Centralise or Federate

How to Scale Your Lakehouse

Centralise or Federate

How to Scale Your Lakehouse

Premium sponsors

VENZO_

INSPARI

valantic companies

itm8®

Eddytor™

dataon

Power your data

Standard sponsors

Backstage

Columbus

TheVIT

DATA - ANALYTICS - AI

twoday



Craig Porteous

Technology Director

analytics masterminds

databricks PARTNER PROGRAM SOLUTIONS ARCHITECT CHAMPION

Microsoft MVP Most Valuable Professional

https://craigporteous.com

https://github.com/cporteou

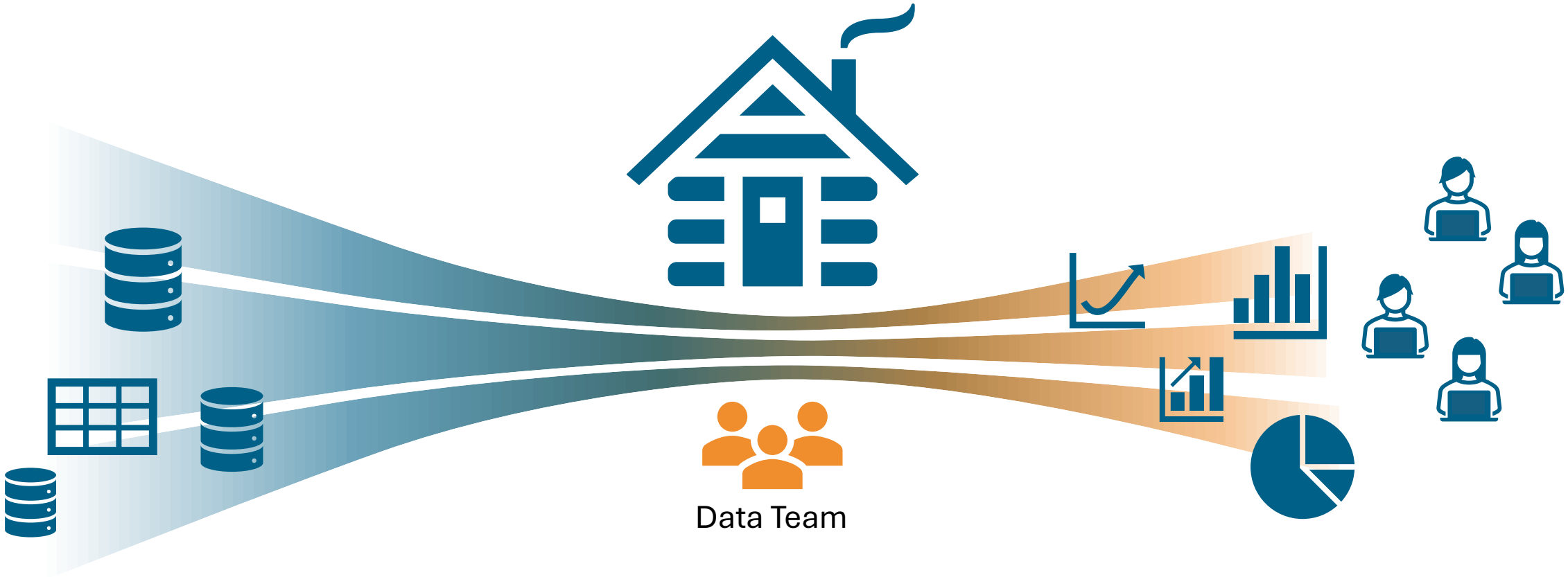
https://www.youtube.com/@cnporteous

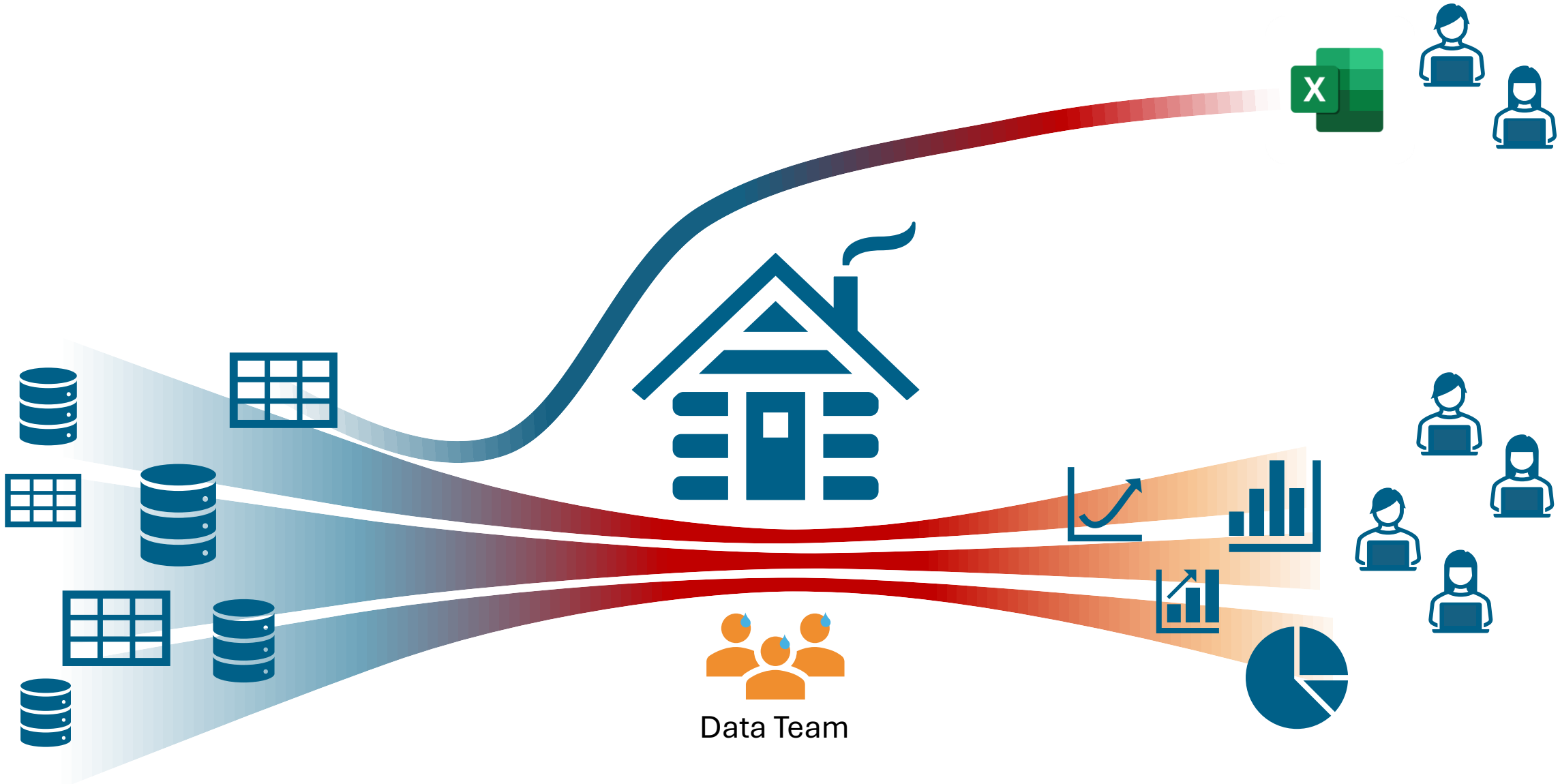
HEY FABRICATORS!

