

# The anatomy of a data product

Johan Ludvig Brattås

Enora

# Agenda

- Storytime
- A closer look at the various components
- What else?



# Storytime

- First public definition of a data product by DJ Patil in 2012 book Data Jiu-jitsu (O'Reilly)

“A data product is a product that facilitates an end goal through the use of data.”

Wait, isn't data products  
a data mesh thing?



# Kind of?

- Zhamak Deghani proposed a new definition in 2019 and refined in 2020

“The node on the mesh that encapsulates three structural components required for its function, providing access to the domain's analytical data as a product.”

# And...

- Then we have Andrew Gioia
- Defined data products in his 2024 book “**Managing Data as a Product: Design and build data-product-centered socio-technical architectures**” Packt
- Builds on DJ Patils definition

“A data product is a product in which the use of data not only supports *but also drives the development of the functionalities* necessary to achieve its goals.”

# Latest and greatest

- Jean-Georges Perrin et al.
- Defined data products in an article in 2025 - **Defining Data Products: A Community Effort**

<wall of text because of consensus...>

# Wall of text made easier

- A **data product** is a reusable, active, and standardized data asset designed to deliver measurable value to its users — whether internal or external — by applying the rigorous principles of product thinking and management.
- It comprises one or more **data artifacts** (e.g., datasets, models, pipelines) and is enriched with **metadata**, including governance policies, data quality rules, **data contracts**, and, where applicable, a **Software Bill of Materials (SBOM)** to document its dependencies and components.
- Ownership of a data product is aligned to a specific domain or use case, ensuring **accountability, stewardship**, and its **continuous evolution** throughout its lifecycle.
- Adhering to the **FAIR principles** — Findable, Accessible, Interoperable, and Reusable — a data product is designed to be discoverable, scalable, reusable, and aligned with both business and regulatory standards, driving innovation and efficiency in modern data ecosystems.



