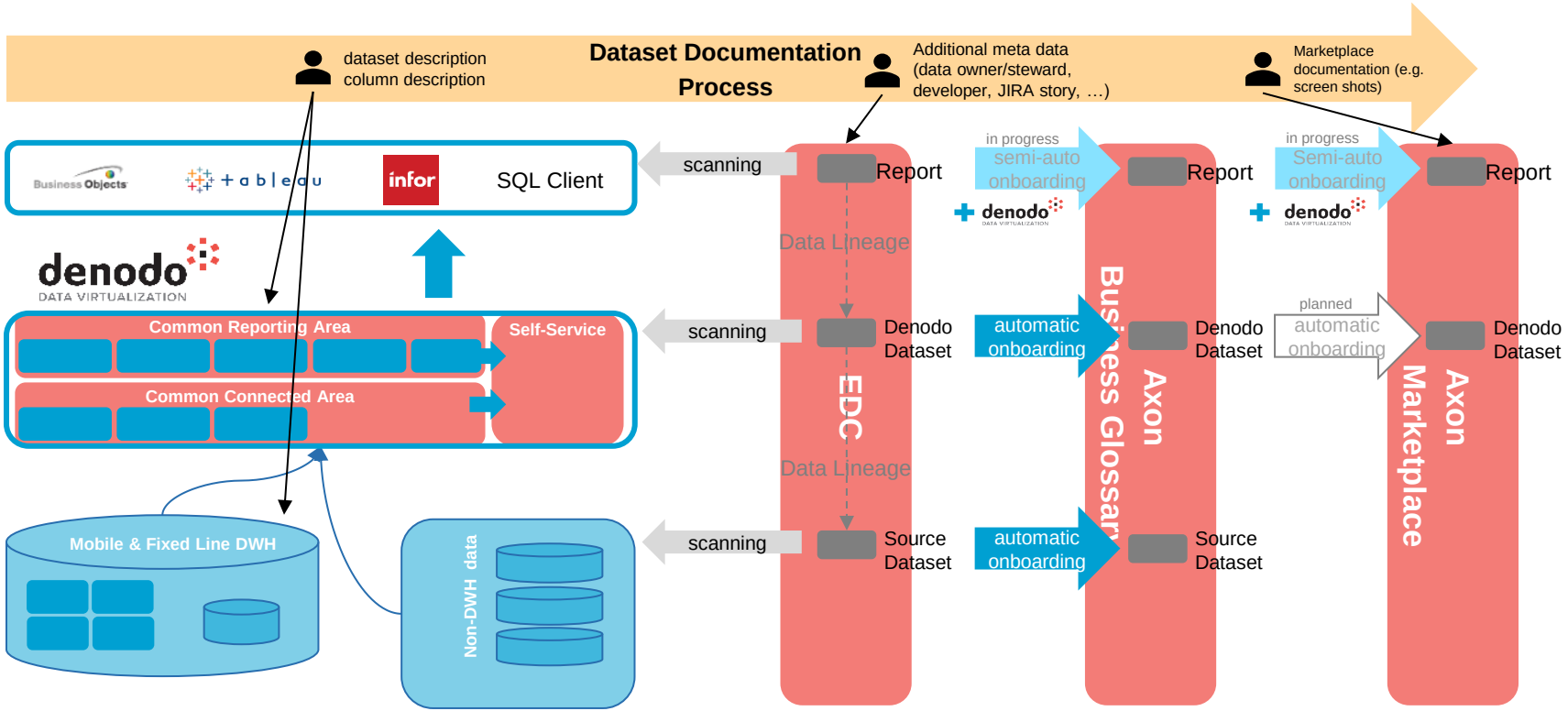
- Slowly **reducing** the amount / need for **shadow reporting**
  - “Just give me the data, I will continue in Excel”
- **Involve** more data analysts in the **data governance processes**
  - Data assignment, Data description, Data access
- Improve **data lineage information** & **metadata** descriptions
  - Including Denodo information in our Enterprise Data Catalog (and vice versa)
  - Making use of inheriting descriptions from source to report views



 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

Data Virtualization and Data Governance **integrated under one architecture**



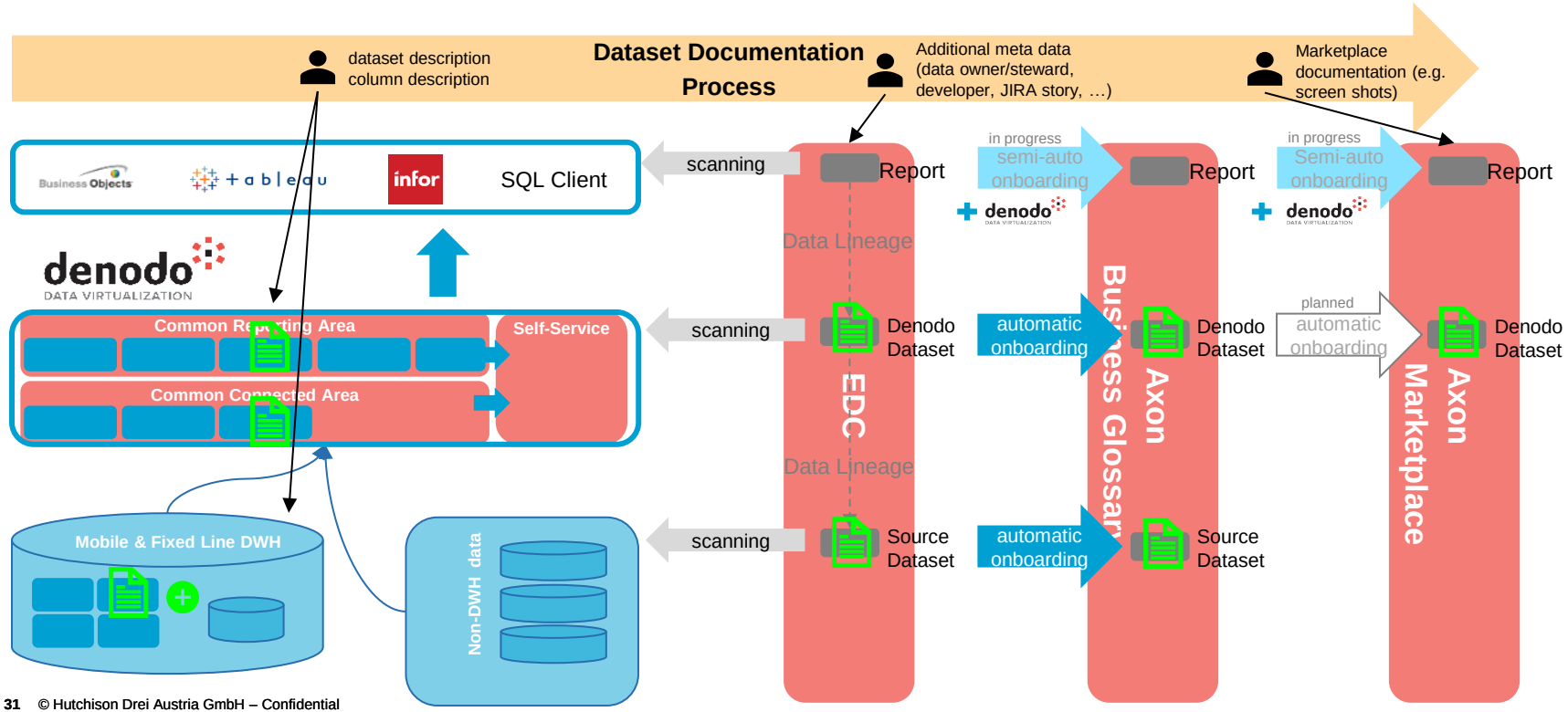
3 - Improve Data Governance

- Documentation of datasets is crucial to our Data Governance guidelines (knowledge management)
- No dataset leaves the “factory” without being documented (dataset & column description)
- **Challenge:**
  - Documentation to be available not only in source system, but wherever and from whomever (meta) data are consumed (DWH, Denodo, Enterprise Data Catalog, Business Glossary, Data Marketplace)
  - How to ensure documentation of datasets coming from source system not obliged for Data Governance standards?
  - How to ensure non-DWH developers in the business areas to document their

© Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

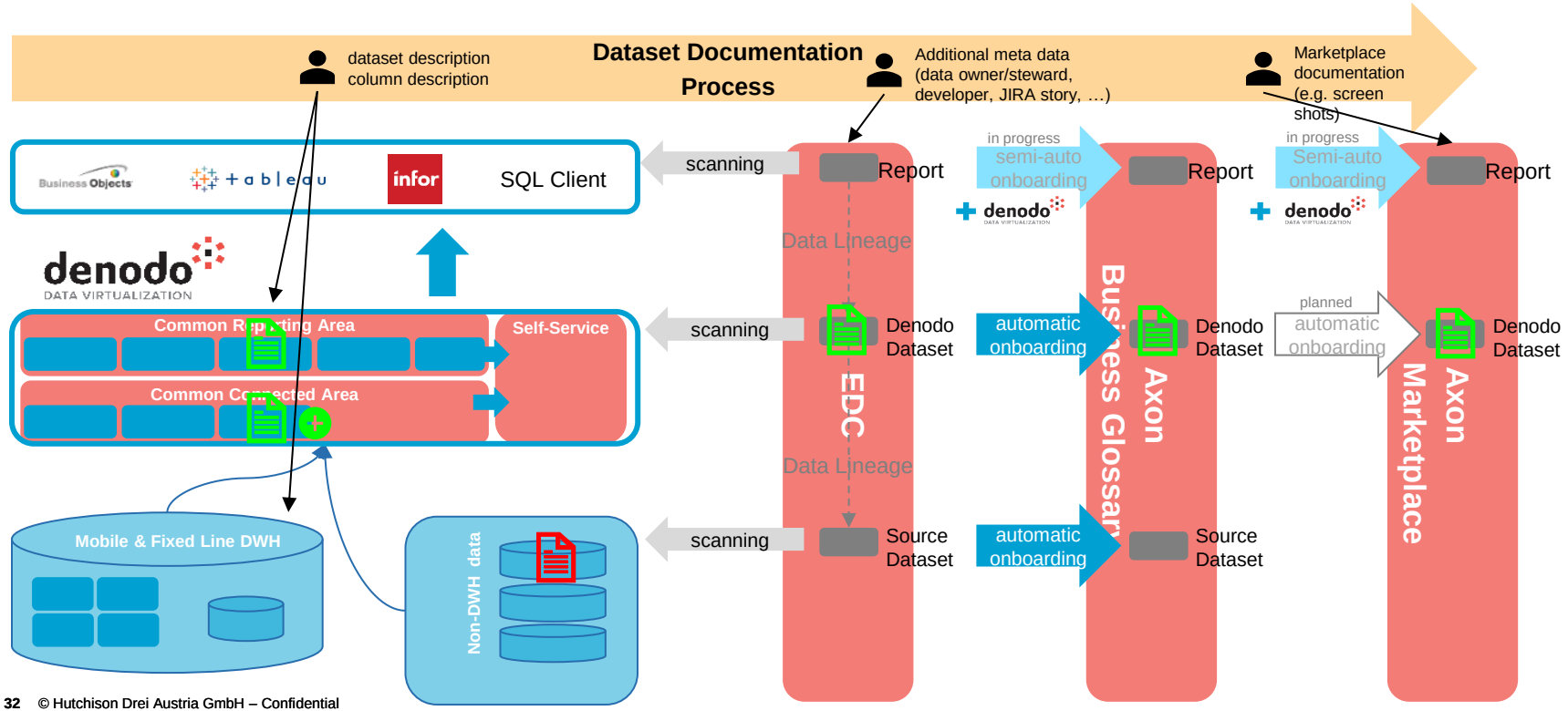
Dataset and column descriptions **automatically available on all levels and applications.**