The Data Platform



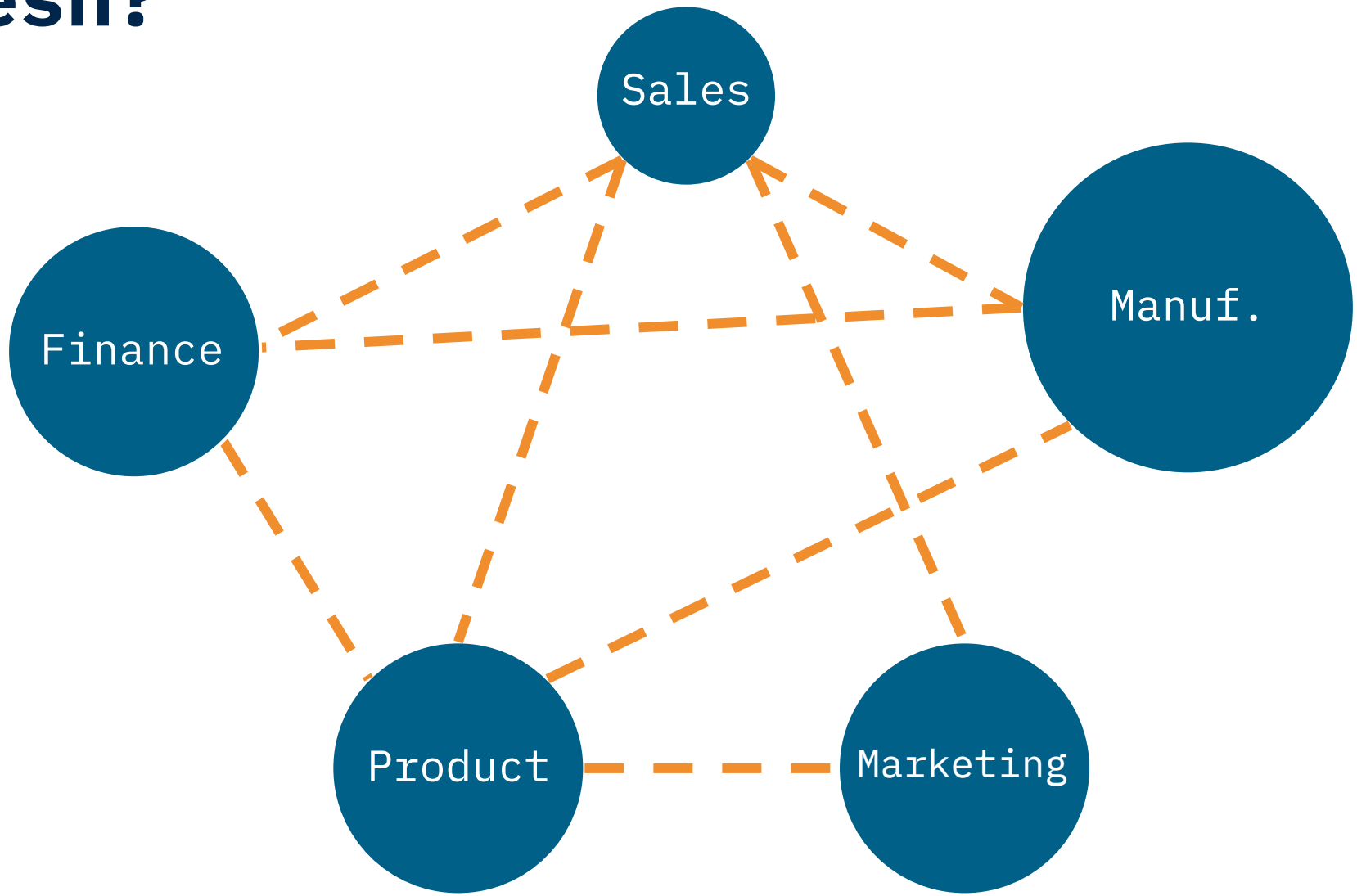
Growth & Demand



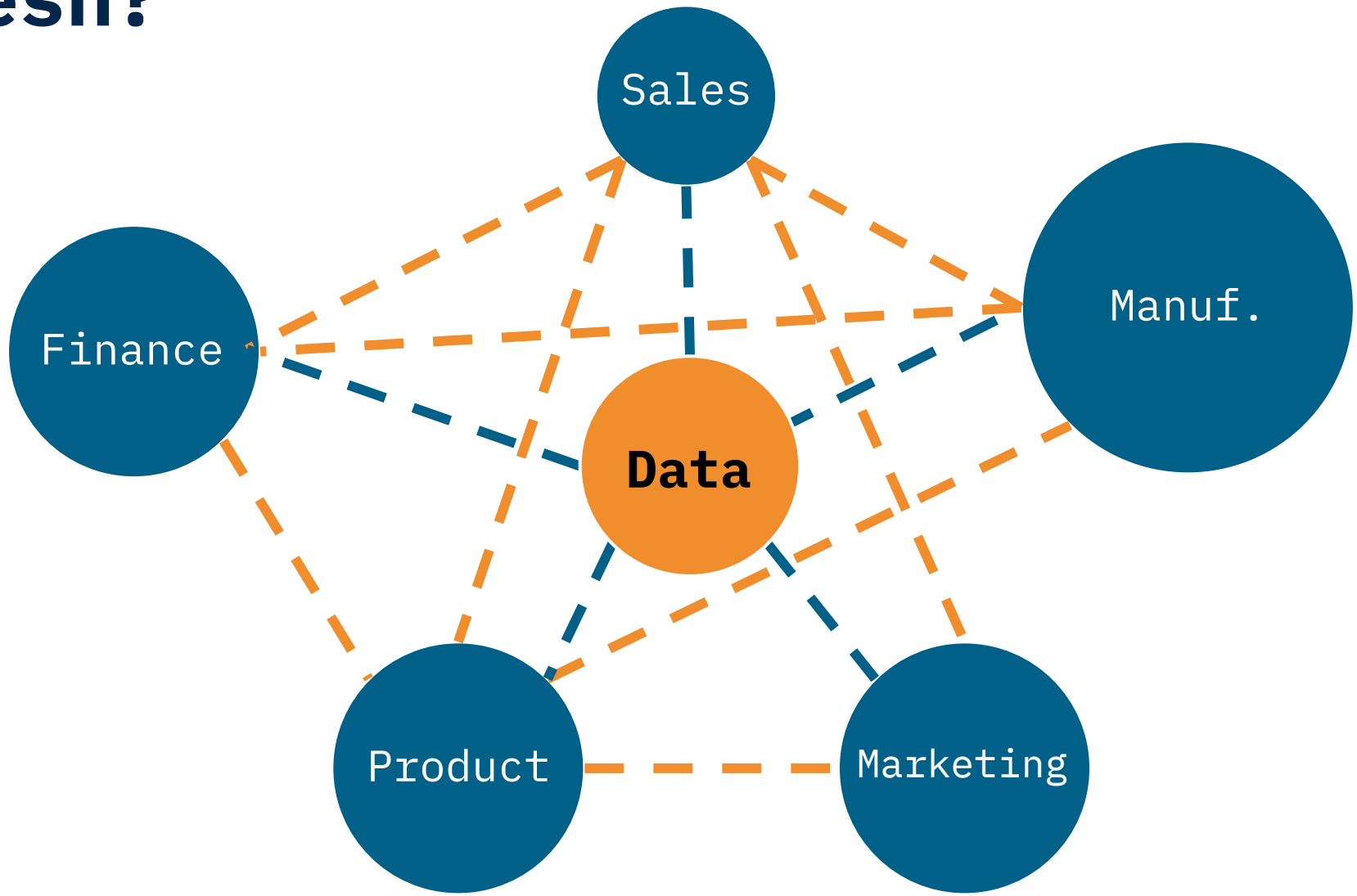

Data Team

What’s next?

Data Mesh?



Data Mesh?



