INSIGHTS & DATA

Webinar Series – Data Meshes & Knowledge Graphs

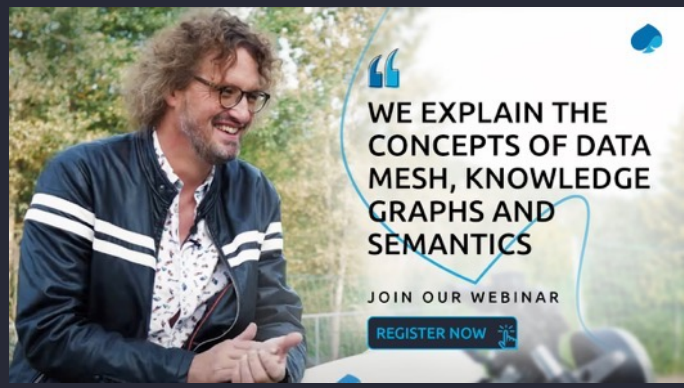
Data Meshes & Knowledge Graphs

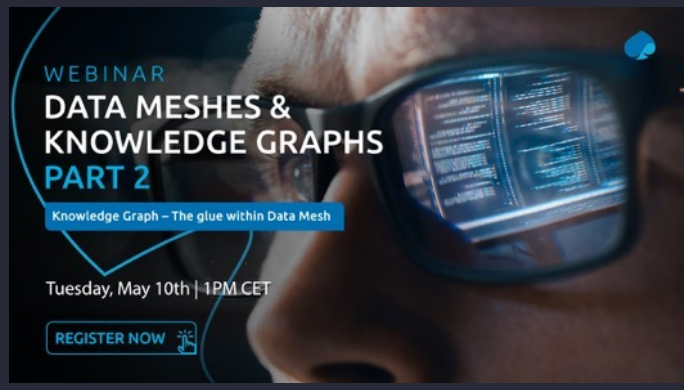
Webinar Series

1. Part 1 – ”How you can add the human way of understanding data better”
   April 6th, Robert Engels
   Replay available

2. Part 2 – ”Knowledge Graph – The glue within Data Mesh”
   May 10th, Arne Rossman
   Todays webinar

3. Part 3 – ”Semantic data mesh - The missing links in your data platform”
   May 20th, speaker Aniruddha Khadkikar
   Hosted by PTC with Capgemini as a partner at the NordicTalks webinar





Capgemini

DATA MESHES AND KNOWLEDGE GRAPHS PART 2

Knowledge Graph – The glue within Data Mesh

Arne Rossmann

WEBINAR
DATA MESHES & KNOWLEDGE GRAPHS

Presented by Capgemini, Insights & Data

Today you will learn about:

- Why the main pitfall of Data Mesh is on the Governance Layer
- How KnowledgeGraph can solve that issue
- What building blocks you need to enable users

Speaker:

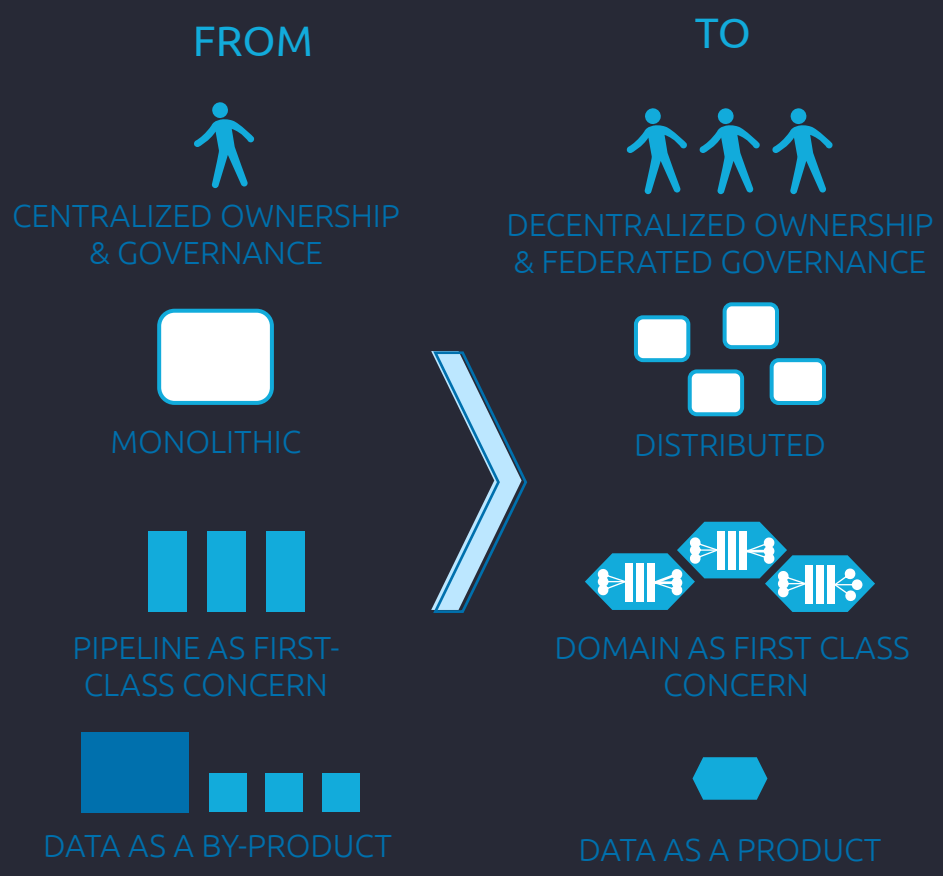
Arne Rossmann

Data & AI Foundation Lead Intelligent Industry for Capgemini, Insights & Data

SHORT INTRODUCTION TO DATA MESH

DATA MESH ON A NAPKIN

The sociotechnical approach to share, access and manage analytical data in complex and large-scale environments - within or across organizations.



Data Mesh Principles

DOMAIN ORIENTED DECENTRALIZATION

Data Mesh essentially refers to the concept of breaking down data lakes and siloes into smaller, more decentralized portions. Much like the shift from monolithic applications toward microservices architectures in the world of software development, Data Mesh can be described as a data-centric version of microservices

DATA AS A PRODUCT

A data product is a product that primarily uses data (e.g. Legacy Data) to contribute value to the achievement of an organization's objectives

SELF-SERVE DATA INFRA AS A PLATFORM

Data self-service depends on the technology used and can be deployed, for example, as Docker, Kafka Service, or Spark code

EVERY “ENTERPRISE” HAS IT’S “DOMAINS” I WANT TO NAVIGATE TO

Juice Bar

Fruits

Juices

Vegetables

https://unsplash.com/photos/AG0iadQfRDg

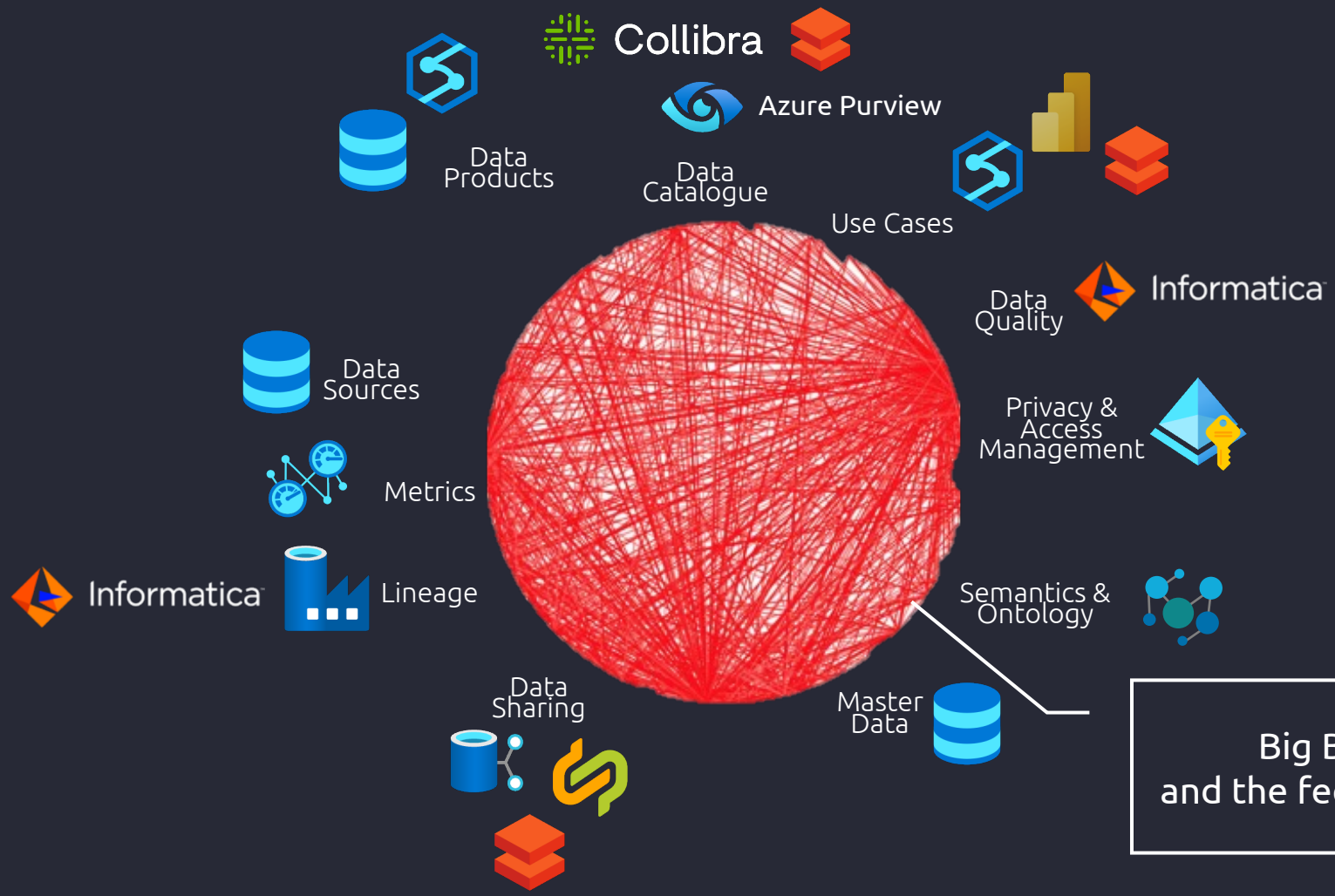
WITHIN A “DOMAIN” THE ”DATA PRODUCT” ARE EXPOSED WITH CLEAR “PROPERTIES”