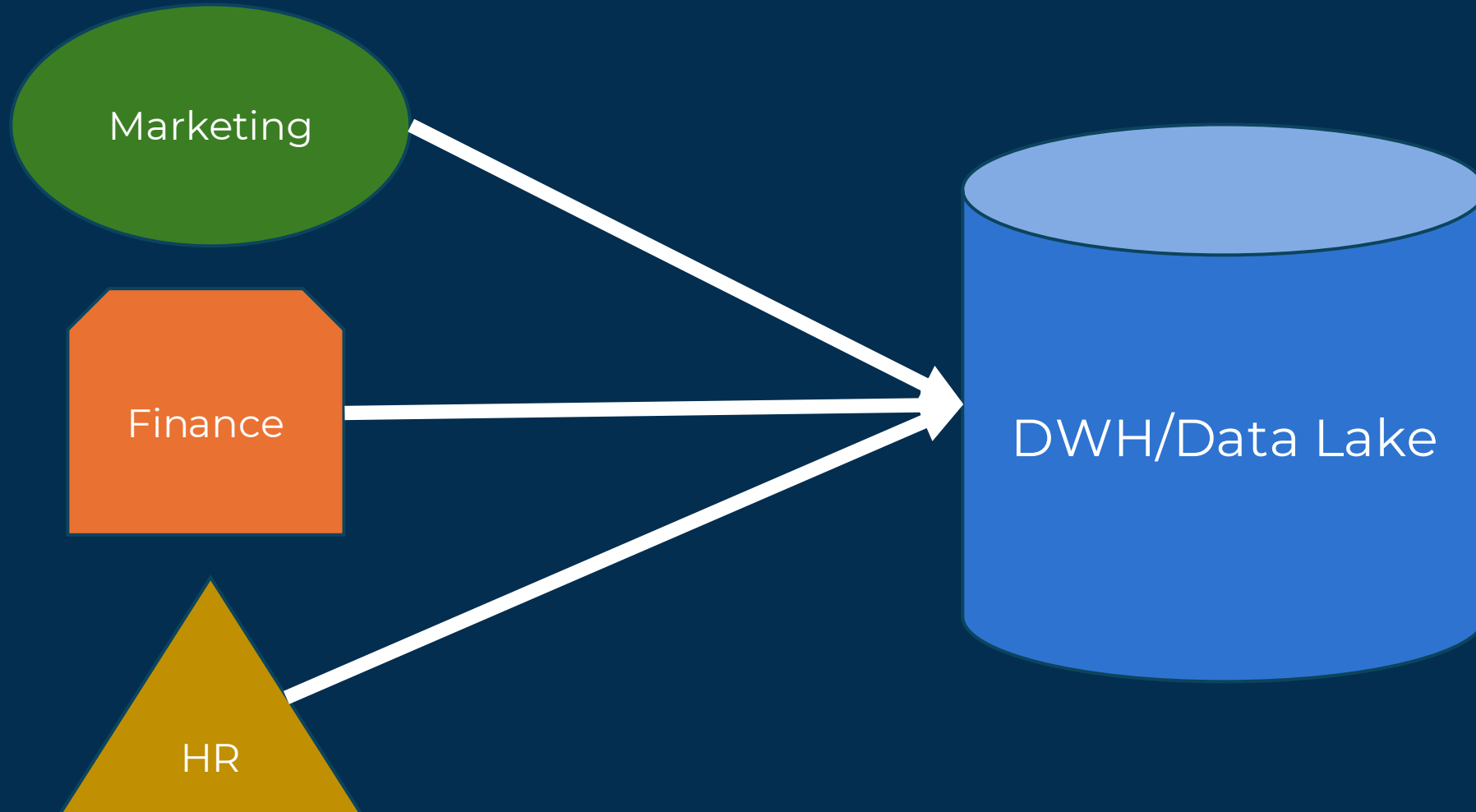
# DATA MESH

Let's have a look at the Data Mesh principles

# Data Mesh Principles

Domain-Oriented Decentralized Data Ownership and Architecture

# Domain-Oriented Decentralized Data Ownership and Architecture



# Domain-Oriented Decentralized Data Ownership and Architecture

Domain data knowledge

Responsible for data quality

Shift from Push & Ingest to Serve & Pull



# Data Mesh Principles

Domain-Oriented Decentralized Data Ownership and Architecture

Data as a Product

# Data as a Product



 Discoverable

 Adressable

 Trustworthy

 Self-describing

 Interoperable

 Secure

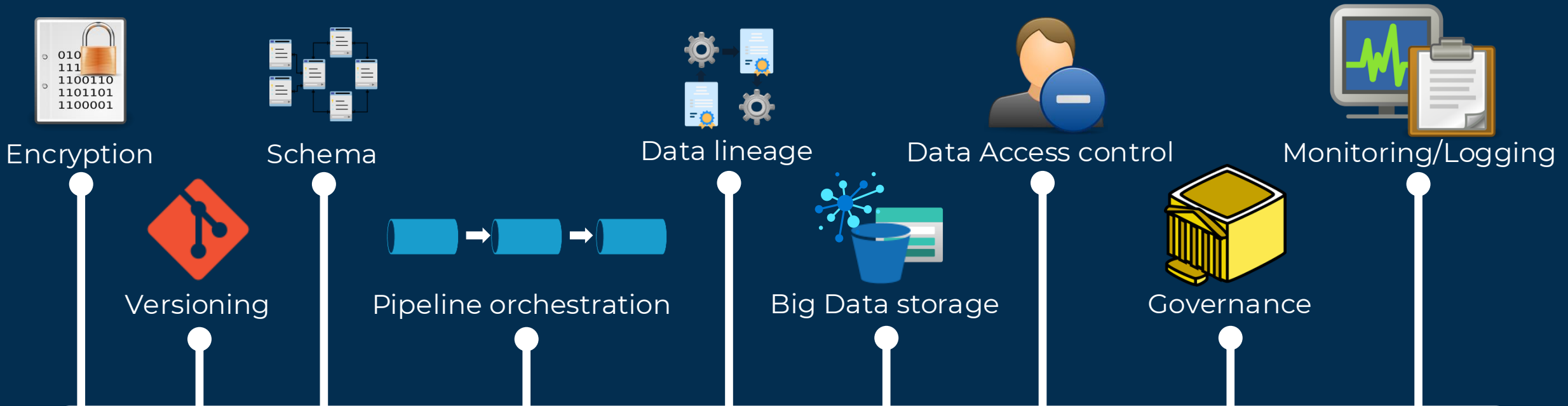
# Data Mesh Principles

Domain-Oriented Decentralized Data Ownership and Architecture

Data as a Product

Self-Serve Data Infrastructure as a Platform

# Self-Serve Data Infrastructure as a Platform



Domain agnostic Data Infra as a Platform

**Success criteria: Lowering lead time to create a new data product**

# Data Mesh Principles

Domain-Oriented Decentralized Data Ownership and Architecture

Data as a Product

Self-Serve Data Infrastructure as a Platform

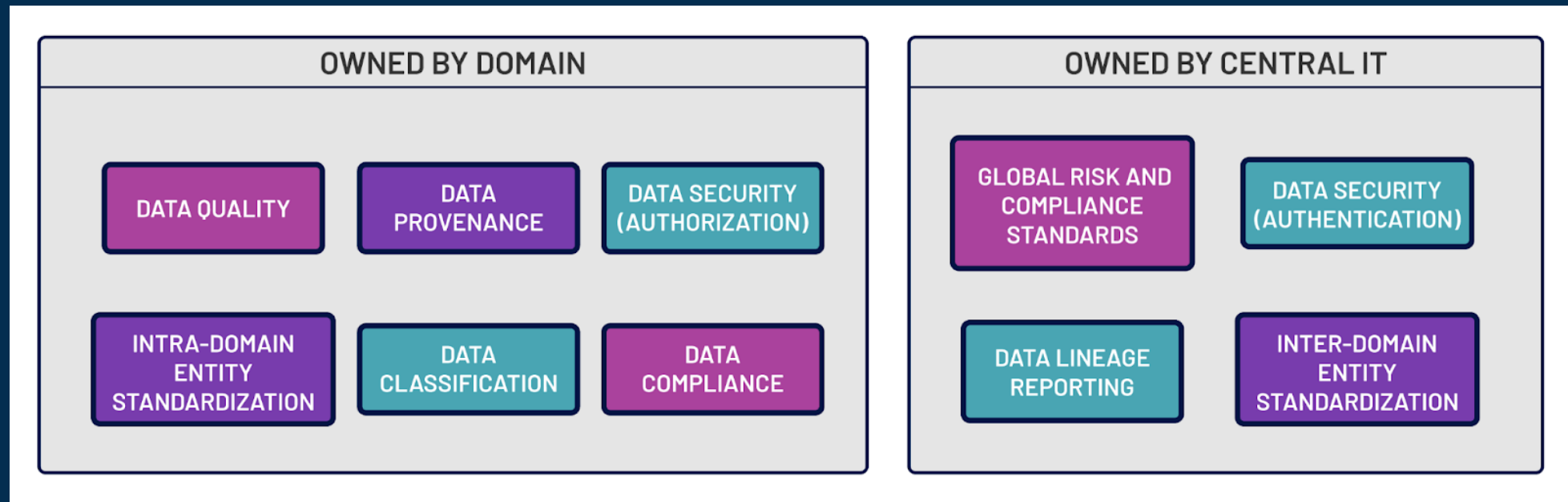
Federated Computational Governance

# Federated computational governance

Shared responsibility between domains and central IT

Focus on interoperability

Example:



# A closer look at data products

---

## FAIR principles

- Findable
- Accessible
- Interoperable
- Reusable



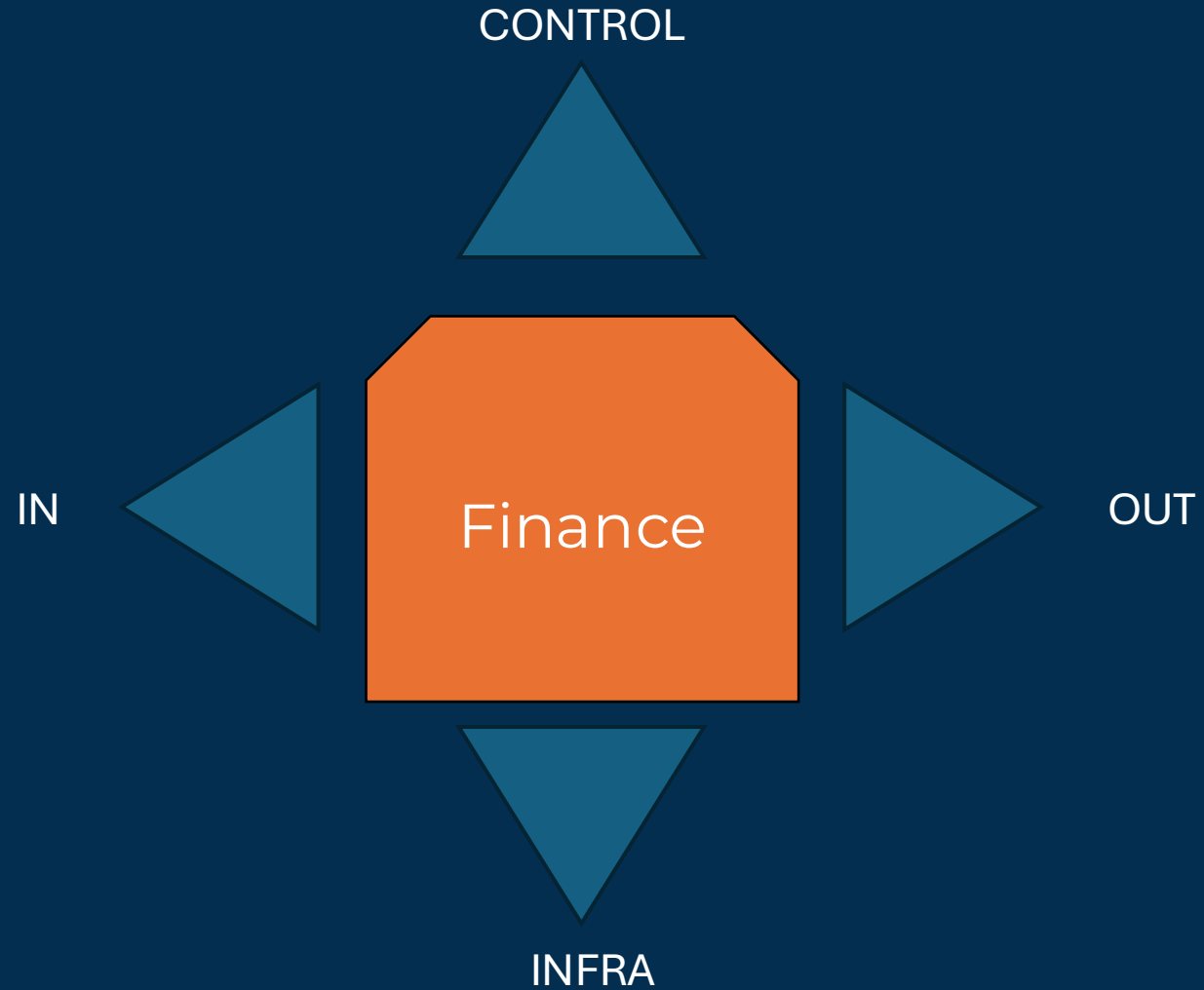
# Data product as shown in the data mesh



Hot-take:

The near religious adherence to data mesh principles have caused more confusion into data products than actually helping the successful adoption of them.

# Data product components



# Data product anatomy - Infra

The baseline of the product.

Storage

Compute

Git, pipelines and monitoring

Provided not by product team, but through common self-service infrastructure



# Data product anatomy - In

The in-port of the product.

Defines the sources used in building your product

Connection methods

Ingestion type (batch, delta, event...)

Ingestion frequency



# Data product anatomy - Out

The out-port of the product.

Who are my consumers?

How do I want consumers to reach my product?

SQL endpoint?

ODBC?

File export?

A report?