 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

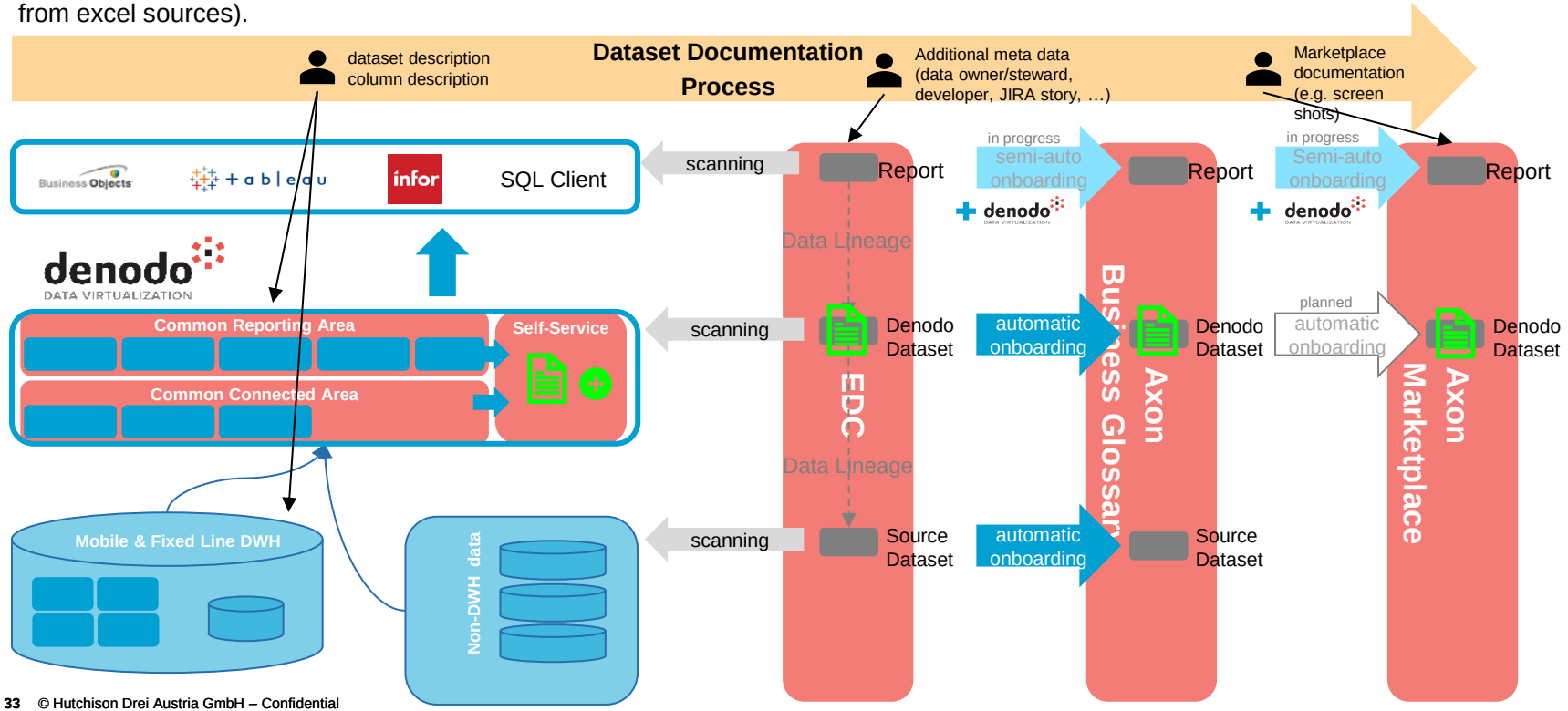
Denodo provides the possibility to **document datasets from sources not included under Data Governance guidelines**.



 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

Denodo enables also **non-DWH developers in Self-Service area** to document their **dataset creations** and even **manual data inputs** (e.g. from excel sources).



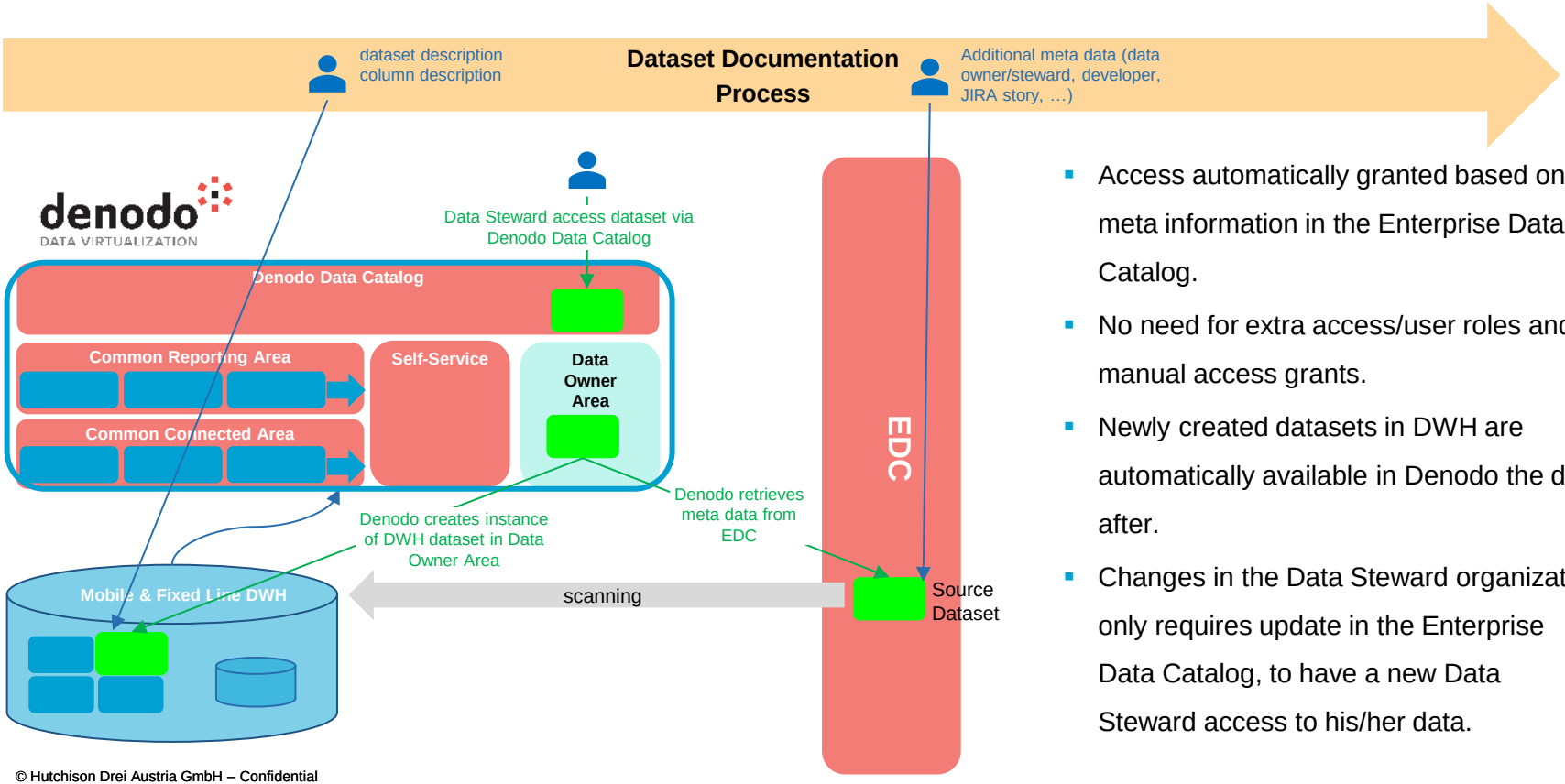
 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

- Every data steward obviously needs access to the data they are responsible for (= subject area)
- **Challenge:**
  - 22 subject areas, 17 data owner, 17 data stewards, 22 data stewards
  - > 2,000 reporting views currently assigned to subject areas
  - Goal: Highly automated/low maintenance solution
  - User friendly possibility to consume data, even for employees who are not heavy data consumer

 © Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

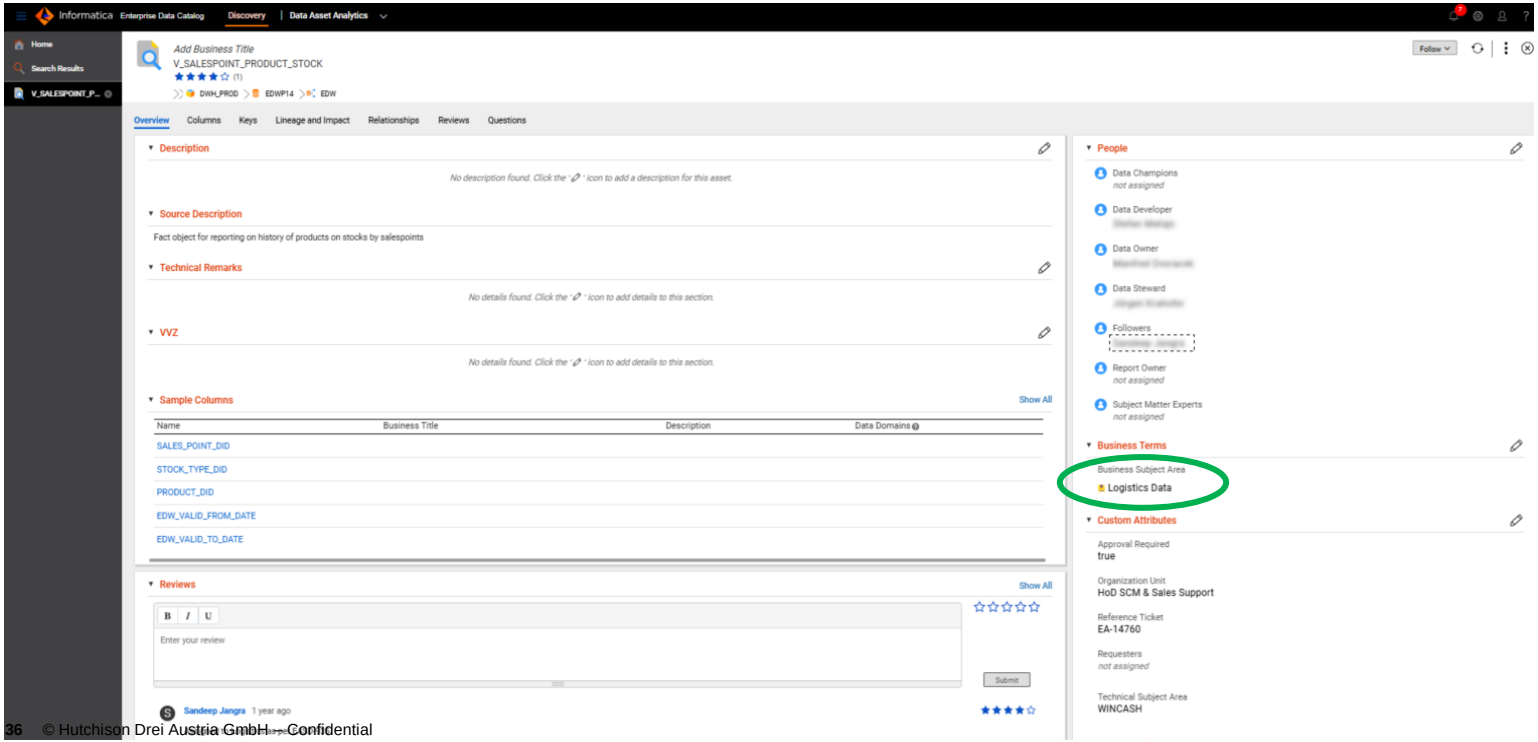


- Access automatically granted based on meta information in the Enterprise Data Catalog.
- No need for extra access/user roles and manual access grants.
- Newly created datasets in DWH are automatically available in Denodo the day after.
- Changes in the Data Steward organization only requires update in the Enterprise Data Catalog, to have a new Data Steward access to his/her data.

© Hutchison Drei Austria GmbH – Confidential

3 - Improve Data Governance

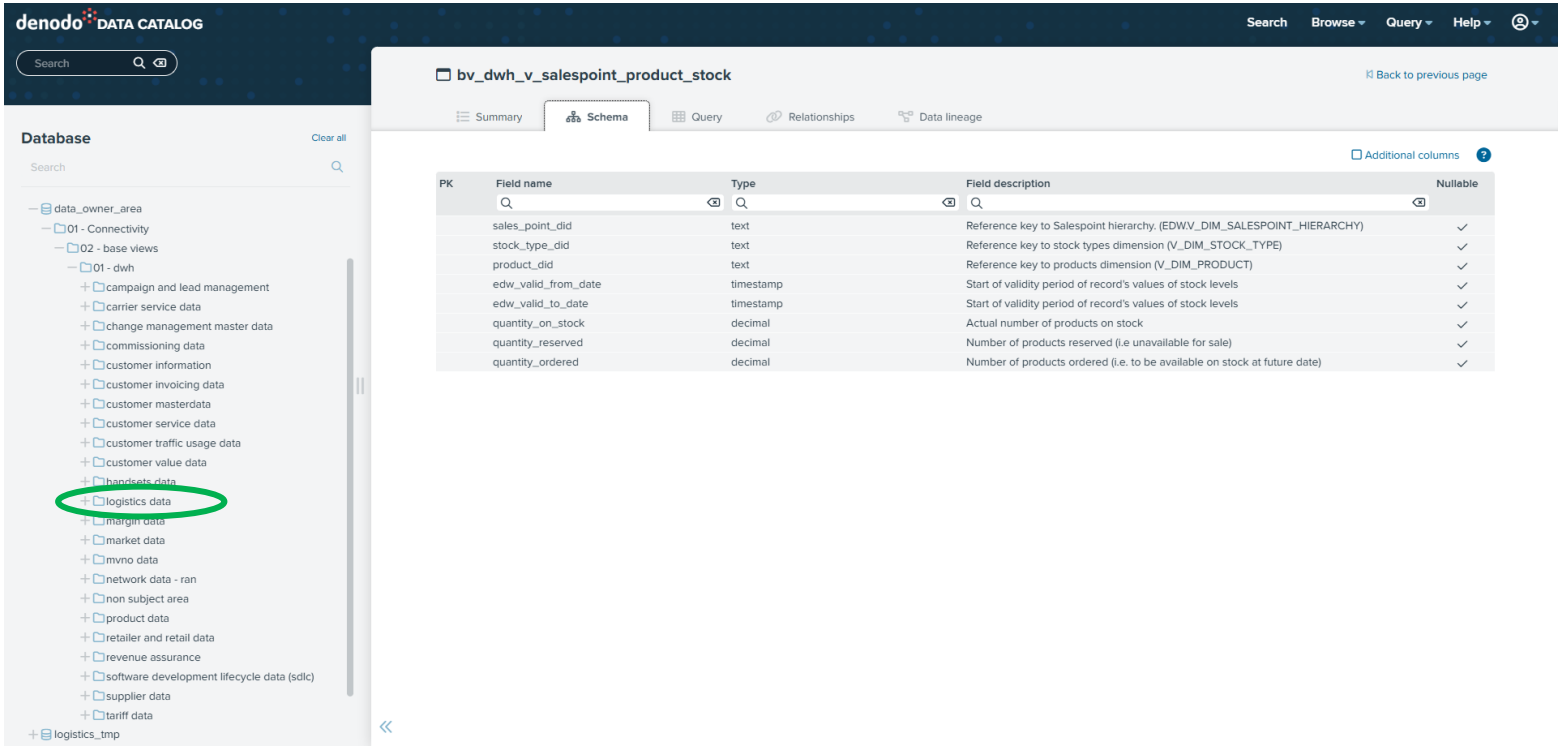
Assignment of **data owner & data steward** (= ownership/subject area) to a data set



 © Hutchison Drei Austria GmbH – Confidential

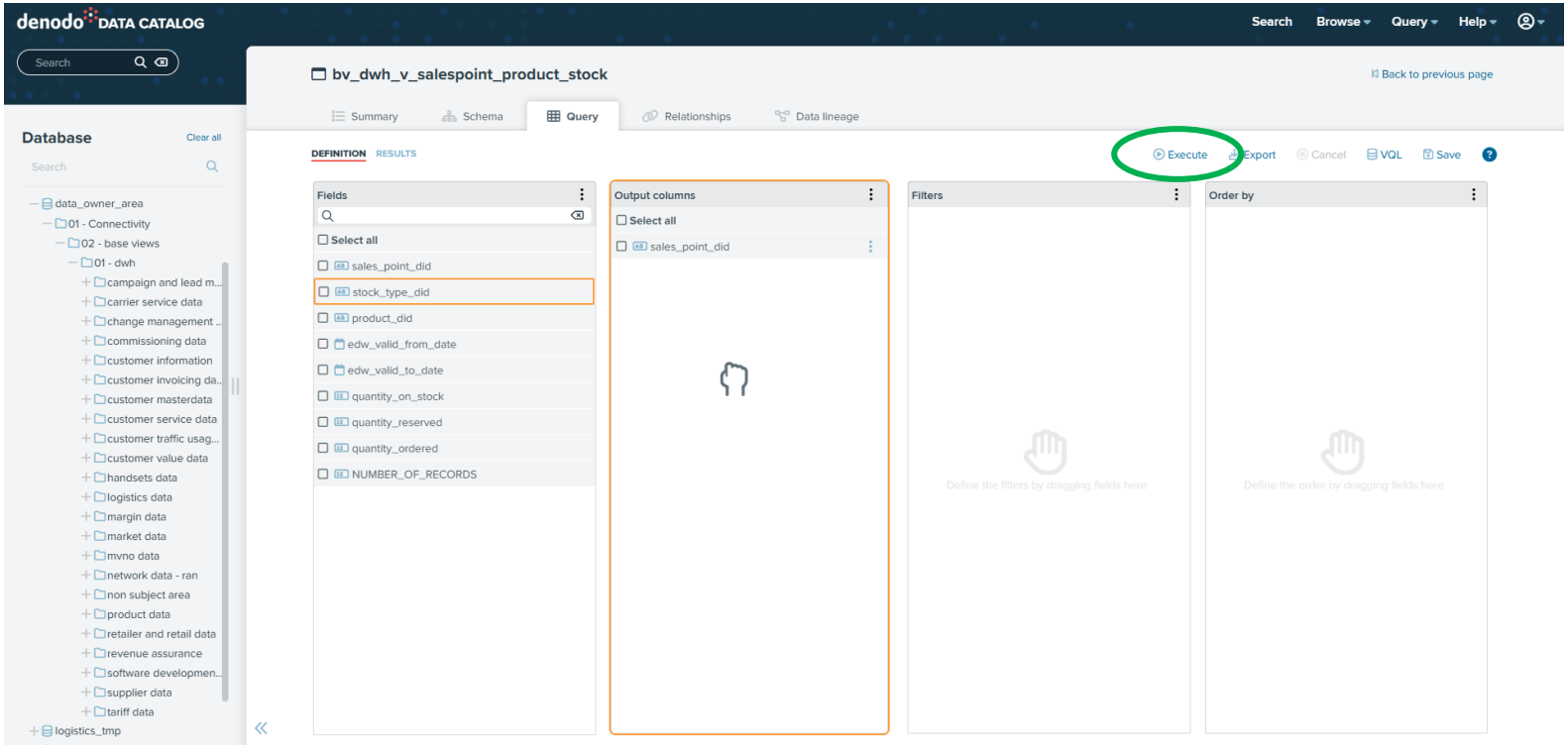
3 - Improve Data Governance

Data Steward's **access automatically granted** in Denodo to newly created datasets in DWH