https://unsplash.com/photos/TcpYjs6qF9o

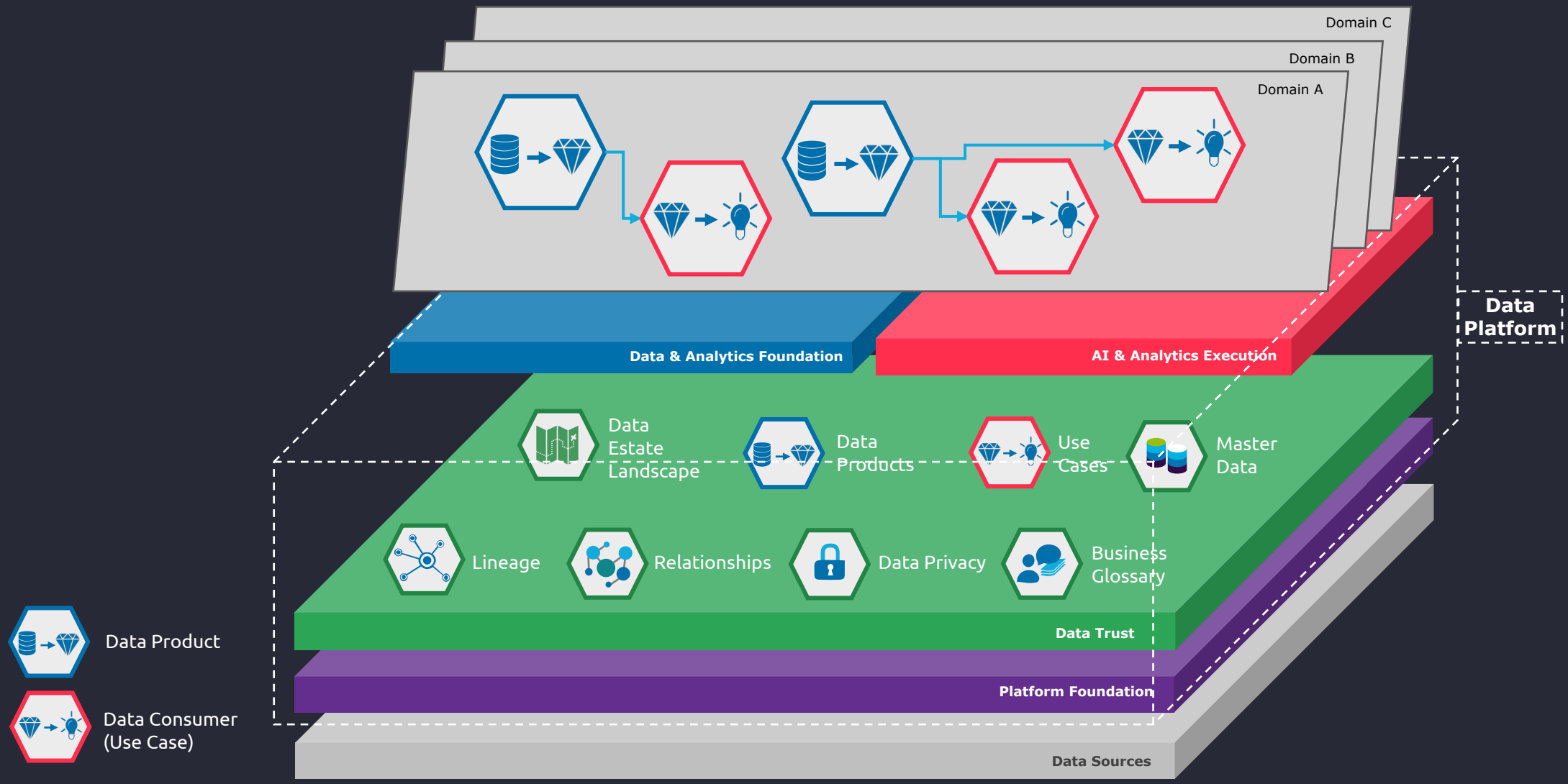
https://unsplash.com/photos/Gk8LG7dsHWA

WHY WE NEED A KNOWLEDGE GRAPH ?

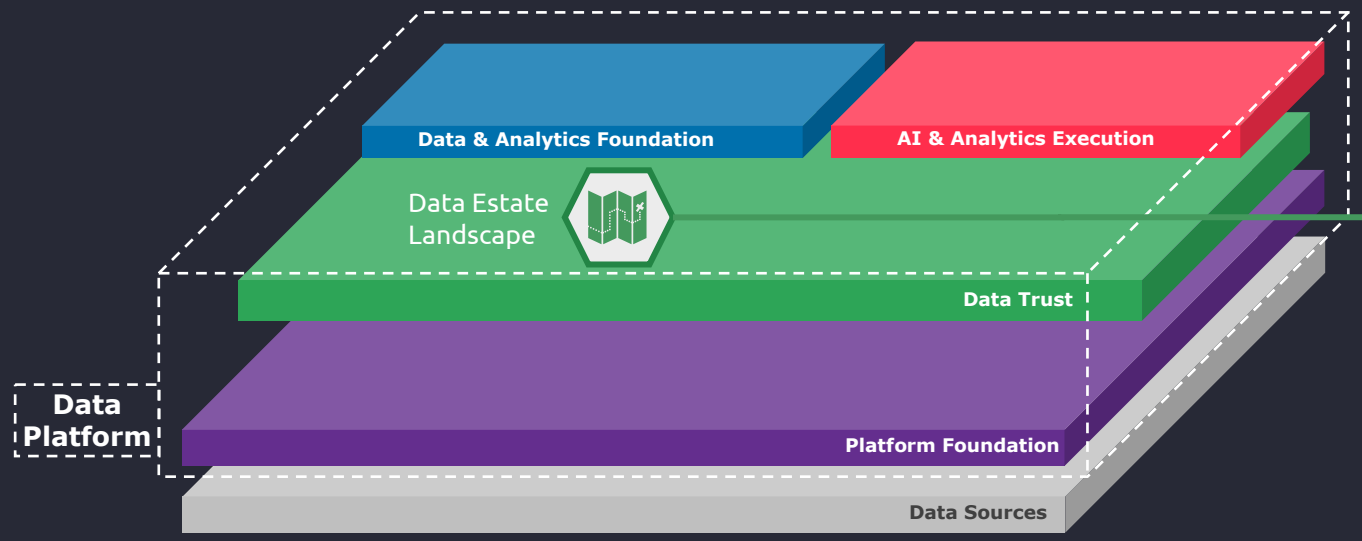
AT ENTERPRISE LEVEL A MULTITUDE OF TECHNOLOGIES HAS TO BE COMPROMISED FOR DATA GOVERNANCE