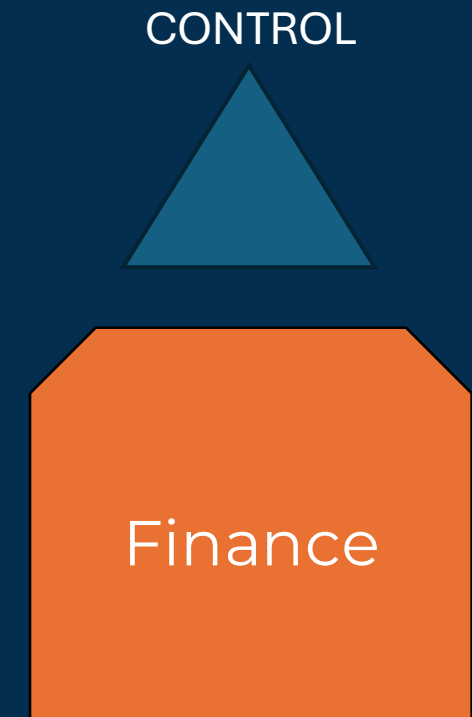
API?



# Data product anatomy - Control

The control-port of the product.

- Product description – “sales pitch”
- Metadata
- Data Quality
- Contracts
- Compliance – GDPR/PII/other regulatory reqs



# Data Contracts

Johan's two types of contract

## 1. Marketplace to Product – SLA

- Is there a defined owner
- Has data quality been quantified
- Have you registered the product in the data catalog
- Is your product rated in terms of compliance etc

## 2. Product to Consumer – EULA

- This product promises to update at this interval...
- This product will only support 2 versions of API
- This product can be used for all purposes/these limited purposes
- Anything built on top of this product is your problem
- Price



Finance

# Data Contracts

Open Data Contract Standard's definition

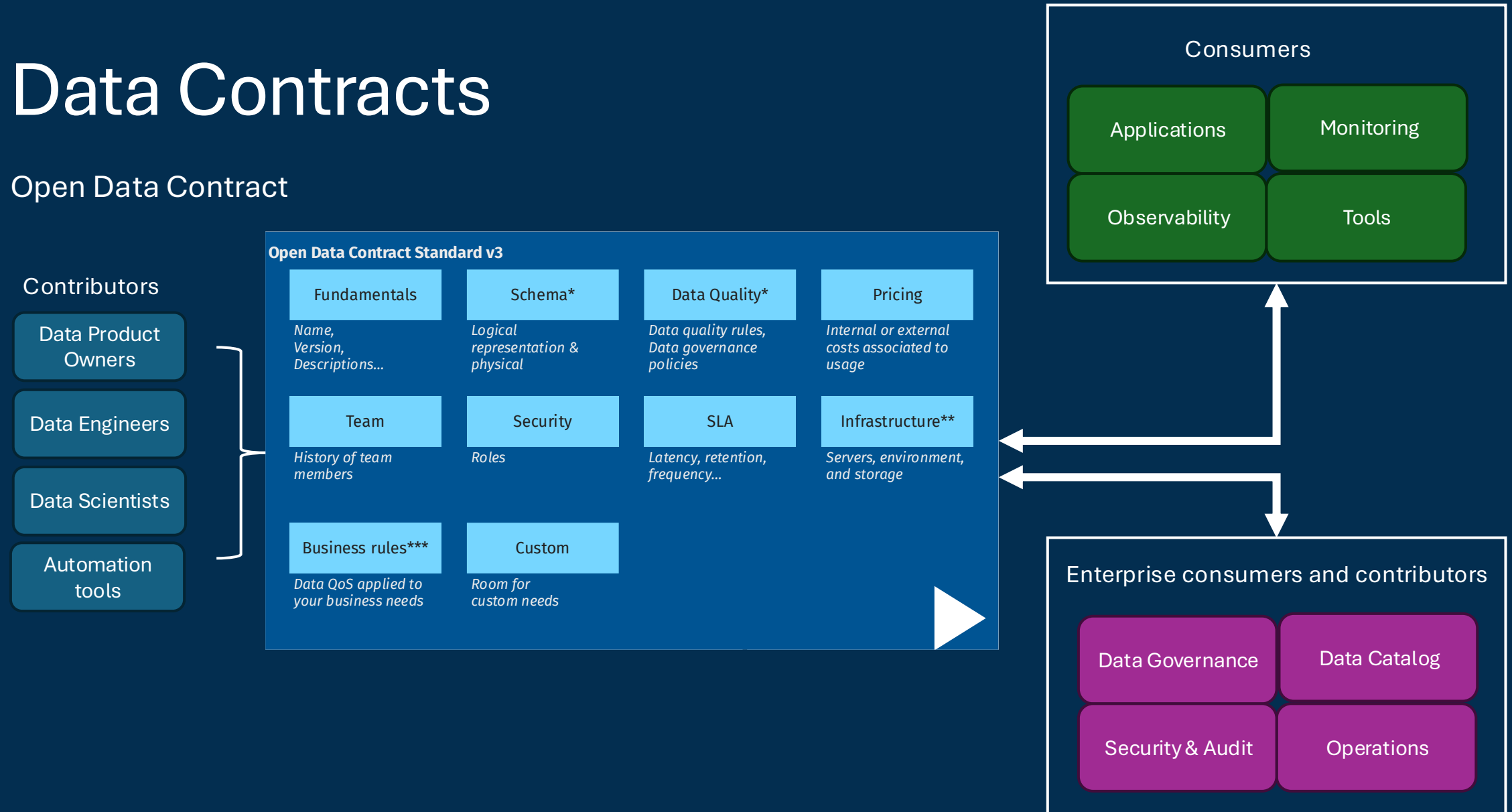
1. Fundamentals
2. Schema
3. Data quality
4. Support & communication channels
5. Pricing
6. Team
7. Security
8. Service-level agreement (SLA)
9. Infrastructures & servers
10. Custom properties.