3 - Improve Data Governance

User friendly consumption of data for Data Stewards in the Denodo Data Catalog



DATA CONSUMERS

ETL Developer | Data Engineer | Citizen Integrator | Data Scientist | Data Analyst | Business Users

Intelligent Data Management Cloud™

- DISCOVER & UNDERSTAND — DATA CATALOG
- ACCESS & INTEGRATE — DATA INTEGRATION & ENGINEERING
- CONNECT & AUTOMATE — API & APP INTEGRATION
- CLEANSE & TRUST — DATA QUALITY & OBSERVABILITY
- MASTER & RELATE — MDM & 360 APPLICATIONS
- GOVERN & PROTECT — GOVERNANCE, ACCESS & PRIVACY
- SHARE & DEMOCRATIZE — DATA MARKETPLACE

CLAIRE®
AI-Powered Metadata Intelligence & Automation

Connectivity
Metadata System of Record

DATA SOURCES

SaaS Apps Sources + Mainframe | Applications | Databases | On-premises Sources + IoT | Machine Data | Logs | Real-time / Streaming Sources

Intelligent Data Management Cloud™

Data Governance & Catalog

Understand and Govern

Govern both data and AI models to help ensure trust

Data Marketplace

Share and Democratize

Package data assets & AI models for consumption

Data Quality & Observability

Cleanse and Trust

Identify, resolve anomalies & issues to grow data pipelines

Data Access Management

Protect and Comply

Secure and share data responsibly for proper use

ONE Metadata Repository

© Informatica. Proprietary and Confidential.

Data Governance

CAPABILITIES

Business-focused Information Consumption and Management

- Intuitive Relationship Visualizations of governance artifacts and the associated business impact to effectively manage and operate governance programs
- Multi-Language for Global rollouts with support for 6 languages
- Customizable and Personalized Dashboards
- Customizable experience by persona
- Multi-hop graph-based searches to browse and navigate complex relationships

Collaboration

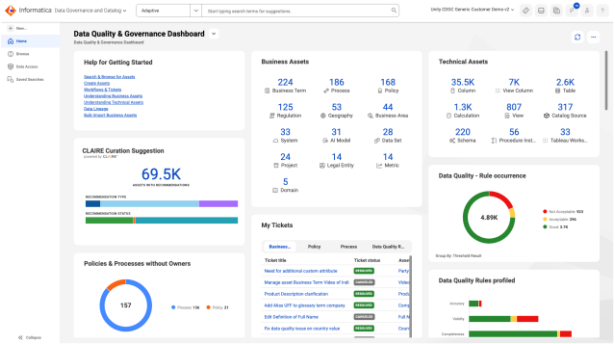
- Crowd-sourced knowledge for continuous feedback, ratings, certifications
- Governance and Technical Asset Ratings to increase trust and confidence
- Discussion Threads for informal interactions/clarifications
- Workflows and Ticket management for structured interactions

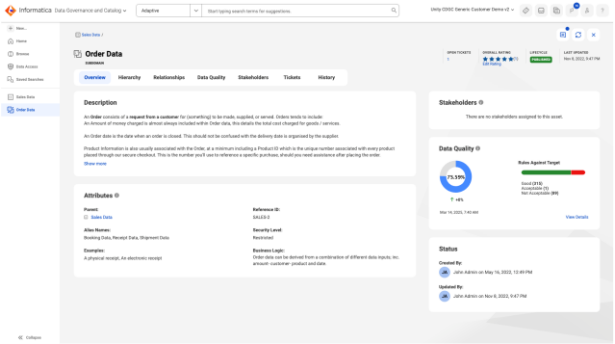
Automation and Integrations to accelerate business insights

- Quality overlays for Business and Technical Lineage to improve understanding and accelerate Quality Coverage
- Data Quality Self Service using NLP-based Rule generation

CLAIRE ®

- Intelligent Glossary Associations to effectively enrich scanned assets with business context
- Centralized dashboard to review and manage CLAIRE generated Glossary Associations and Classifications
- Data Quality Automation to assess sources against enterprise standards





Data Catalog

CAPABILITIES

Advanced Scanners and Curation

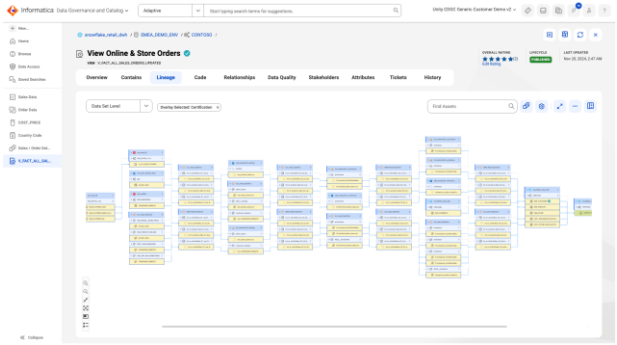
- Broadest metadata coverage across all key ecosystems: Snowflake, Microsoft, AWS, Google, SAP, Oracle and Salesforce/Tableau and more
- Advanced Scanning capabilities include extracting data lineage from Stored Procedure Scanning, Python Code Scanning, and Embedded SQL in Data Pipelines
- Bring your own metadata with extendable Custom Scanner Framework
- Bulk Curation to eliminate human task and accelerate data labeling and glossary association outcomes
- High-granularity Sensitive Data Classification for visibility, tracking and reporting

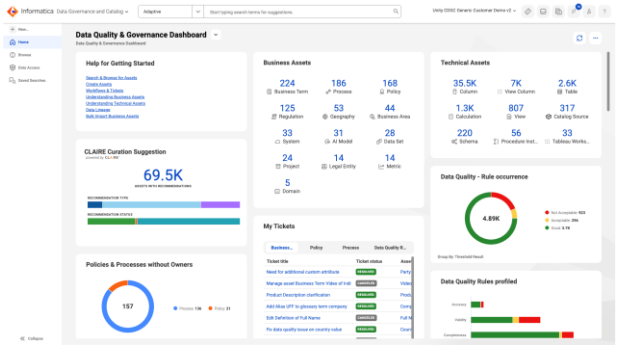
CLAIRE®

- 225+ Pre-built based Data Elements like national IDs, dates, locations, names
- Classification builder to create custom classifications
- CLAIRE generated classifications
- Aggregate elements into entities to classify and discover business entities

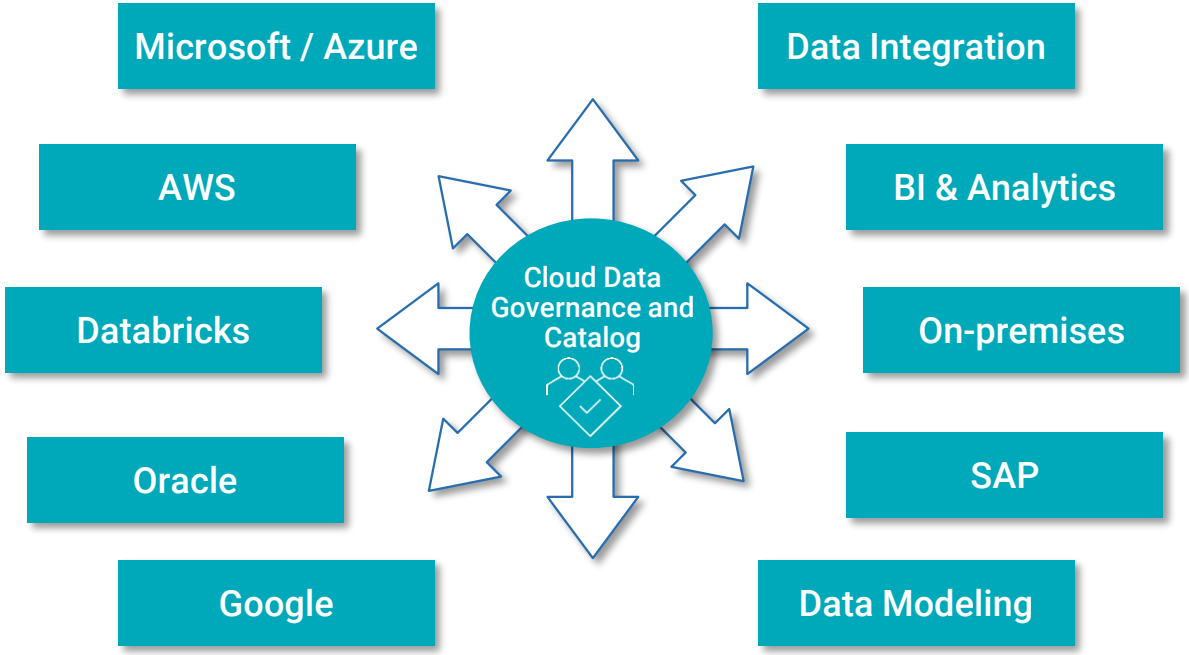
Lineage

- Automated Column-Level Lineage. views with business overlays including Data Quality, Policies, Glossaries, and Processes
- Advanced lineage scanning including parsing transformations hidden in SQL code, ETL/ELT code, Python Code, BI Tools and more
- Detailed Impact Analysis reports for both upstream and downstream changes





Broad & Deep Metadata Connectivity



Advanced capabilities

- Automated extraction of granular metadata on Table and Column level from complex enterprise systems
- Advanced parsing of source code from multi-vendor sources
- Dynamic code support and proper parameters values tracking based on custom catalog tables and operational metadata for selected catalog sources
- Incremental metadata scan support for selected catalog sources
- Metadata extraction from other catalogs
- Serverless metadata scans for selected sources
- Automated metadata synchronization from selected IDMC services with the catalog
- Rule-based and automated inferred lineage