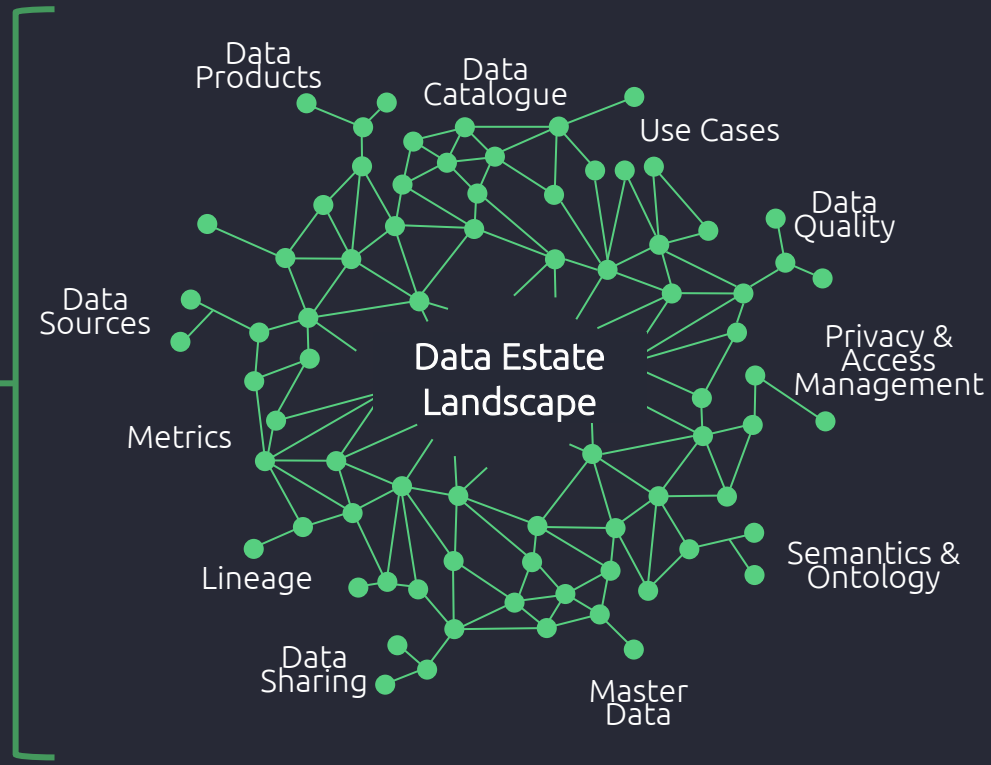


THE DATA TRUST LAYERS ENABLES GOVERNANCE & TRANSPARENCY WITHIN THE ORGANIZATION



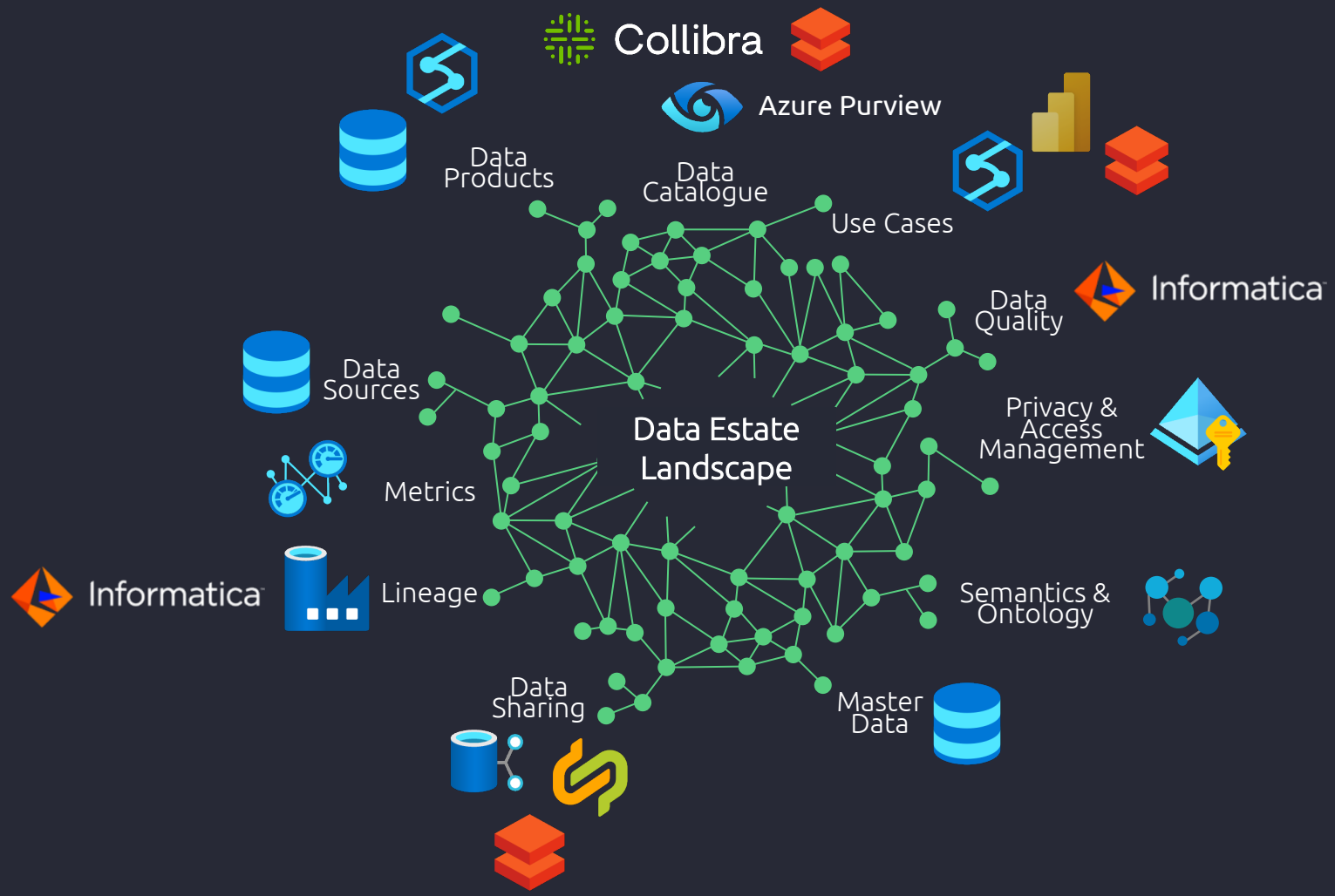
A GRAPH BASED DATA ESTATE LANDSCAPE ADDRESSES THESE ISSUES AND PROVIDES TRANSPARENCY AND AGILITY





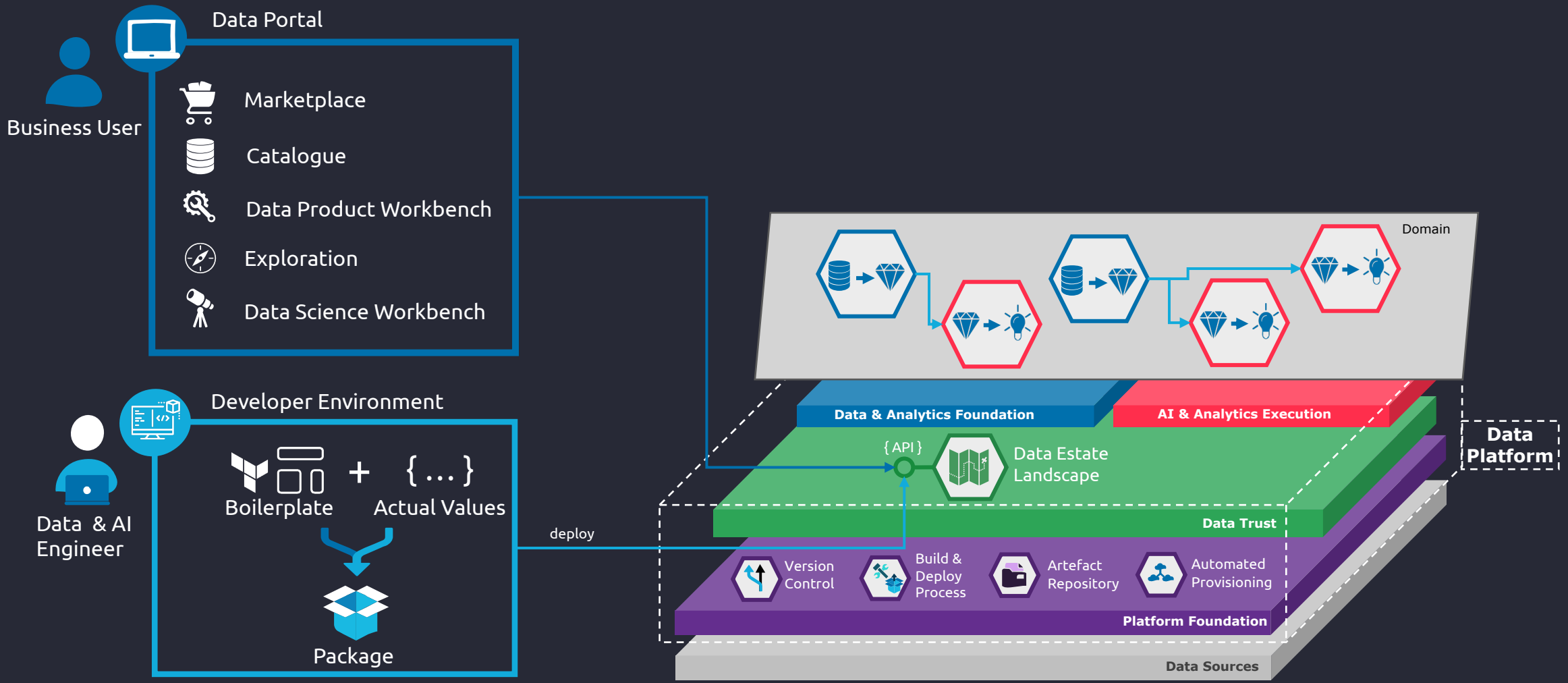
- The Data Estate Landscape brings transparency to the Data Platform Ecosystem by maintaining the relationship between Sources, Data Products and Use Cases
- This enables faster Use Case development and reuse of data and resources
- Establishment of the Data Estate Landscape ensures overall governance over the Data Ecosystem