Finance

# Data Contracts

## Open Data Contract



# Data product

- Now we have our products – but how do consumers find them?
- Data Catalog
- Marketplace
  
- How do we enforce a common governance model?
- How do we align on a common governance model?
  
- Computational Governance is the key word



Finance

# Data product

There are products out there to support you – i.e. Witboost

- Provides a marketplace
- At the core a policy enforcement engine
- However, forces you to build connectors that they don't support

We ended up with building our own solution using Backstage.io

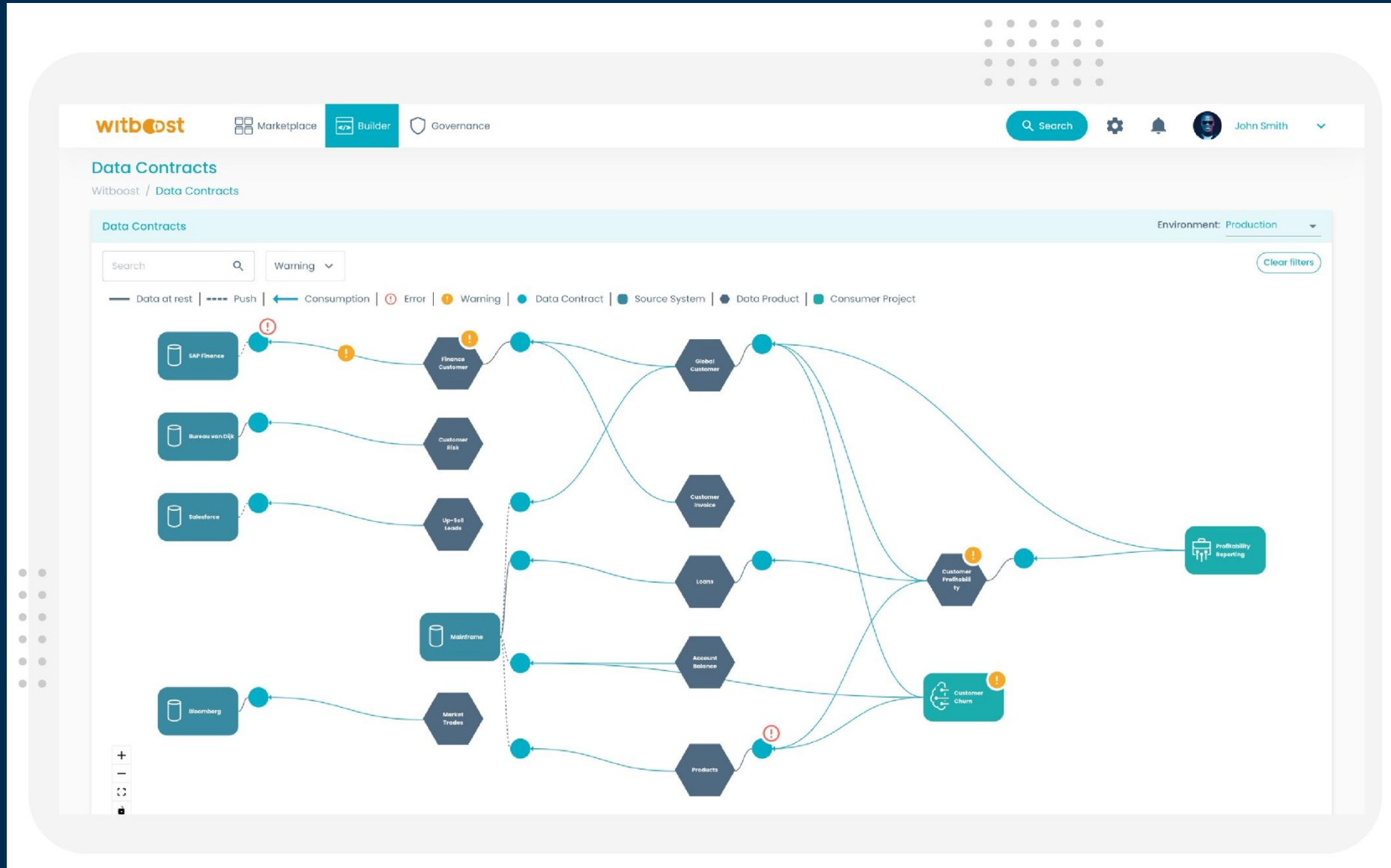
However – at the heart of any data product implementation are the people and the organization.