Metadata Scanners Coverage

Area	Support
AWS	AWS S3 | Redshift | AWS RDS (Oracle, MS SQL Server, PostgreSQL and MySQL) | DynamoDB+ | AWS Glue | Athena
Microsoft	Azure SQL DB | Azure Synapse | Azure ADLS Gen 2 | Azure Blob | Azure Data Factory | OneDrive | SharePoint | Cosmos DB+ | PowerBI | SSAS | SSRS | Fabric Data Warehouse | Fabric Lakehouse | Fabric OneLake
Google	Google BigQuery | Google Cloud Storage | Google Looker
Oracle	Oracle ADB | Oracle Cloud Object Storage | Oracle Golden Gate | OBIEE | Oracle Data Integrator
Catalogs	Microsoft Purview | Apache Atlas | Databricks Unity Catalog | AWS Glue
DW	Snowflake DW | Databricks Delta Lake | Apache Hive/HDFS | SingleStore+
On-Premises	Oracle | Microsoft SQL Server | IBM DB2 z/OS | IBM DB2 LUW | MySQL | Teradata | PostgreSQL | JDBC | IBM Netezza+ | MongoDB | Kafka | Local/Shared Filesystem | SFTP | Windows Share/CIFS | Sybase ASE+ | Sybase IQ+ | MariaDB | MySQL | Apache Cassandra+
SAP	SAP BW | SAP BW/4HANA | SAP ECC | SAP S/4HANA | SAP HANA DB | SAP Analytics Cloud | SAP Datasphere | SAP SuccessFactors
Applications	Salesforce | Marketo+ | Dynamics CRM+ | Workday+ | Informatica MDM B360 | Informatica R360 | NetSuite+
BI & Analytics	Tableau | Microsoft PowerBI | Microsoft SSRS | Qlik View | Qlik Sense | Qlik Sense Cloud | Databricks Notebooks | IBM Cognos | Google Looker | Microsoft SSAS | Tibco Spotfire | MicroStrategy | SAP BO | OBIEE | Microsoft Azure Analysis Services | Tableau Cloud
Data Integration	Informatica Cloud Data Integration | Informatica PowerCenter | Microsoft SSIS | Talend DI | Microsoft Azure Data Factory | IBM DataStage | dbt | Microsoft Azure Synapse Analytics | Informatica Developer
Data Modelling Tools	erwin Data Modeler | SAP Power Designer | IDERA ER/Studio Data Architect
Accelerators	SAS Base Libraries (Accelerator) | SAS Base Programs (Accelerator) | IBM Mainframe JCL (Accelerator)

© Informatica. Proprietary and Confidential.

+ Supported through JDBC

IDMC Advantages: Multiple drivers for added value…

Designed for Modern Data Architectures

- Data Fabric
- Data Mesh
- Role-based UX

CLAIRE®-powered Data Management

- AI-Automated
- Copilot
- GPT Search

Scale Up Your Digital Business

- Consumption / IPUs
- Volume/Performance
- Integrated Tools

Cost Efficient, Unified Approach

- AI-enabled performance
- No CapEx
- Low OpEx

Empower Data Consumers

- Data Democratization
- Access Management
- Last-Mile Delivery

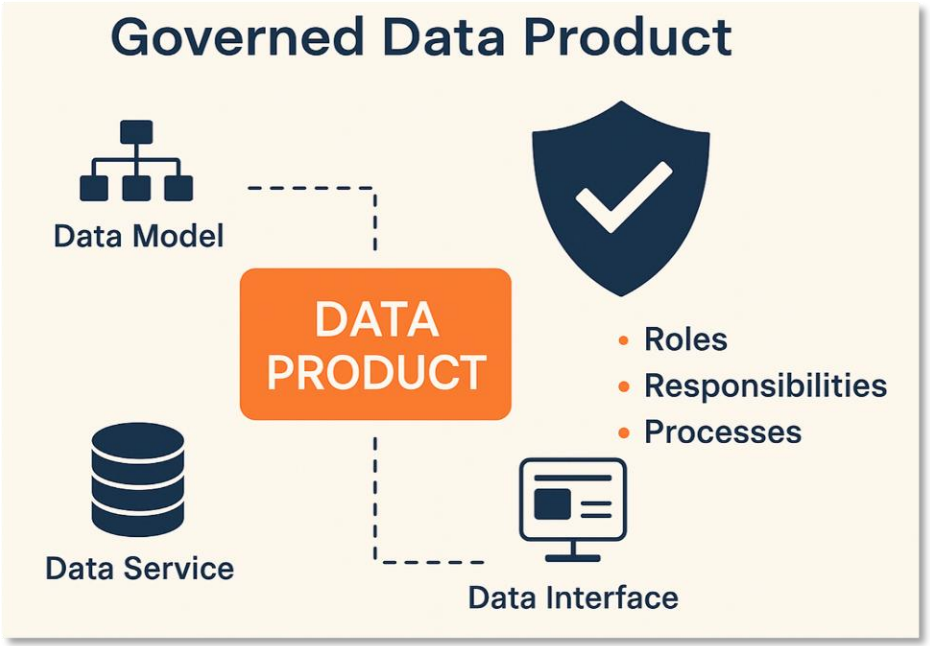
© Informatica. Proprietary and Confidential.

Things for the future

 © Hutchison Drei Austria GmbH – Confidential

Data Product principles

- A dashboard is not a Data Product
- Data Governance applies on all elements
- A data management plattform is the foundation for data products
- No Data Products without a managed product lifecycle



© Hutchison Drei Austria GmbH – Confidential

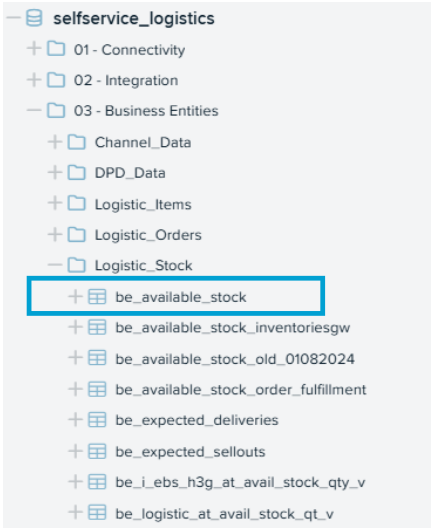
Data Product Concept (End-to-End)

- Step 1: Definition
  - Using Data Product Canvas

 © Hutchison Drei Austria GmbH – Confidential

Data Product Concept (End-to-End)

- Step 2: Technical Build-Up (decentrally and in responsibility of Data Product Team)