AT ENTERPRISE LEVEL THE DATA ESTATE LANDSCAPE HAS TO COMPROMISE A MULTITUDE OF TECHNOLOGIES



HOW TO ENABLE BUSINESS WITH THAT TECHNOLOGY

BY PROVIDING API ENDPOINTS THE DATA ESTATE LANDSCAPE SUPPORT DEVELOPERS AND BUSINESS USERS ON THEIR WORK



THE DATA PORTAL ACTS AS THE ENTRY POINT FOR THE USERS

https://data-portal.company.com

Introduction

This is the Data Portal in which you can find:

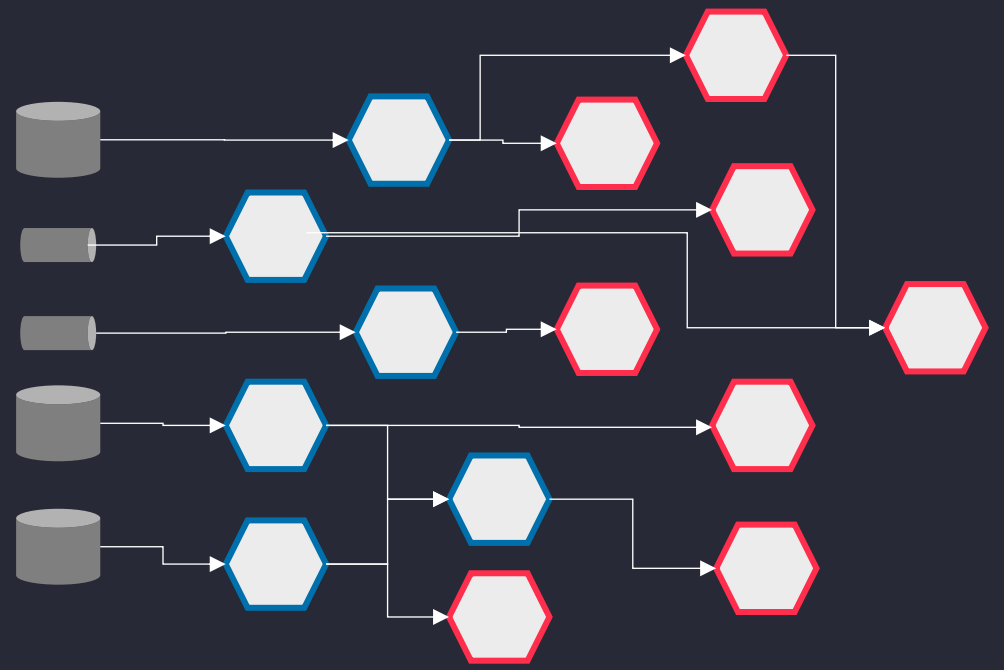
- Marketplace
- Catalogue
- Data Product Workbench
- Exploration
- Data Science Workbench

Tutorial

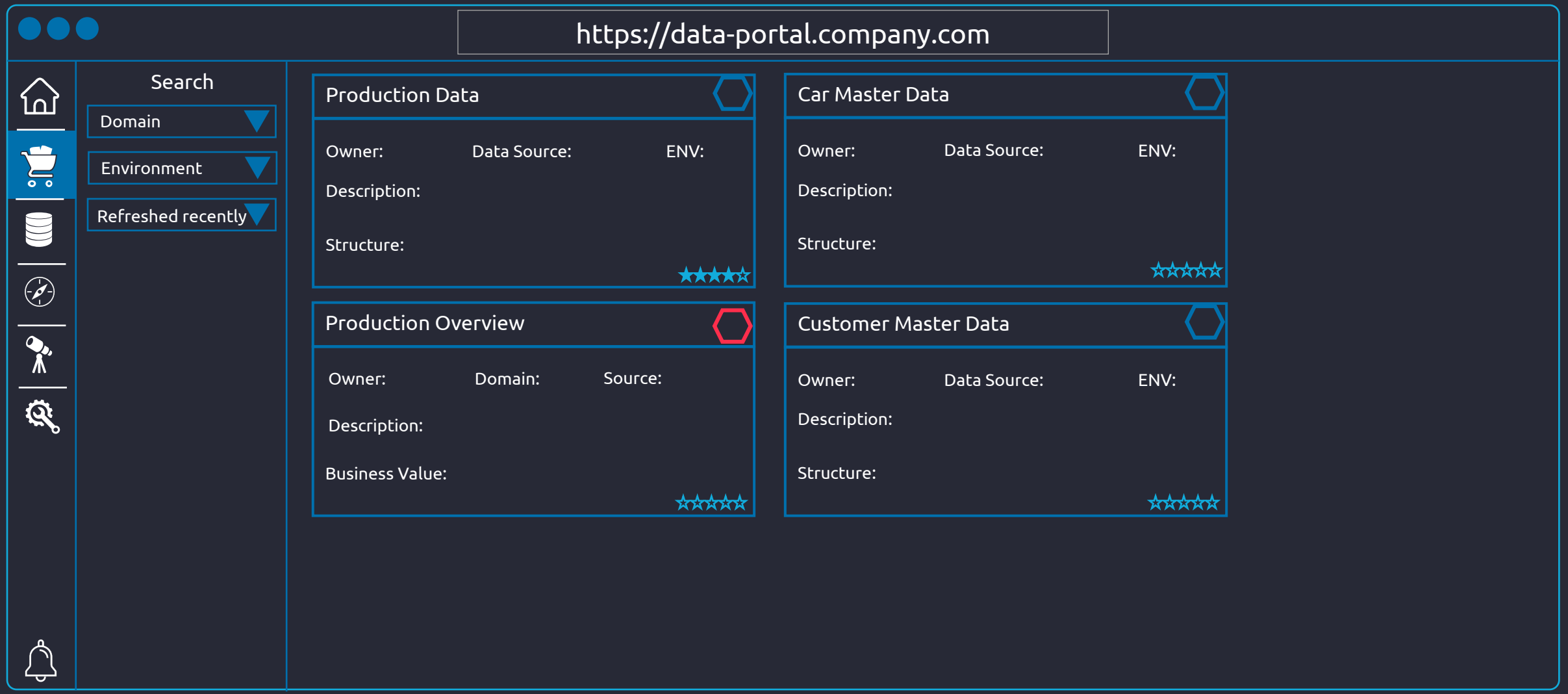
With this tutorial you'll get a step-by-step guide on how to work within the Data Portal and find the data you're looking for. Click here to continue.