No company will become data driven or data product focused if there is no buy-in and support from top level execs – but also with the will to enforce support at all levels of the organization



Finance

# Data product



Search

Filters

[Clear filters](#)

▼ DOMAIN

Select Domain or Sub-dom... >

EMPTY DOMAINS

Show Empty Domains

▼ DATE

Select date range >

SYSTEM TYPE

- Data Product
- Consumer View
- Dashboard
- ML Model
- Staging Dataset

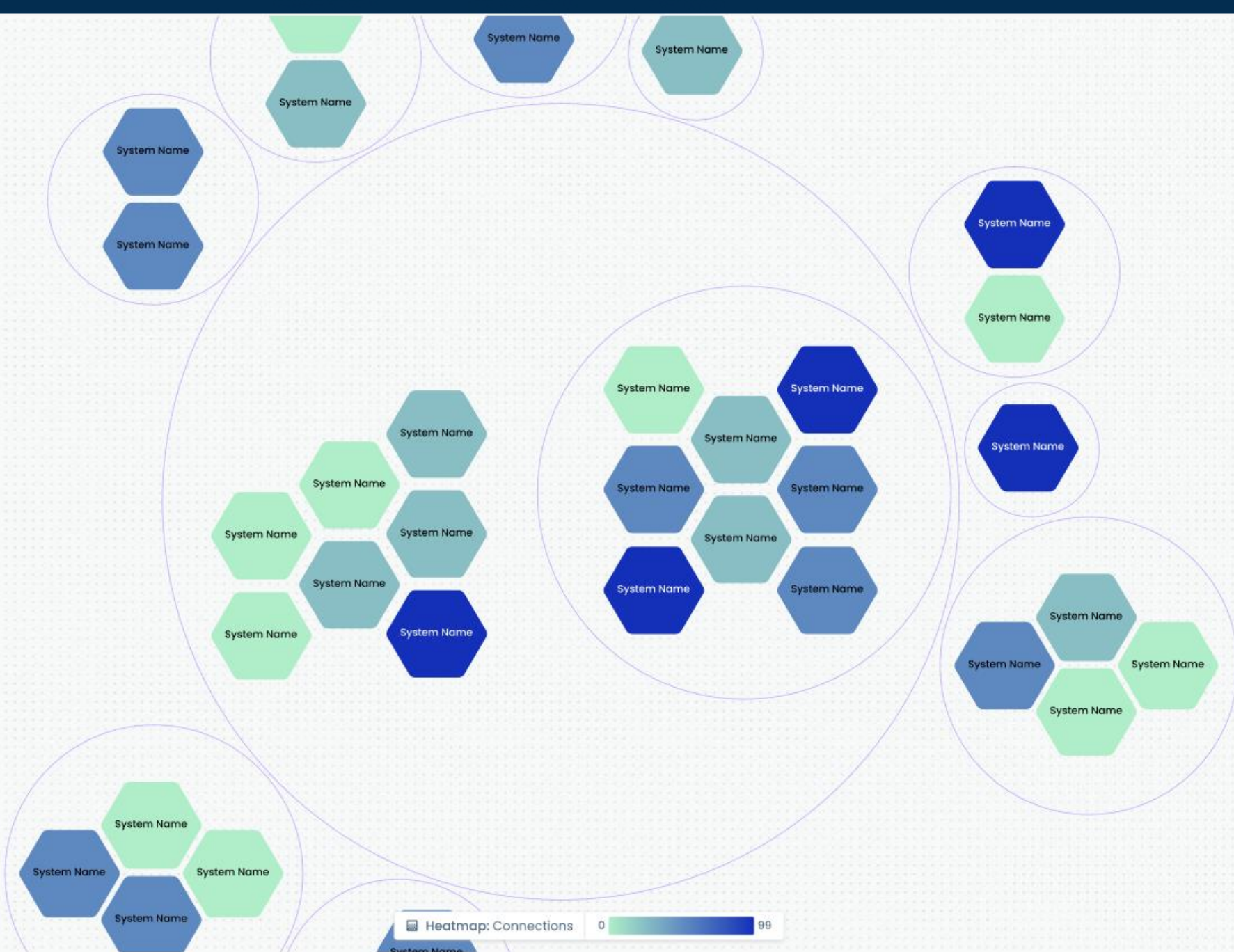
[Show all 25](#)