Consistency
Control

Data Team

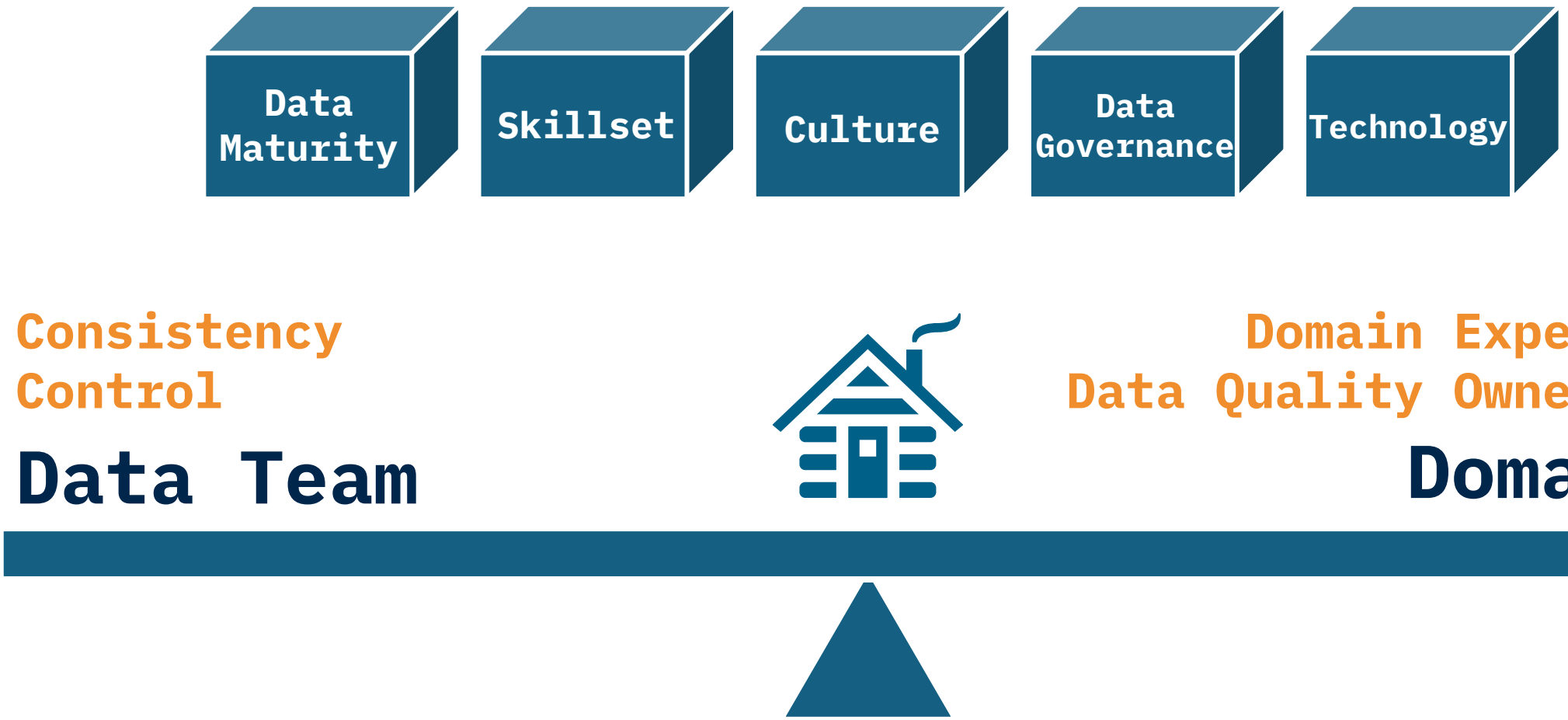


Domain Expertise
Data Quality Ownership

Domains



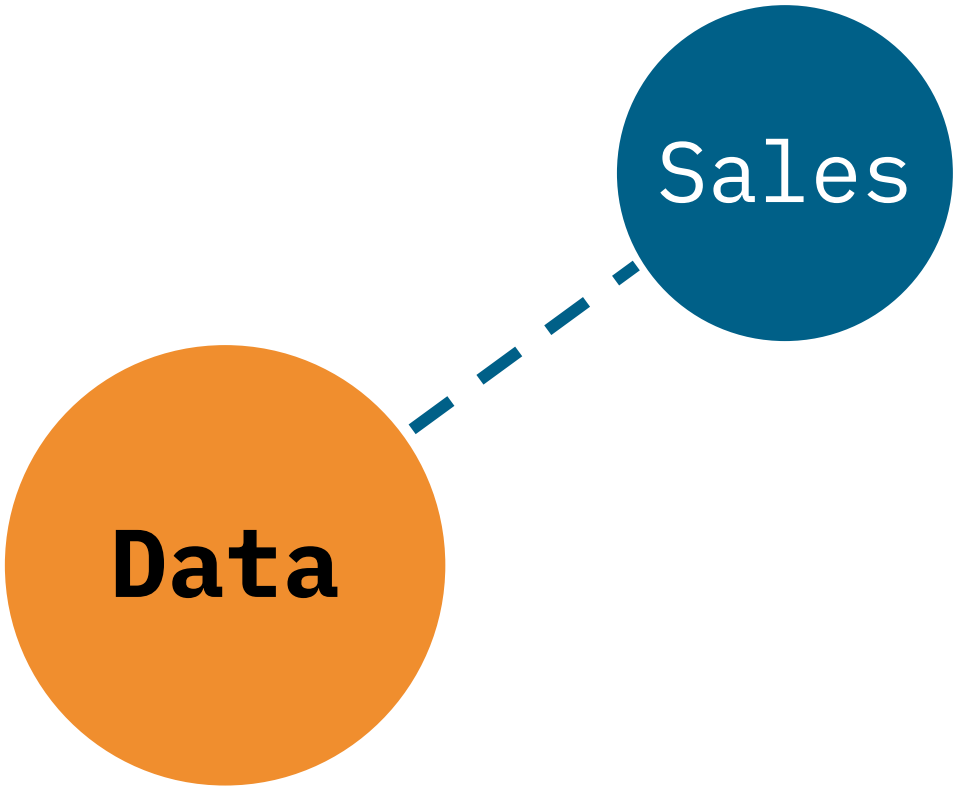



Data Maturity
Skillset
Culture
Data Governance
Technology
Consistency