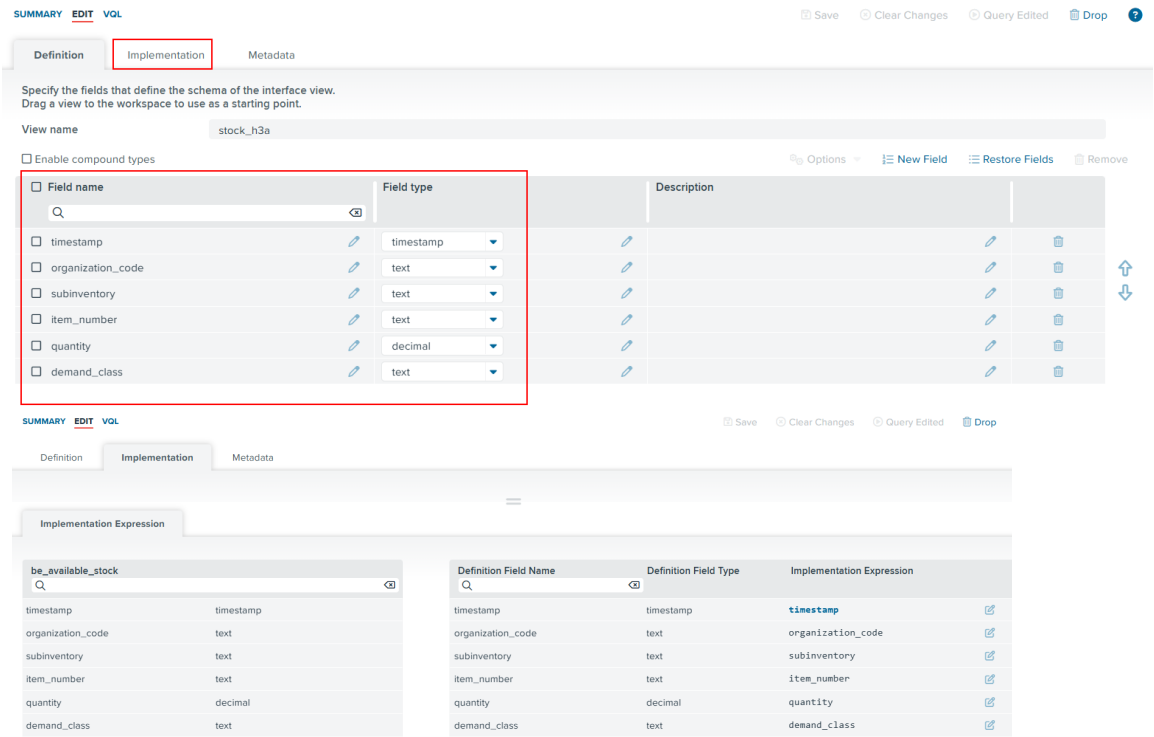


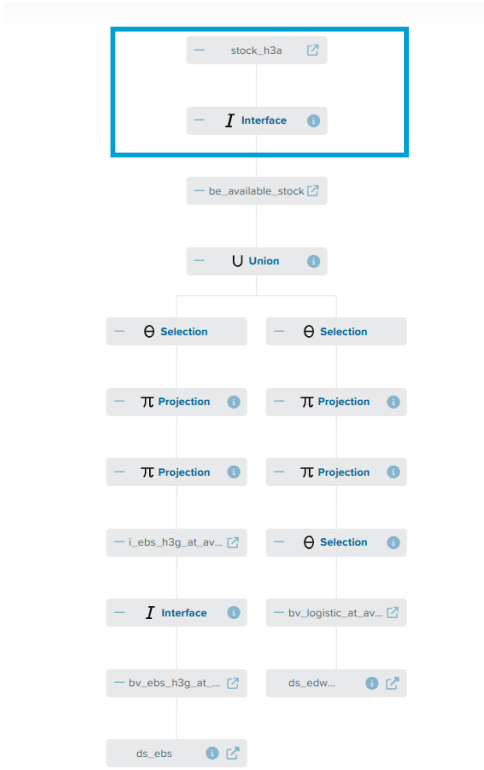
```
# Generated with Denodo Platform 9.1.1.

DROP VIEW IF EXISTS be_available_stock CASCADE;

CREATE VIEW be_available_stock
    FOLDER = '/03 - business entities/logistic_stock' AS
    WITH live_stock AS (
            SELECT creation_date AS timestamp,
                   organization_code,
                   subinventory_code AS subinventory,
                   item AS item_number,
                   quantity,
                   quantity_type AS demand_class
            FROM i_ebs_h3g_at_avail_stock_qty_v
        ),
        dwh_stock AS (
            SELECT date_id AS timestamp,
                   organization_code,
                   subinventory,
                   item AS item_number,
                   quantity,
                   quantity_type AS demand_class
            FROM bv_logistic_at_avail_stock_qt_v
            WHERE (
                    date_id = (SELECT max(date_id) AS max FROM bv_logistic_at_avail_stock_qt_v)
                    AND subinventory not in('FERTIG', 'DOM-FERTIG')
                )
        )
    SELECT *
    FROM live_stock
    SQL UNION ALL
    SELECT *
    FROM dwh_stock;
```


© Hutchison Drei Austria GmbH – Confidential

Data Product Concept (End-to-End)

- Step 3: Interface Definition (centrally)

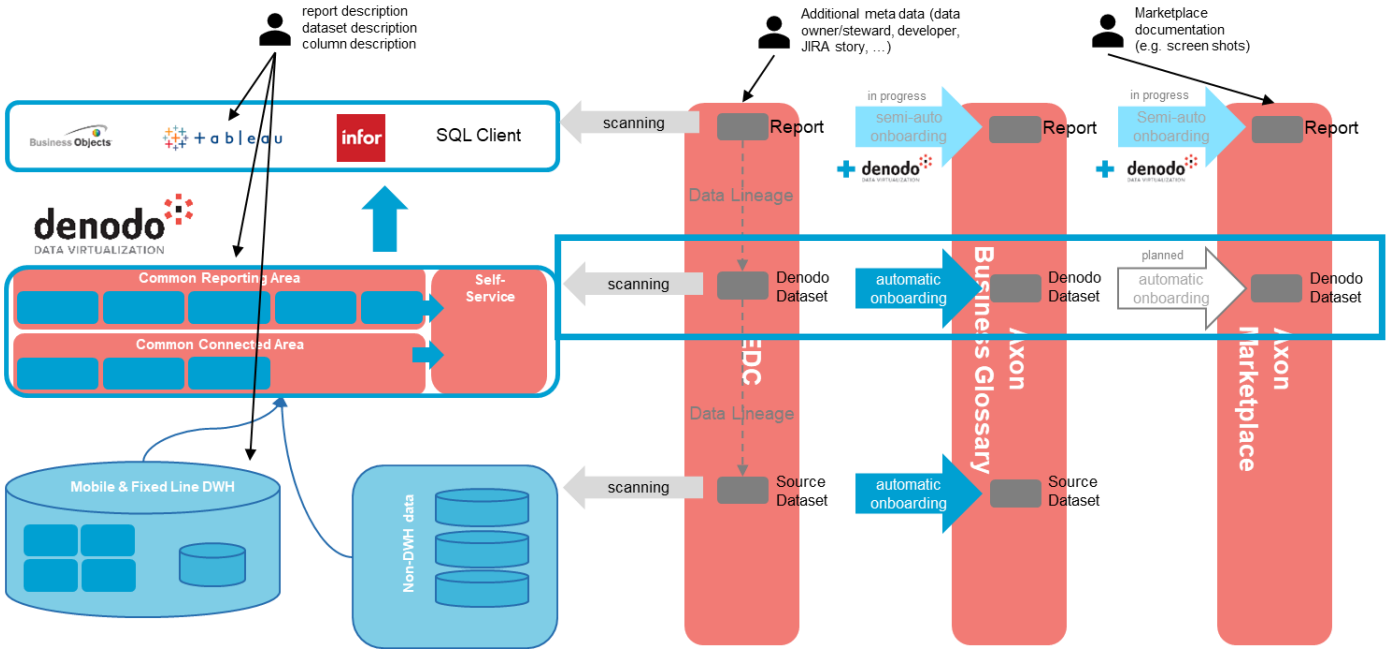




 © Hutchison Drei Austria GmbH – Confidential

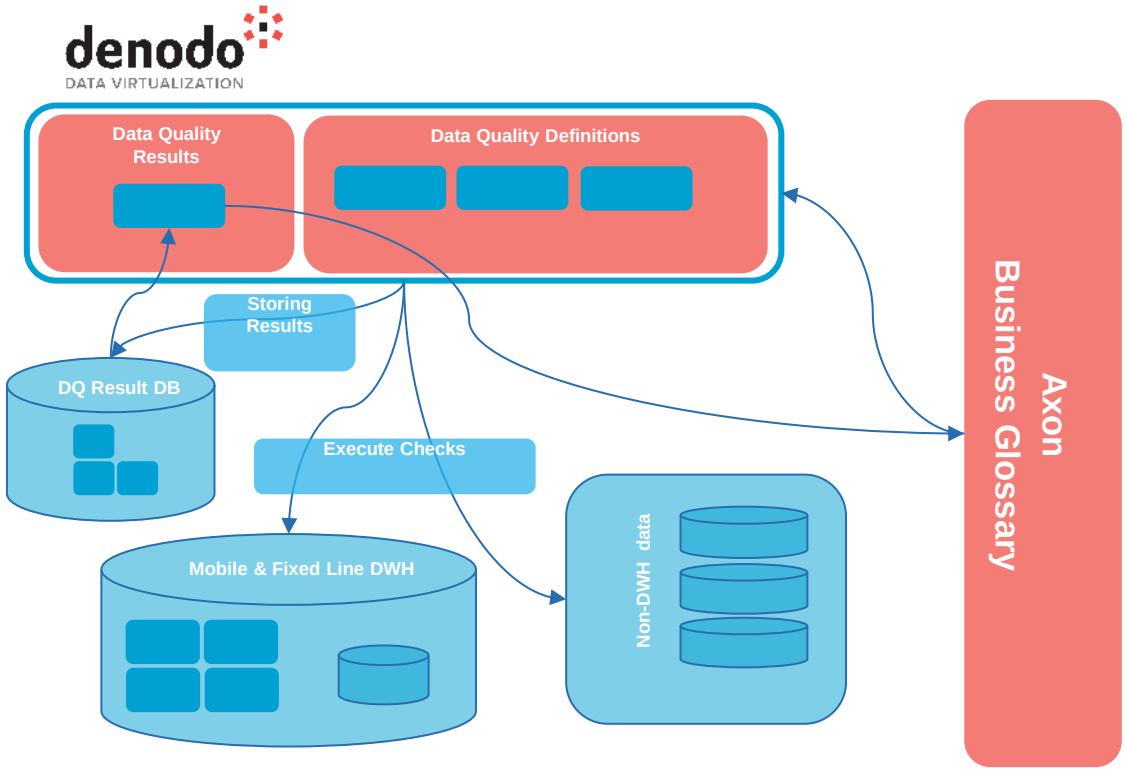
Data Product Concept (End-to-End)

- Step 4: Integrate in Data Governance process incl. Onboarding in Data Marketplace



 © Hutchison Drei Austria GmbH – Confidential

Data Quality Concept with Axon & Denodo

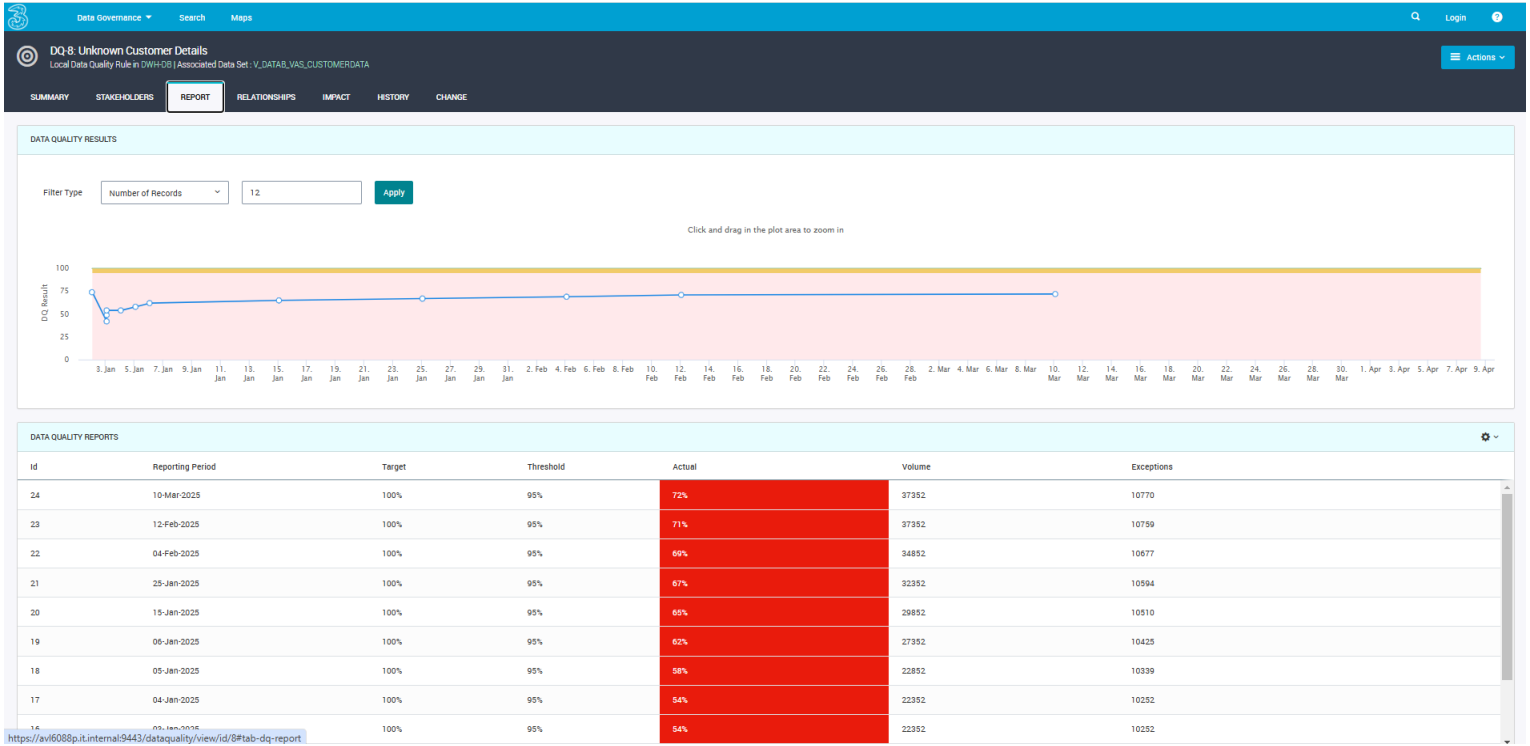


- Data Quality Rules on Data Sets defined in Axon
  - Incl. roles (Data Provider, Steward etc.)
- Availability of History & Reporting
- Definitions are pulled in Denodo and executed against the defined source
- Results are stored in a dedicated DB and shared back to Axon