Data Estate Landscape

Domain A | Domain B | Domain C



THE DATA PORTAL ACTS AS THE ENTRY POINT FOR THE USERS



THE DATA PORTAL ACTS AS THE ENTRY POINT FOR THE USERS

https://data-portal.company.com

Search

Domain

Environment

Type

Production Data

Owner: Dagobert Duck

Data Source: MES Siemens

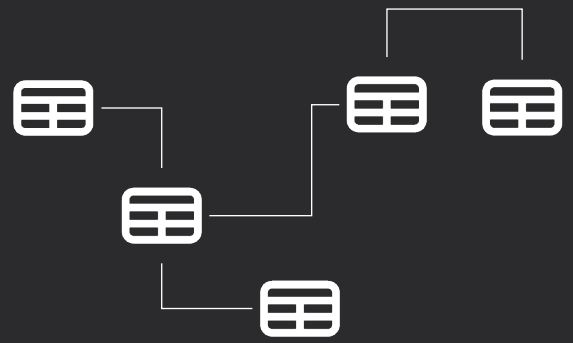
Type: Relational

ENV: DEV

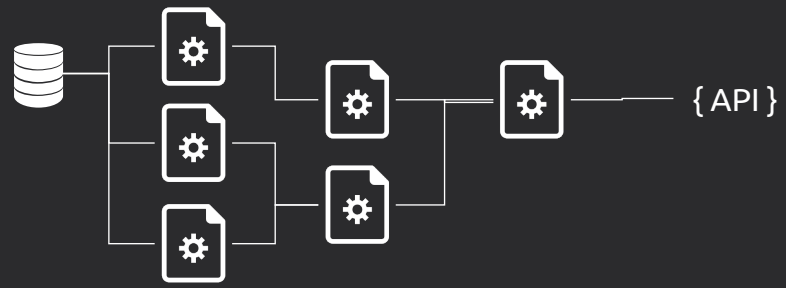
Description: Lorem ipsum sit dolor

Metrics:
Last loaded: 2022-04-12-21:30:00
Load frequency: every hour
Data quality: 100% matched records

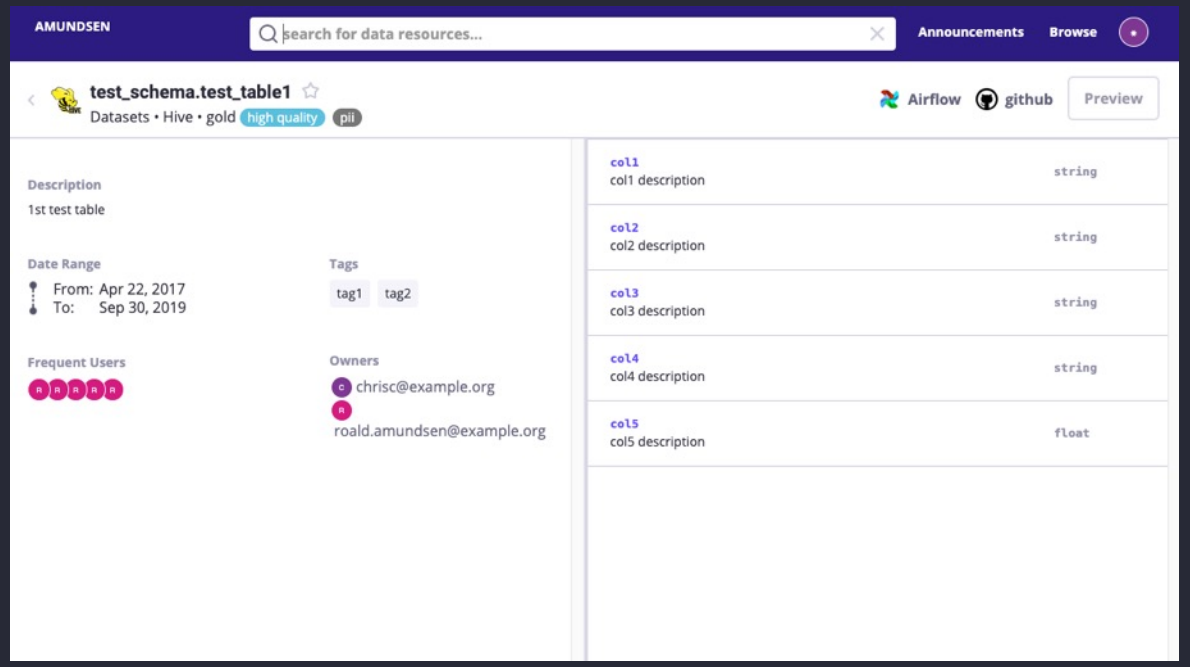
Structure:

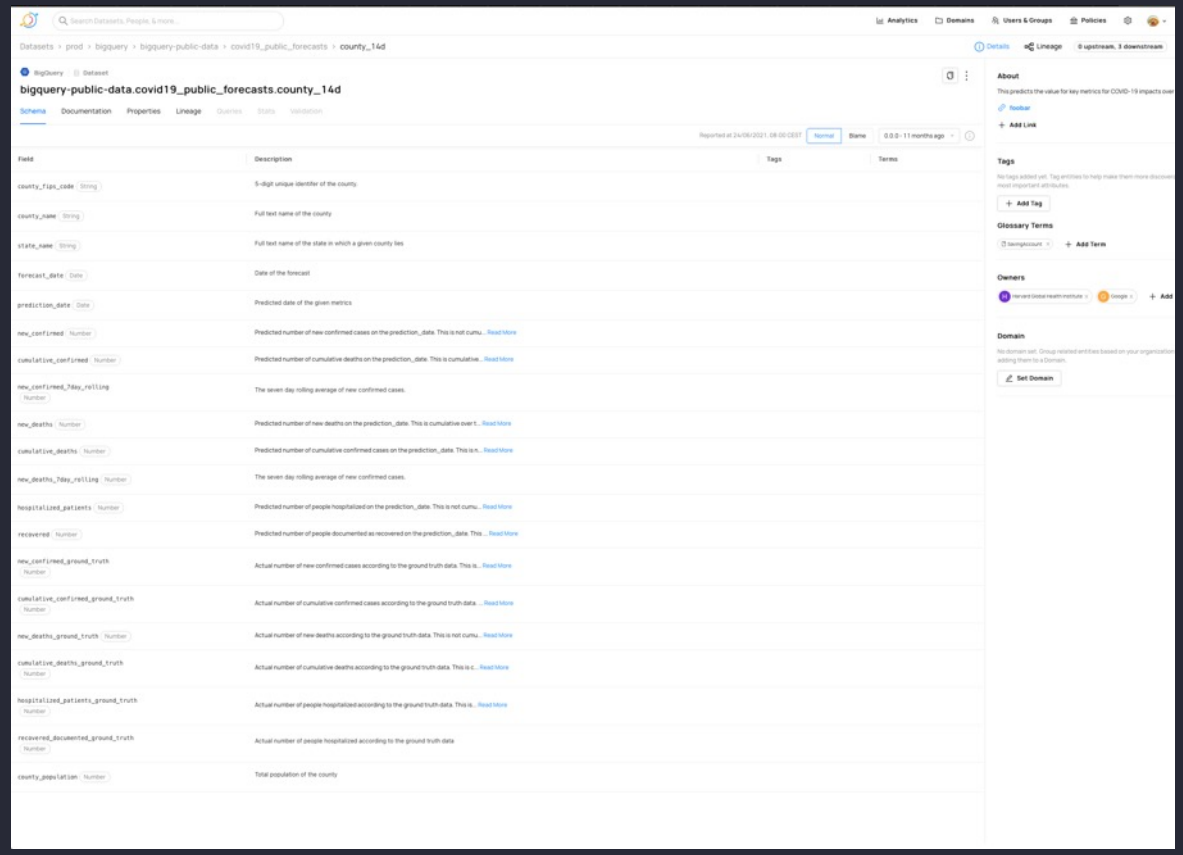


Lineage:

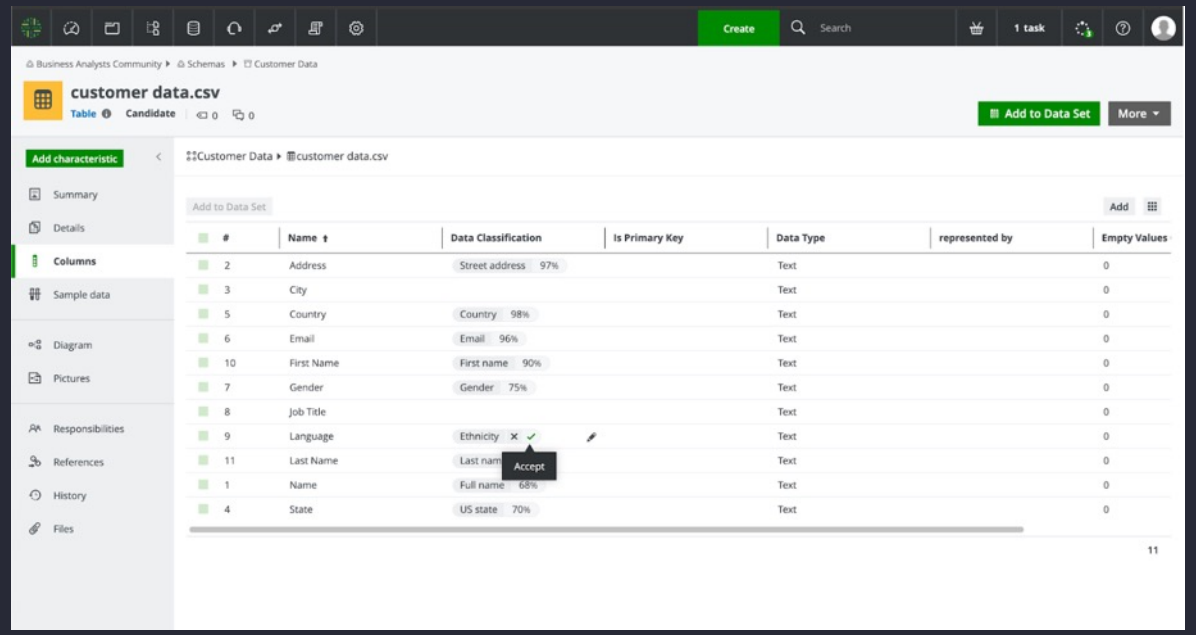


EXAMPLES OF DATA CATALOGUES

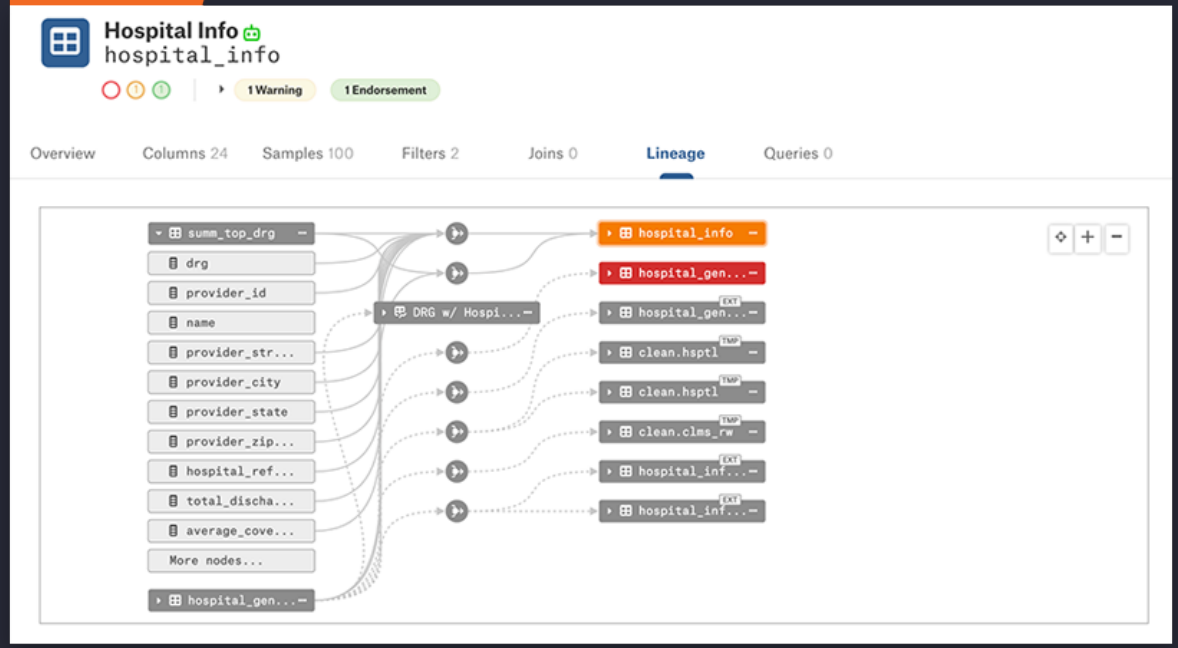




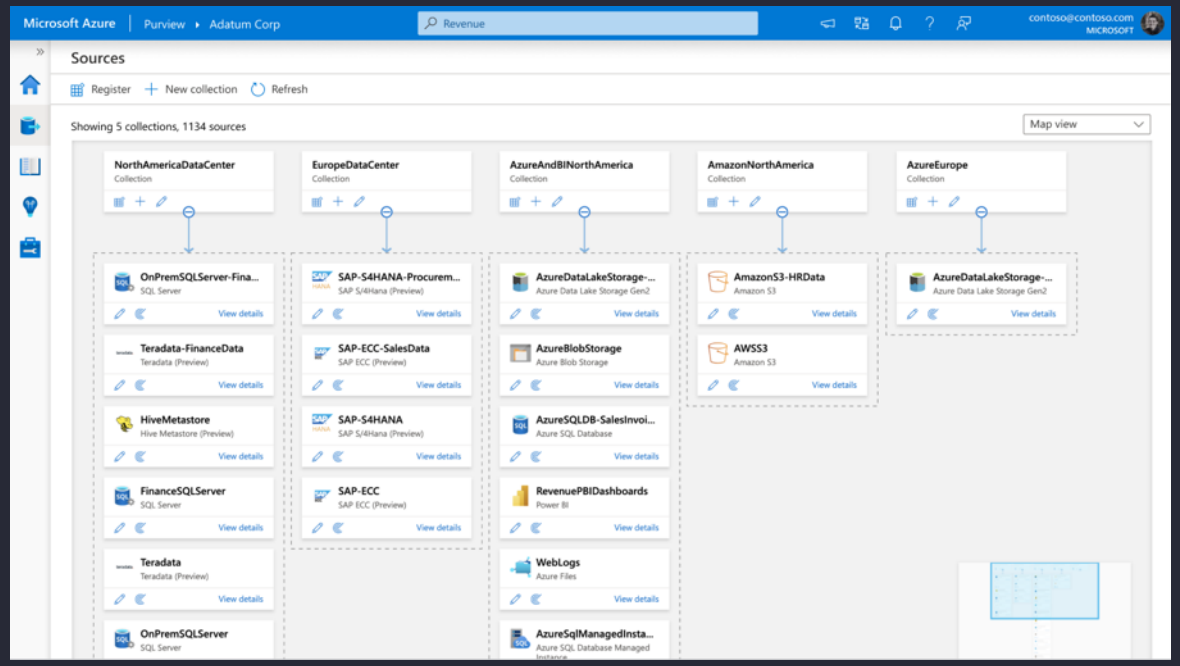
EXAMPLES OF DATA CATALOGUES

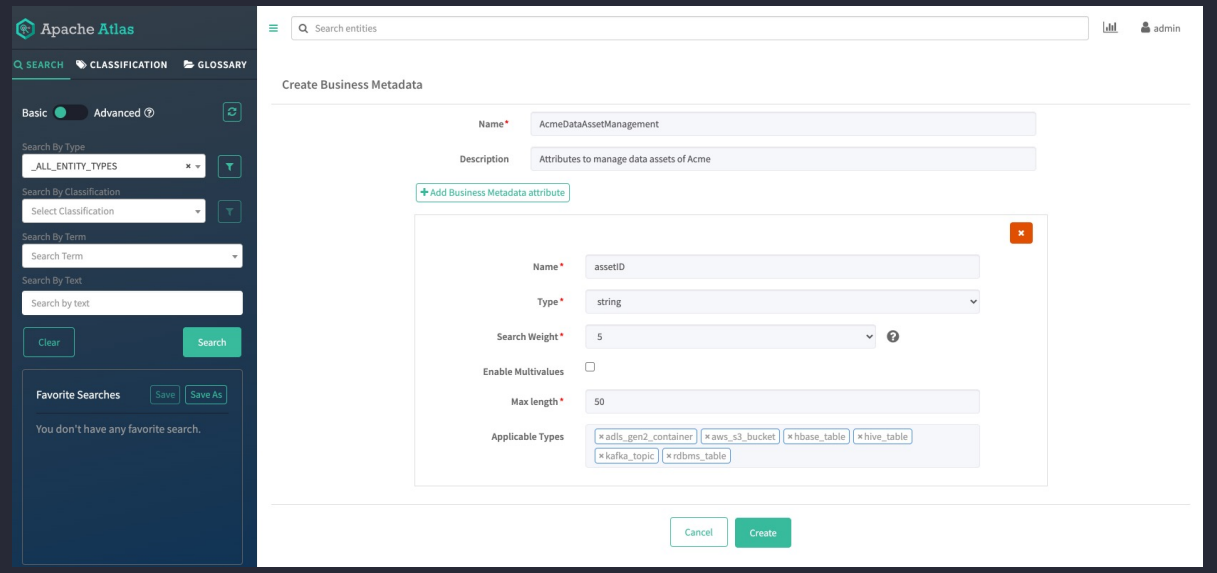


DATA LINEAGE



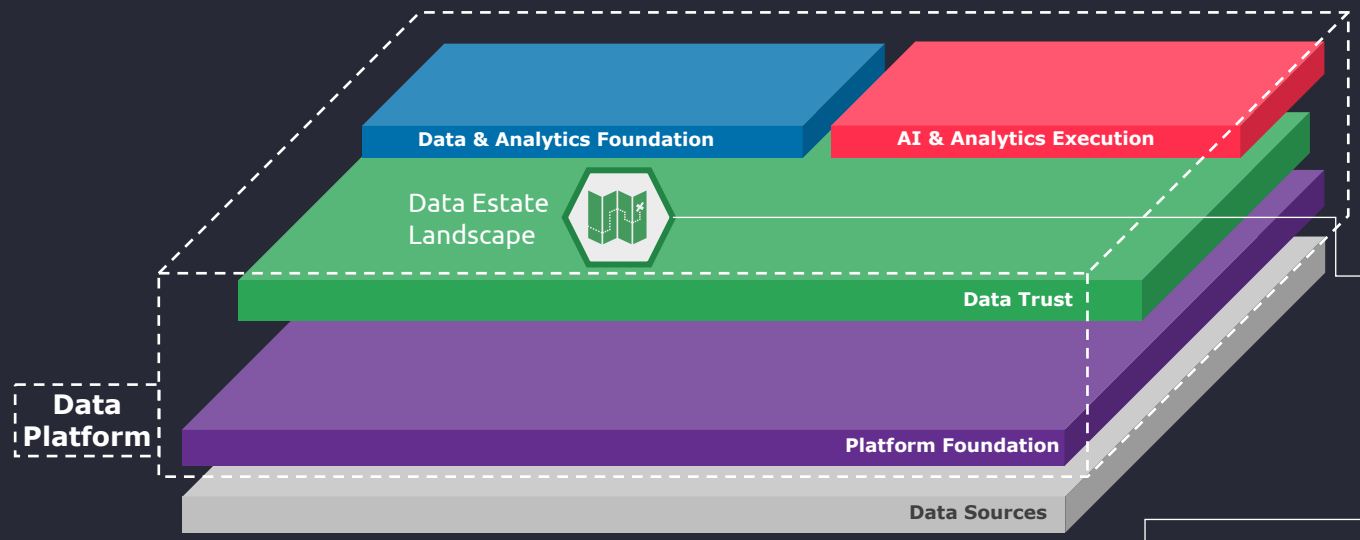
EXAMPLES OF DATA CATALOGUES

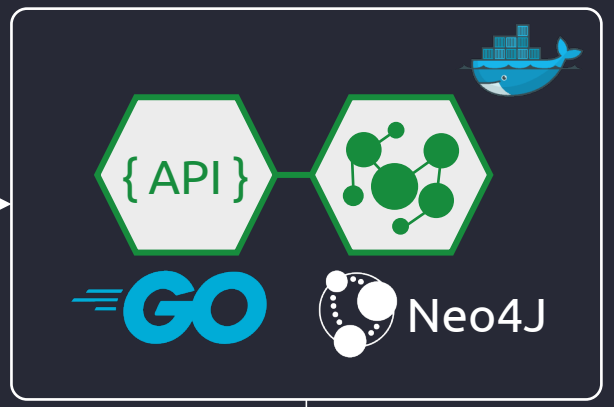


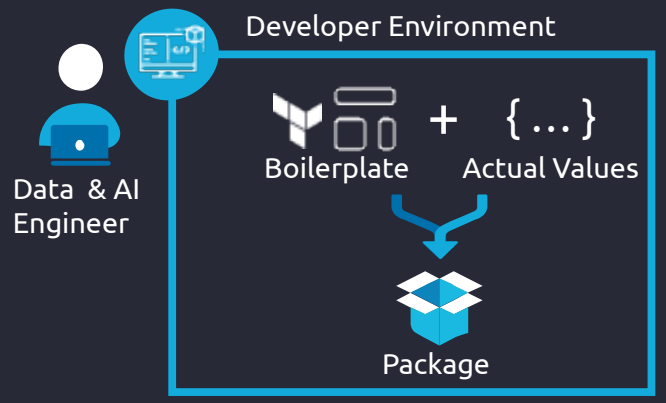


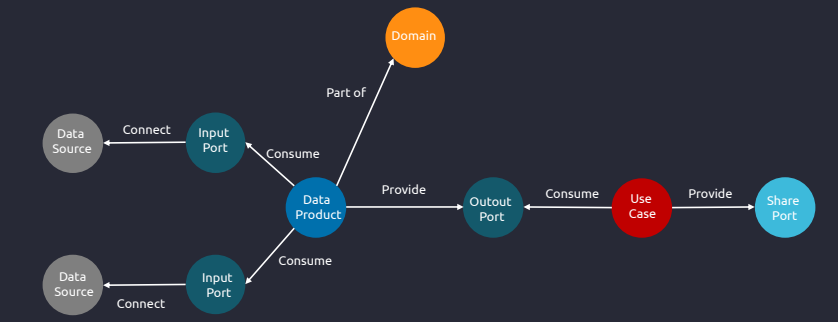
EXEMPLARY IMPLEMENTATION

A GRAPH BASED DATA ESTATE LANDSCAPE ADDRESSES THESE ISSUES AND PROVIDES TRANSPARENCY AND AGILITY

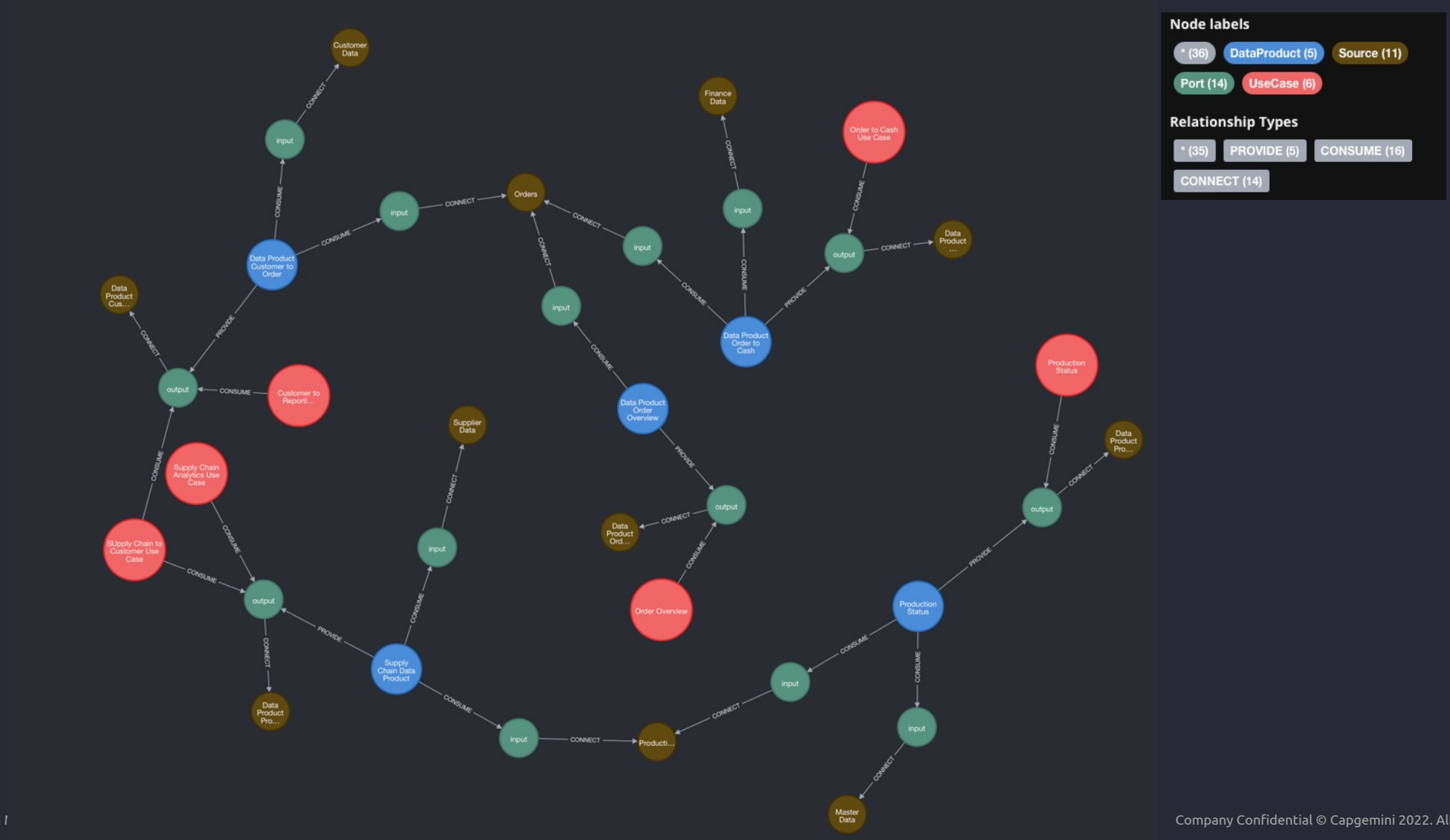




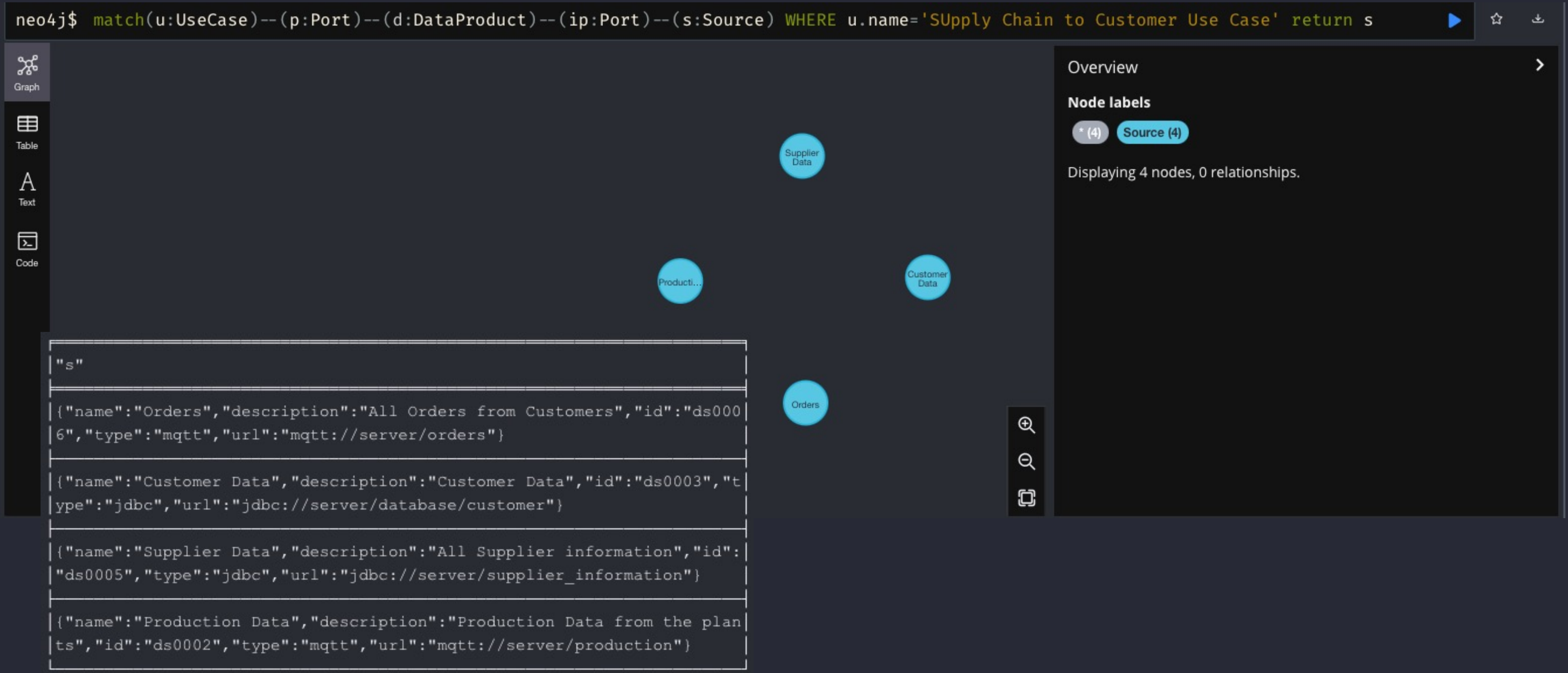




KNOWLEDGE GRAPH OF EXEMPLARY DATA MESH

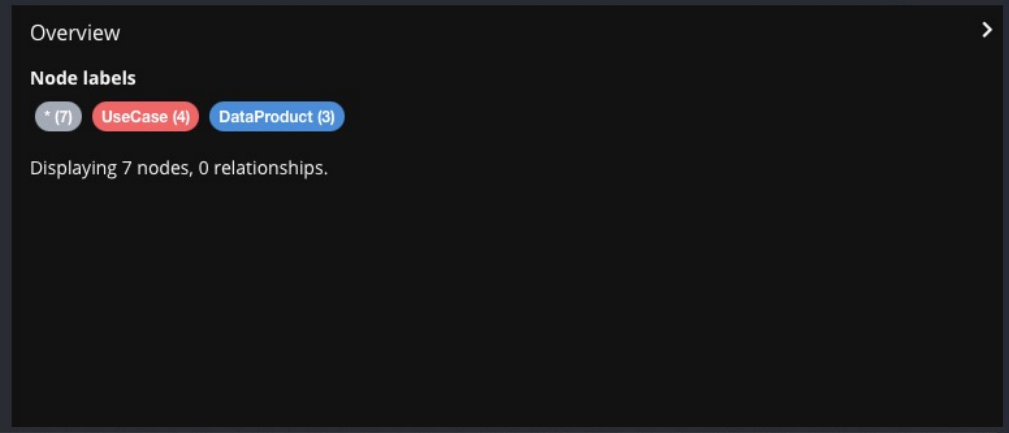


GETTING ALL INVOLVED DATA SOURCES FOR SELECTED USE CASE



GETTING ALL DATA PRODUCTS AND USE CASES OF DATA SOURCE

```
neo4j$ match(u:UseCase)--(p:Port)--(d:DataProduct)--(ip:Port)--(s:Source) WHERE s.name='Orders' return u,d
```



"u"	"d"