▼ PRODUCT OWNER

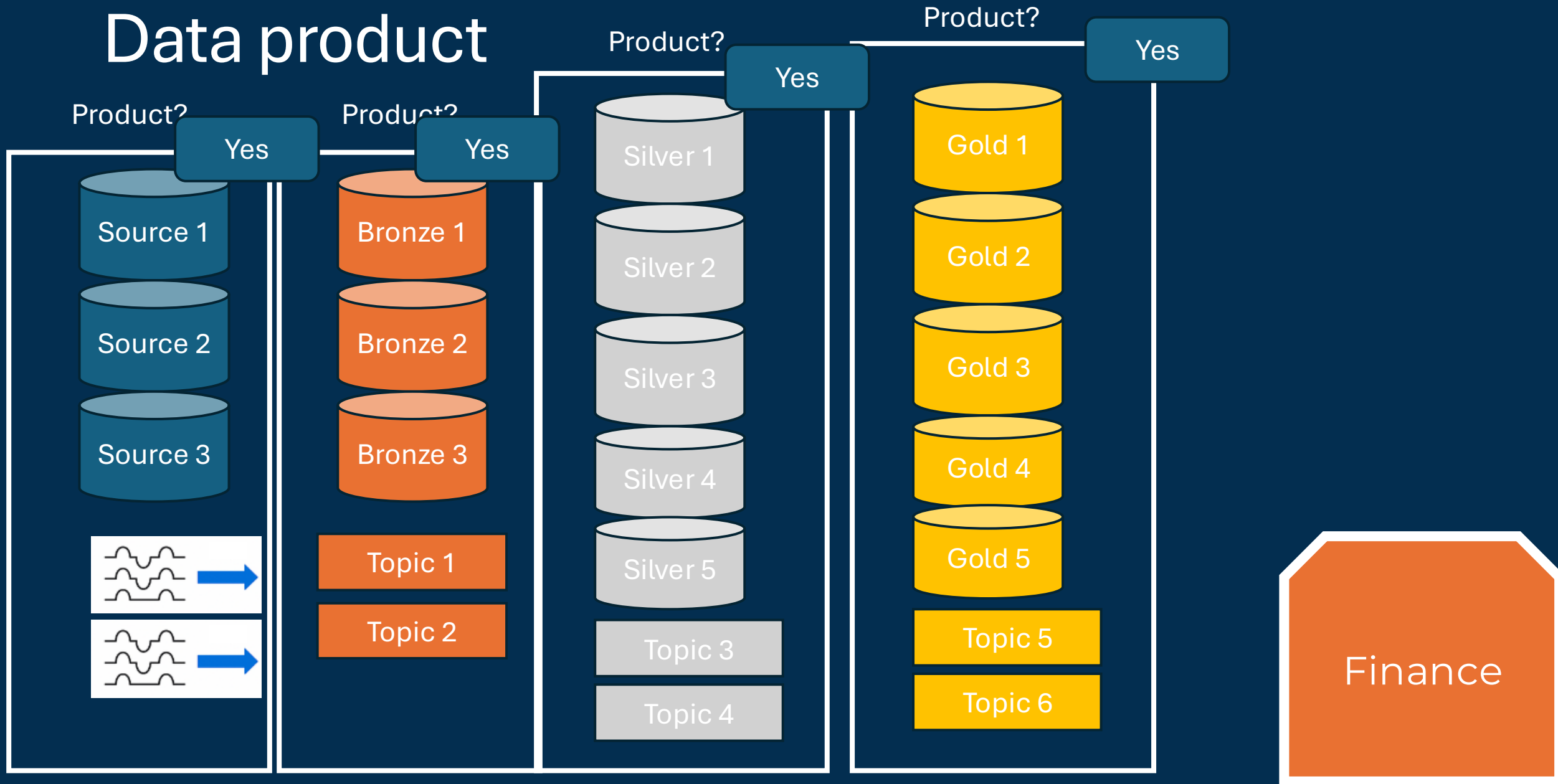
Search User

[Clear](#)

- Tracy Heathcote
- Ms. Tomas Kunde
- Joy Treutel III
- Marlene Schoen Sr.
- Lowell Swift
- Dallas Abshire
- Lila Klocko