© Hutchison Drei Austria GmbH – Confidential

Data Quality Concept with Axon & Denodo

Denodo feeds the results of the data quality checks back to Axon, where it can be discovered by stakeholders.



UDAL What's Next ?.

- **1** - Improve **processes** & **automation**
  - Development Process (creation of objects, CI/CD pipeline etc.)
  - Self-Service (creation of roles & permissions, descriptions etc.)
  - Training for specific roles & users

- **2** - Widen the scope to include **Big Data**
  - Results for Reporting based on Big Data Use Cases
  - Providing input for Data Engineers (Denodo as a Source)
  - Use Big Data cluster for MPP

- **3** - To everyone's surprise… **AI**
  - What can AI do for Denodo
    - As a support for developing & interacting with Denodo (specifically for Self-Service)
    - ad-hoc queries on the most used KPIs using standard dimensional information
  - What can Denodo do for AI
    - As part of our enterprise wide AI initiatives around integrating enterprise data into AI services

 © Hutchison Drei Austria GmbH – Confidential

Zentrale Datenbereitstellung für KI mit Denodo

Denodo Assistant und AI SDK

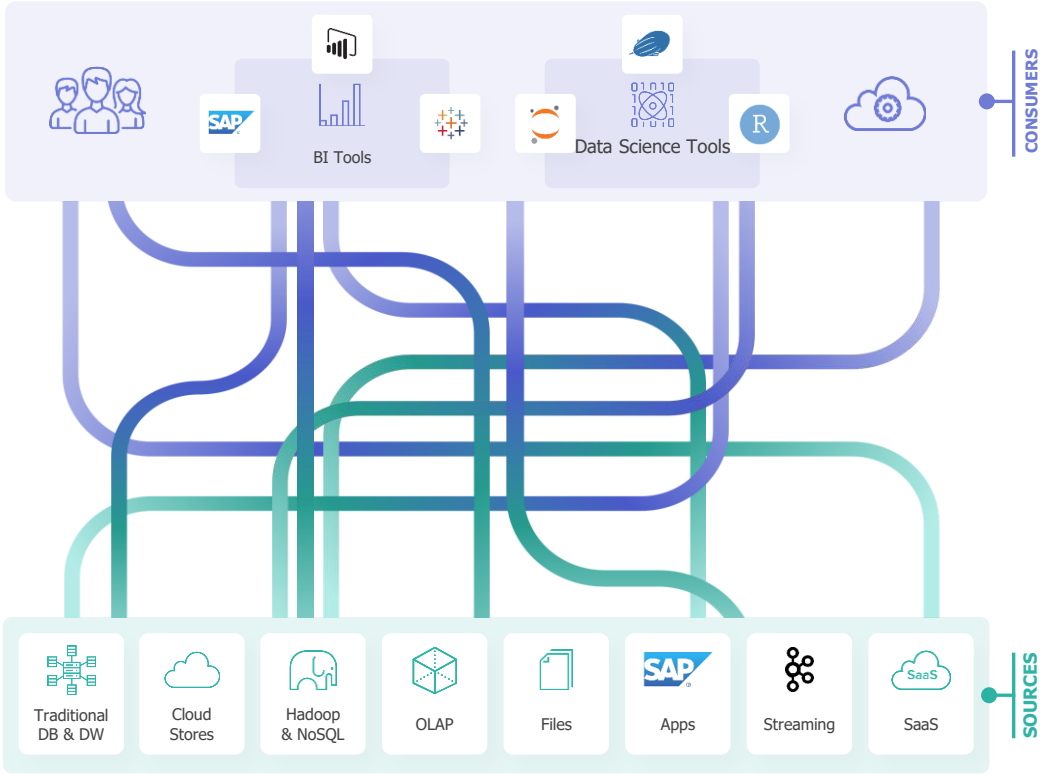
Jan Ulrich Maue
Sr. Sales Engineer
jmaue@denodo.com
(m) +49 (1525) 7969 540

10. April 2025

Distributed Data Landscape

Point-to-point data integration approaches are challenging:

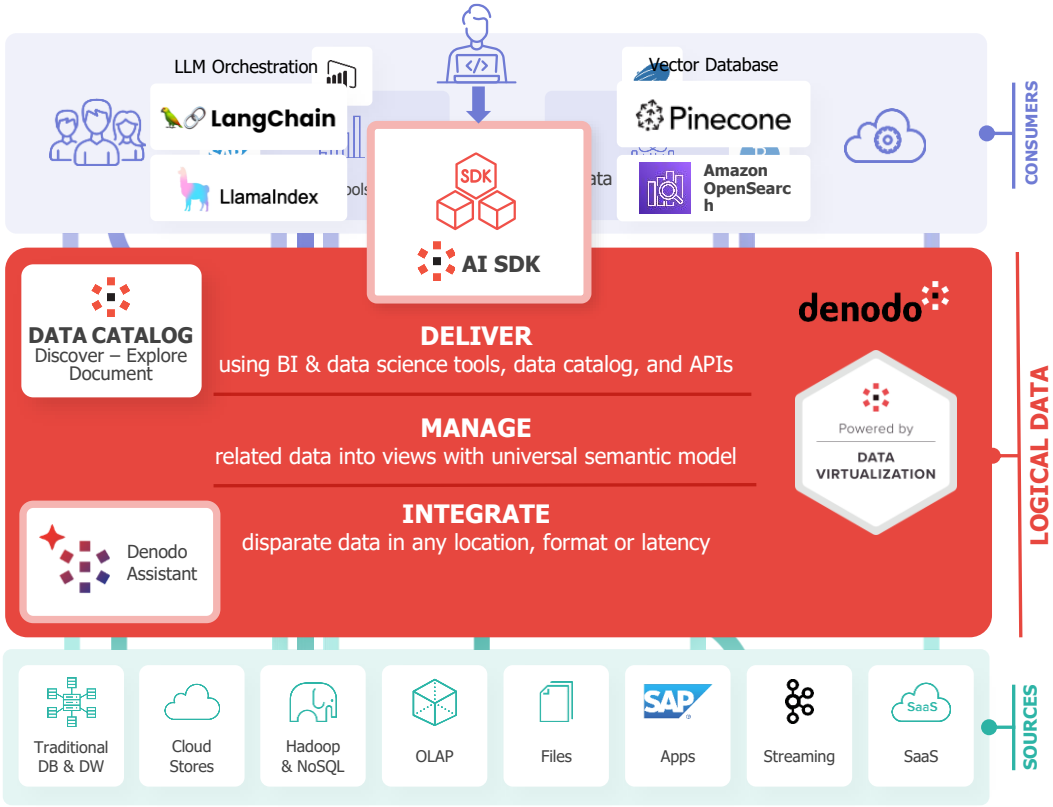
- Extracting and moving data increases latency and cost, and decreases quality
- Every project solves data access and integration in a different way
- Solutions are tightly coupled to data sources, impacting flexibility and agility



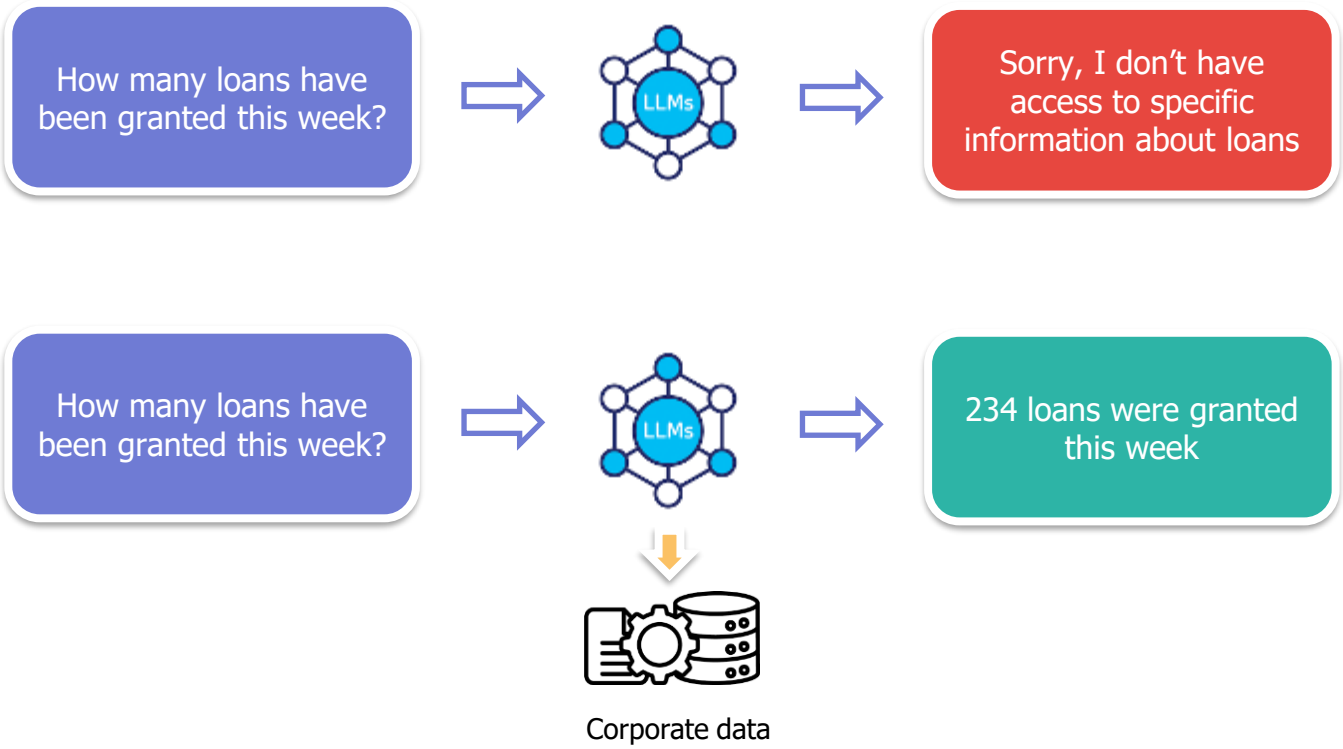
Denodo Platform: One logical platform for all your data