Control
Data Team
Domain Expertise
Data Quality Ownership
Domains

How?

1 Create your Champion

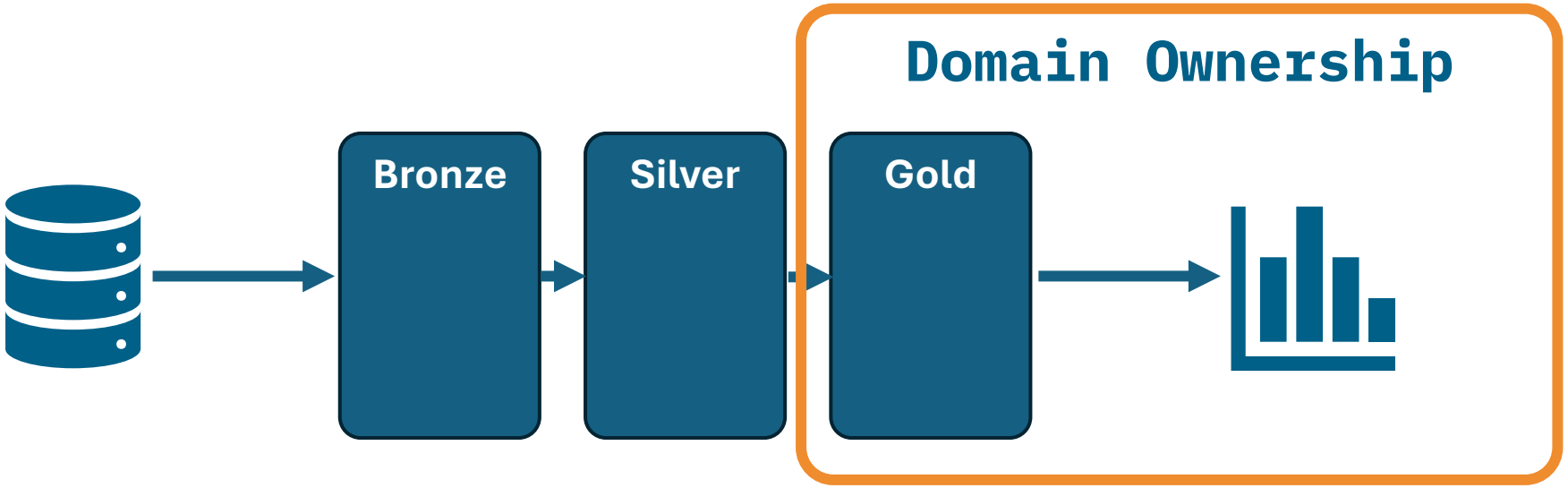
Culture



- Understand their needs
- Show them the benefits
- Build a playground
- Support & Train