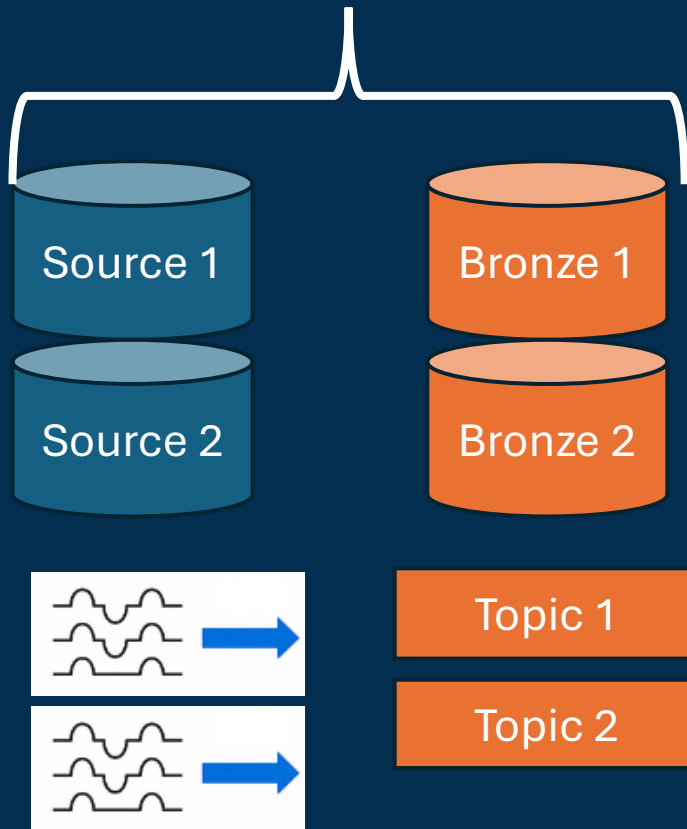


# Data product

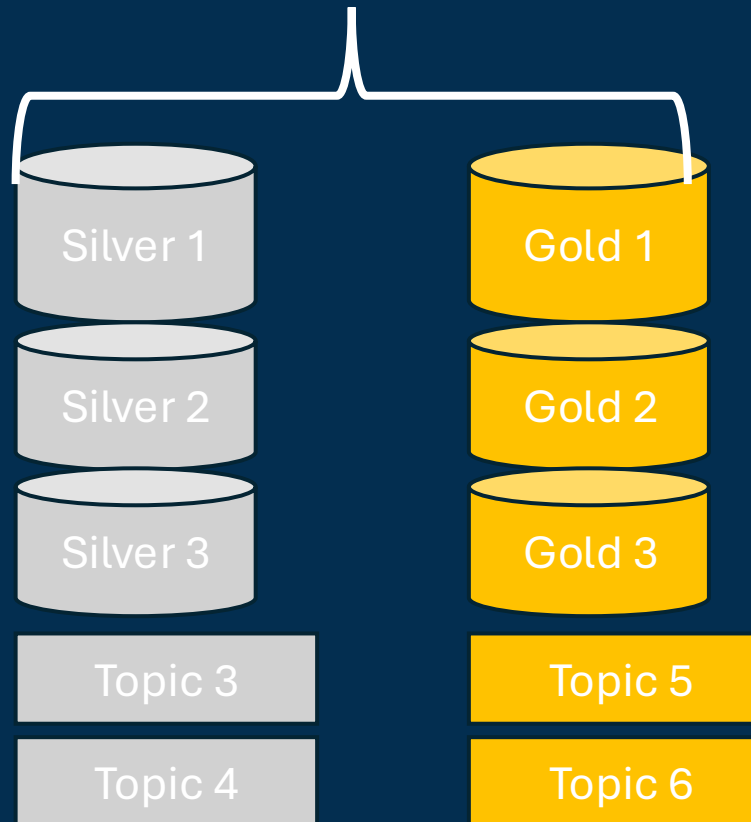


# Data product

Source Aligned

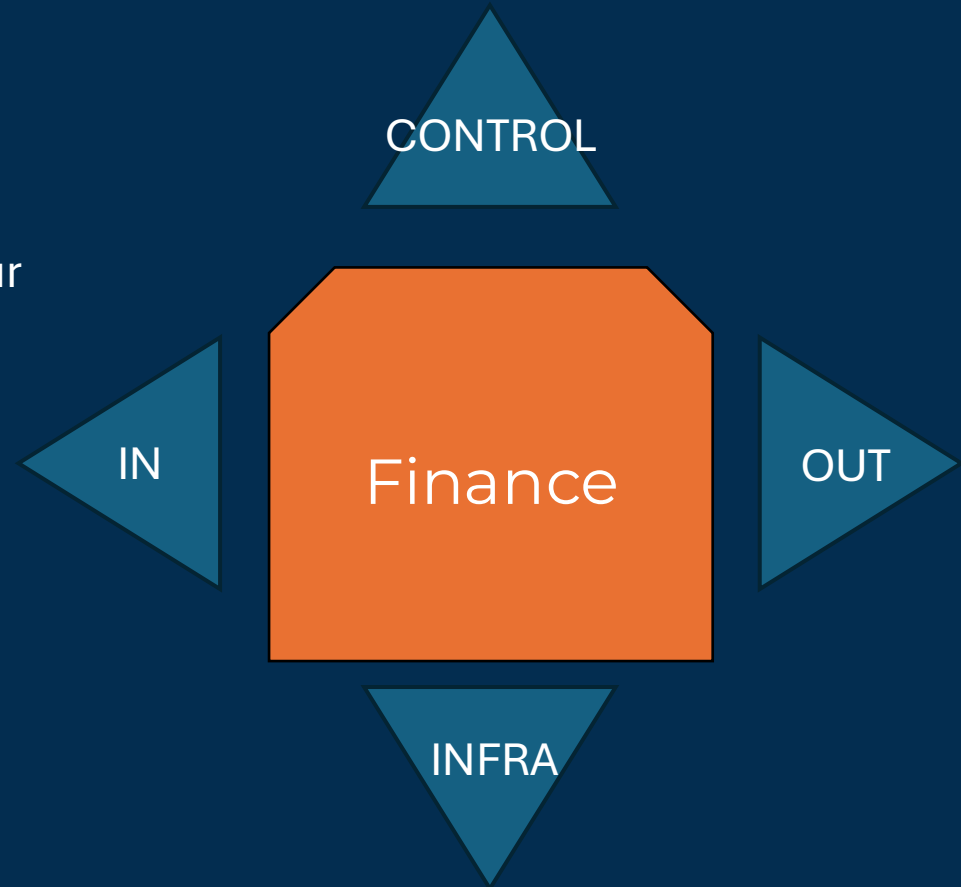


Consumer Aligned



# What we do

- CEO driving the change and pushing for change
- Created a new data strategy
- Established a central data governance team
- Removed system owners' ownership of data – it's not your data, it's the company's data.
- Building a self-service data platform for labs as well as production-ready data
- Domain-oriented teams





# A big thank you to our partners



**Loved it? Learned something? Tell us!**  
**Share your feedback in just 1 minute**




▣ The anatomy of a data product



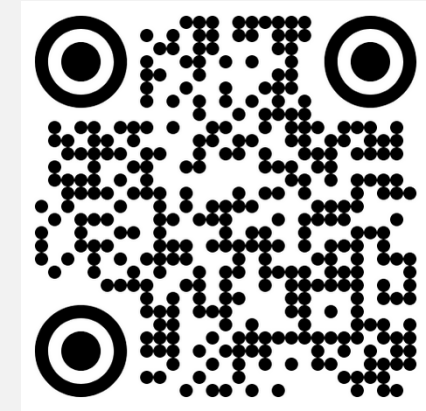
# Johan Ludvig Brattås

## Head of Data & AI @ Enora

 /johanludvig

 @sqlclause.no

 johan@brattas.no



GitHub

### Chronic volunteer

Co-organizer – DataSaturday Oslo

President – MDPUG Oslo

Frequent volunteer in general

### When not geeking out over new tech

Teaching coeliacs how to bake gluten free

Baking

Hiking

Gardening