The Denodo Platform offers:

- Semantic Layer
- Centralized Data Governance
- Data Security and Masking
- Self-Service for Business Users
- Denodo Assistant for Recommendations and assisted Queries
- AI SDK enables development of Intelligent Applications and Chatbots



AI and dynamic data

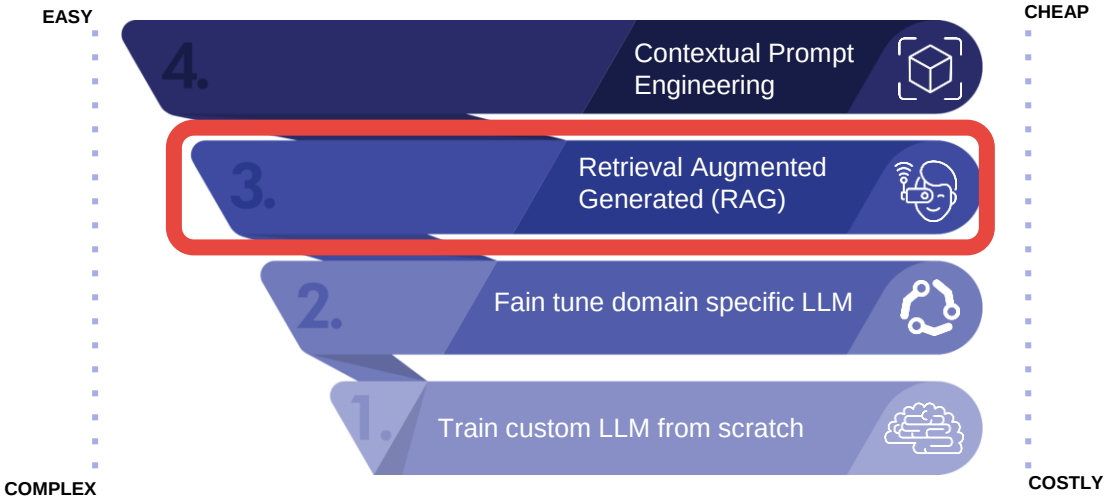


RAG – Retrieval-Augmented Generation

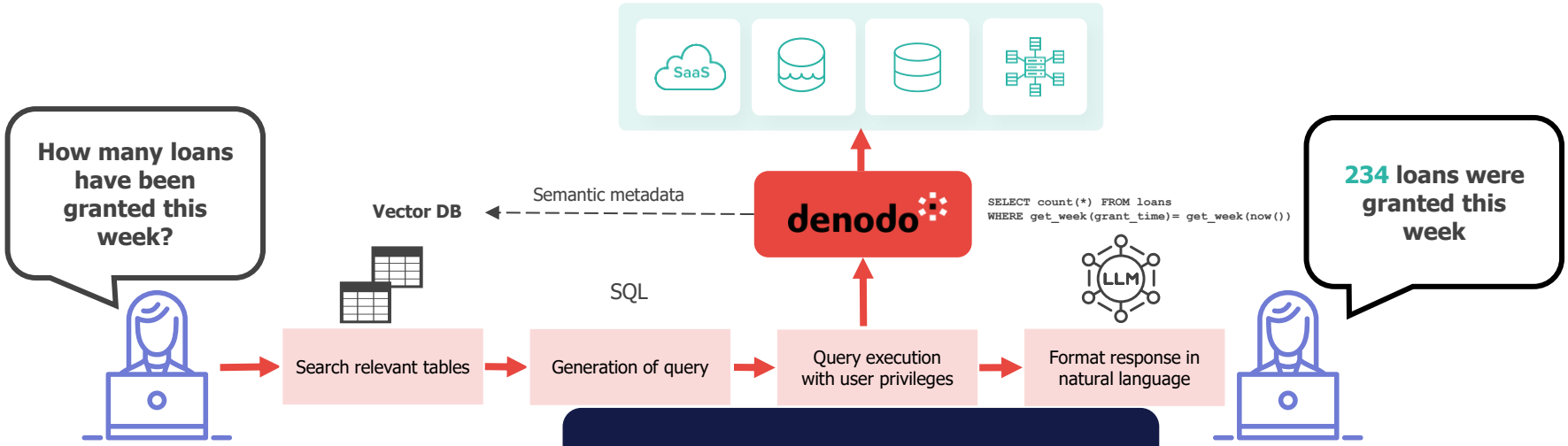
RAG is the standard for building Smart Apps with GenAI, grounded in facts using enterprise data.

WHY RAG?

- **Enhances AI Accuracy**:
  - access to real-time enterprise data
  - reduces hallucinations
  - precise, trustworthy outputs
- **Contextual Intelligence**:
  - domain-specific knowledge
  - smarter user interactions
- **Scalable Innovation**:
  - supports diverse applications
  - accelerates GenAI-driven initiatives



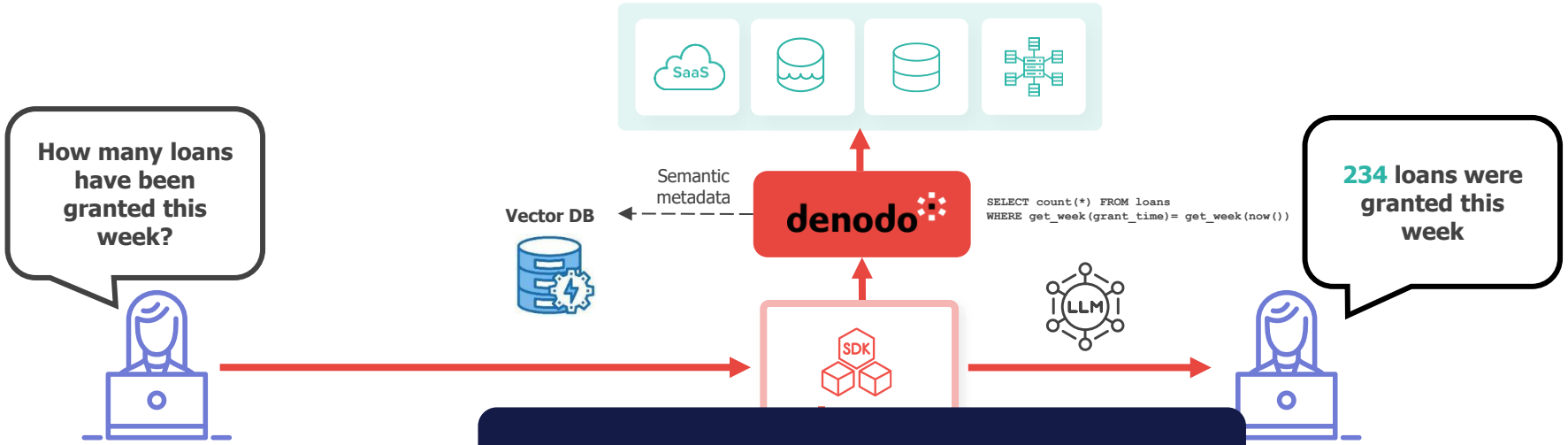
What are the steps to retrieve trustworthy answers?



Denodo provides:

- Single access point for all data
- Fine-grained access control
- Access to fresh data in real time

Denodo is uniquely equipped



Denodo AI SDK provides:

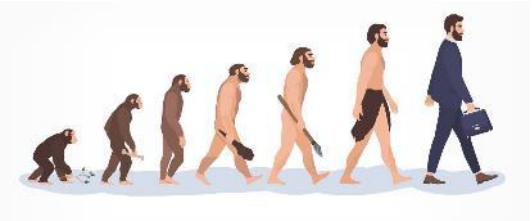
- Encapsulate the process in an easy-to-use SDK
- Simplifies genAI adoption for developers

Things we learned

 © Hutchison Drei Austria GmbH – Confidential

Lessons learned

- **1** - Evolution – not revolution
  - Take your time & “let the use cases come to you”
  - “You don’t have to change everything now, but look what it can do for you!”



- **2** - Find Influencers
  - Invest time in early adopters and be open about the (non existing) processes
  - “If Mike* thinks it can help us, it’s probably not bad!”

- **3** - Don’t mention Denodo
  - Most users don’t care about new tools or concepts, they want to have something that can help them in their daily tasks

*his name was not really Mike

 © Hutchison Drei Austria GmbH – Confidential

Data Virtualization benefits:

Why we will steadily increase the usage of Denodo

- Data virtualization can **hide technical complexity** regarding access to data
  (like storage-point, storage-structure, API's, query language and storing technology)
- Enable access to different data-sources in **one logical access-point**
- Possibility to do some **format changes and aggregation** of the data on the fly
- Possibility to **combine query-result-sets** across multiple source system
  - ***Reduce DWH-systems workload*** *by replicating data only when necessary, therefore also reducing (future) data-storage requirements*
- Possibility to offer direct or "**live**" **reporting** on the data sources when required
- Possibility to **publish result-sets** as View or also as Rest **services**

Provide technical framework for a data mesh *and increase transparency of and responsibility for data products*

© Hutchison Drei Austria GmbH – Confidential

Thank you.

Macht's einfach.