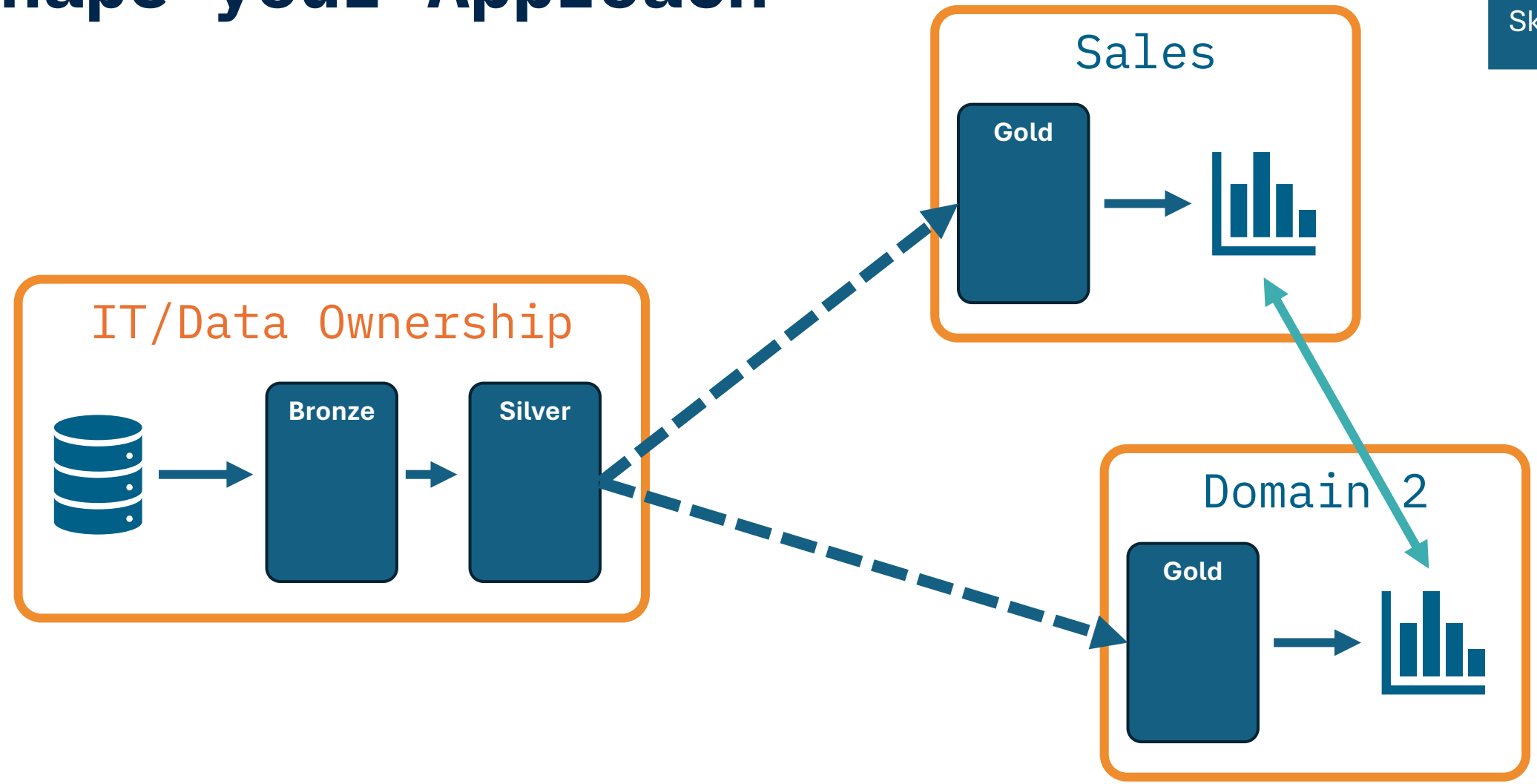
2 Shape your Approach

Skillset

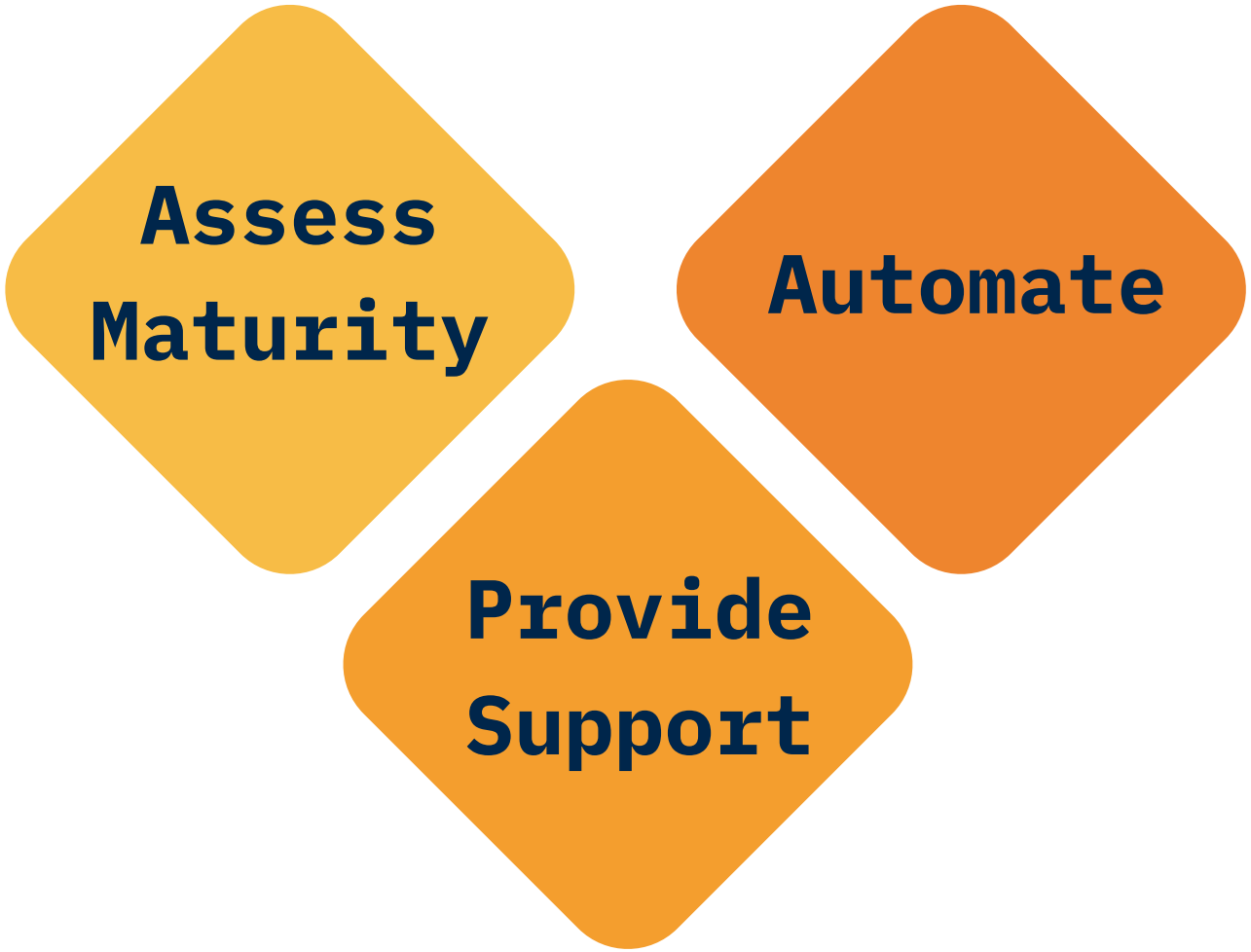


2 Shape your Approach

Skillset

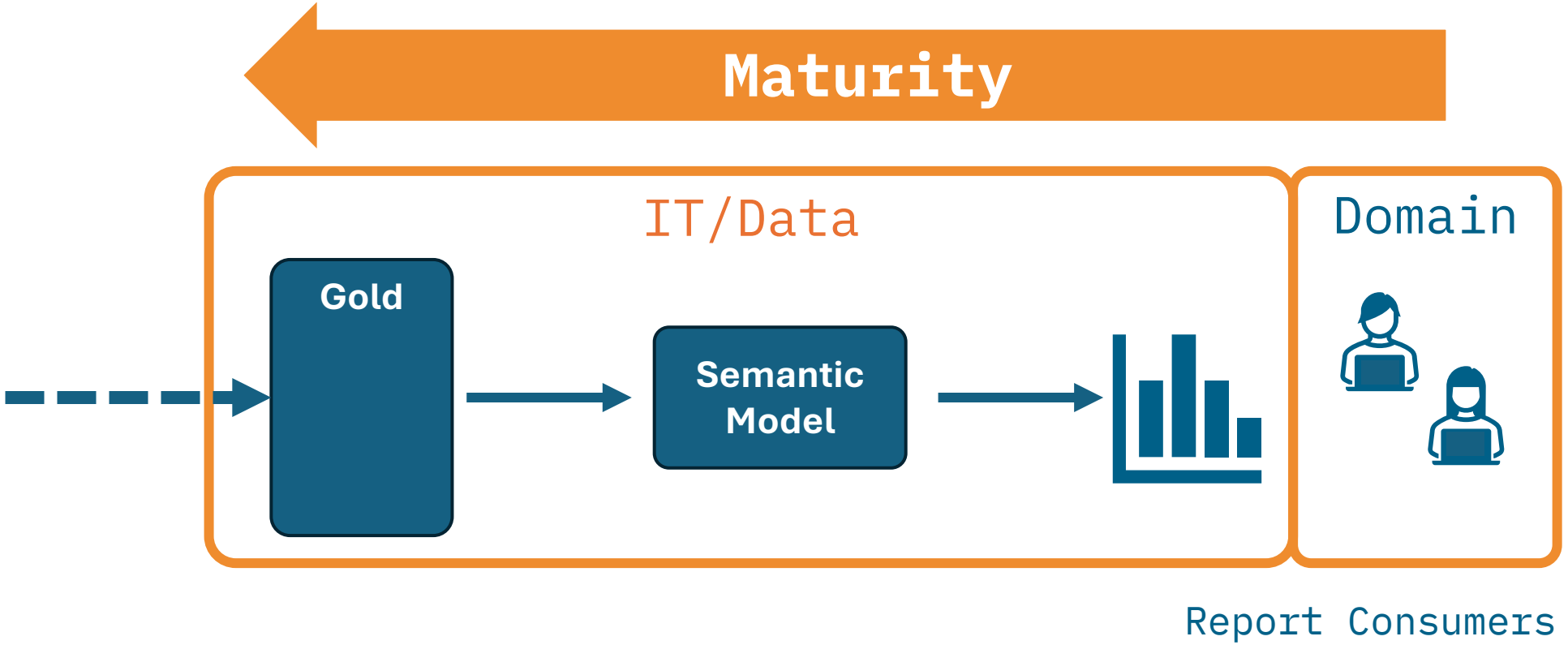


3 Build the Journey



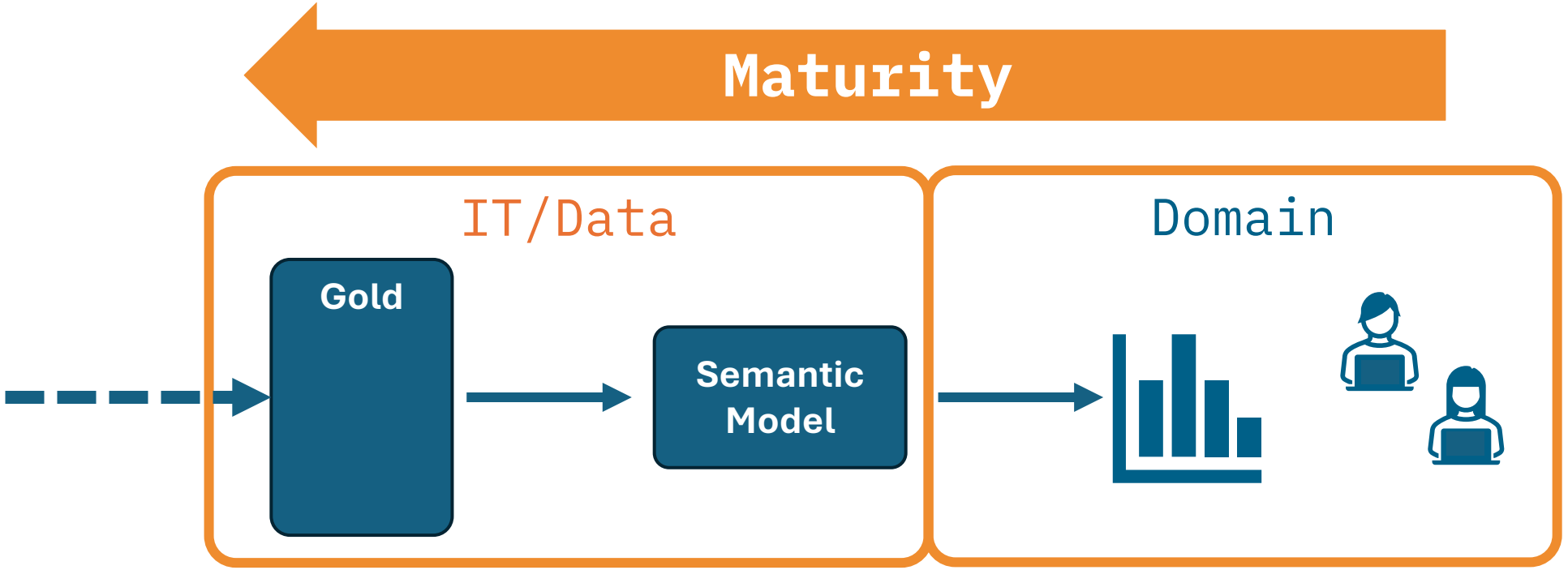
Assess Maturity

Data Maturity



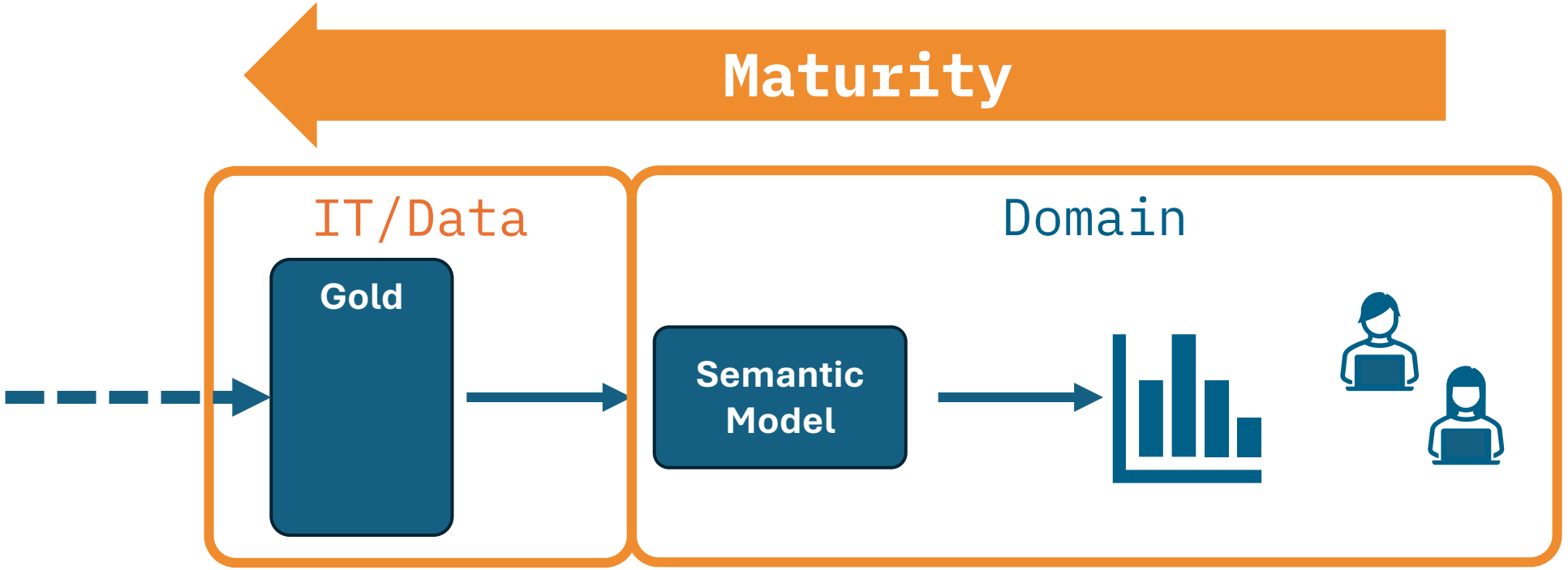
Assess Maturity