{"owner":"Dagobert Duck","createdDate":"2022-03-07 10:00:00","deputy":"Donald Duck","name":"Order Overview","description":"Having the overview on orders","lastModified":"2022-03-07 10:00:00","id":"uc00006","version":"1.0.0"}	{"owner":"Darth Sidious","createdDate":"2022-03-14 10:00:00","collectedSince":"2022-03-14 10:00:00","name":"Data Product Order Overview","description":"Data Product for Order Overview","lastModified":"2022-03-07 10:00:00","id":"dp00005","version":"1.0.0"}
{"owner":"Dagobert Duck","createdDate":"2022-03-07 10:00:00","deputy":"Donald Duck","name":"Order to Cash Use Case","description":"Making the CFO happy","lastModified":"2022-03-07 10:00:00","id":"uc00005","version":"1.0.0"}	{"owner":"Darth Sidious","createdDate":"2022-03-14 10:00:00","collectedSince":"2022-03-14 10:00:00","name":"Data Product Order to Cash","description":"Data Product for making orders to cash transparent","lastModified":"2022-03-07 10:00:00","id":"dp00004","version":"1.0.0"}
{"owner":"Dagobert Duck","createdDate":"2022-03-07 10:00:00","deputy":"Donald Duck","name":"Customer to Order Reporting","description":"Customer to Order","lastModified":"2022-03-07 10:00:00","id":"uc00004","version":"1.0.0"}	{"owner":"Sith Loard","createdDate":"2022-03-14 10:00:00","collectedSince":"2022-03-14 10:00:00","name":"Data Product Customer to Order","description":"Data Product for connecting Customer to Orders","lastModified":"2022-03-07 10:00:00","id":"dp00003","version":"1.0.0"}
{"owner":"Dagobert Duck","createdDate":"2022-03-07 10:00:00","deputy":"Donald Duck","name":"SUpply Chain to Customer Use Case","description":"Connecting SUpply Chain and Customer Data","lastModified":"2022-03-07 10:00:00","id":"uc00003","version":"1.0.0"}	{"owner":"Sith Loard","createdDate":"2022-03-14 10:00:00","collectedSince":"2022-03-14 10:00:00","name":"Data Product Customer to Order","description":"Data Product for connecting Customer to Orders","lastModified":"2022-03-07 10:00:00","id":"dp00003","version":"1.0.0"}

CONCLUSIONS

DATA MESH WITHOUT DATA GOVERNANCE

Will end up in muddy games

Capgemini

GET THE FUTURE YOU WANT

capgemini.com

About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of over 325,000 team members more than 50 countries. With its strong 55-year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fueled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2021 global revenues of €18 billion.

Get The Future You Want | www.capgemini.com