Data Maturity



Assess Maturity

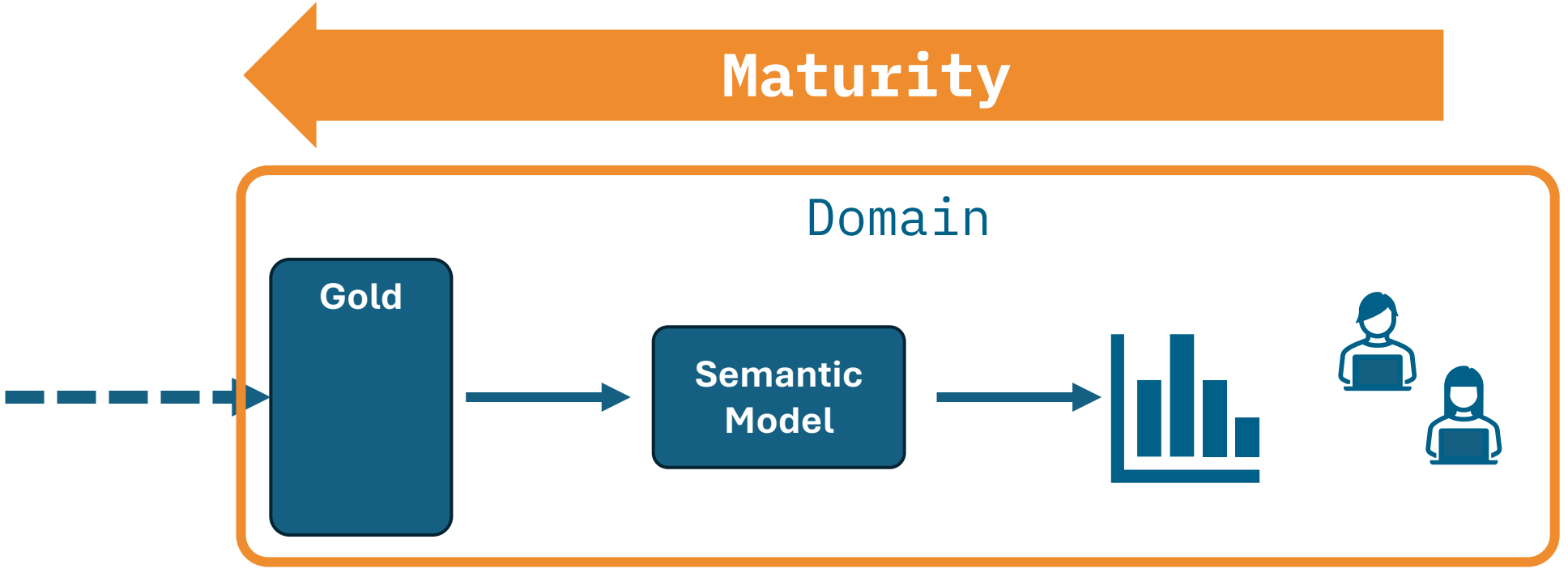
Data Maturity



Report Consumers
Report Building
Data Modelling

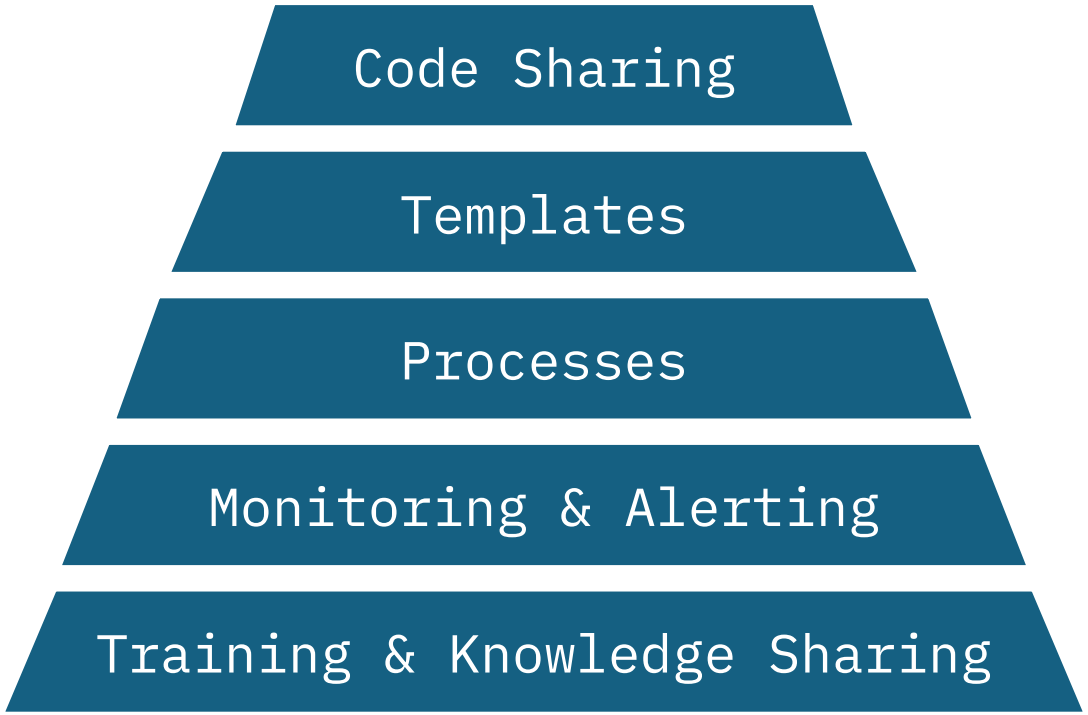
Assess Maturity





Report Consumers
Report Building
Data Modelling
Data Curation
Business Logic
Data Aggregations

Provide Support



Data Governance

Automate

Technology

- Infrastructure
- Security Roles
- Policies
- Tools



4 Control > Guardrails



Control		Guardrail
Approval Gates	→	Automated Policy Enforcement
Central Capacity & Costs	→	Budgets & Throttling
Performance Protections	→	Education and Training
Data Product ownership	→	Certification and discoverability

Centralise or Federate



Most domains haven't adopted self-service reporting

At least some domains showing strong ownership of data

No appetite to expand technical skills

Some technical skills exist in domains and appetite to learn

Ownership and Data Governance seen as a burden

Domains interested in data ownership and eager to champion way of working

Limited guardrails, policies, and engagement in place

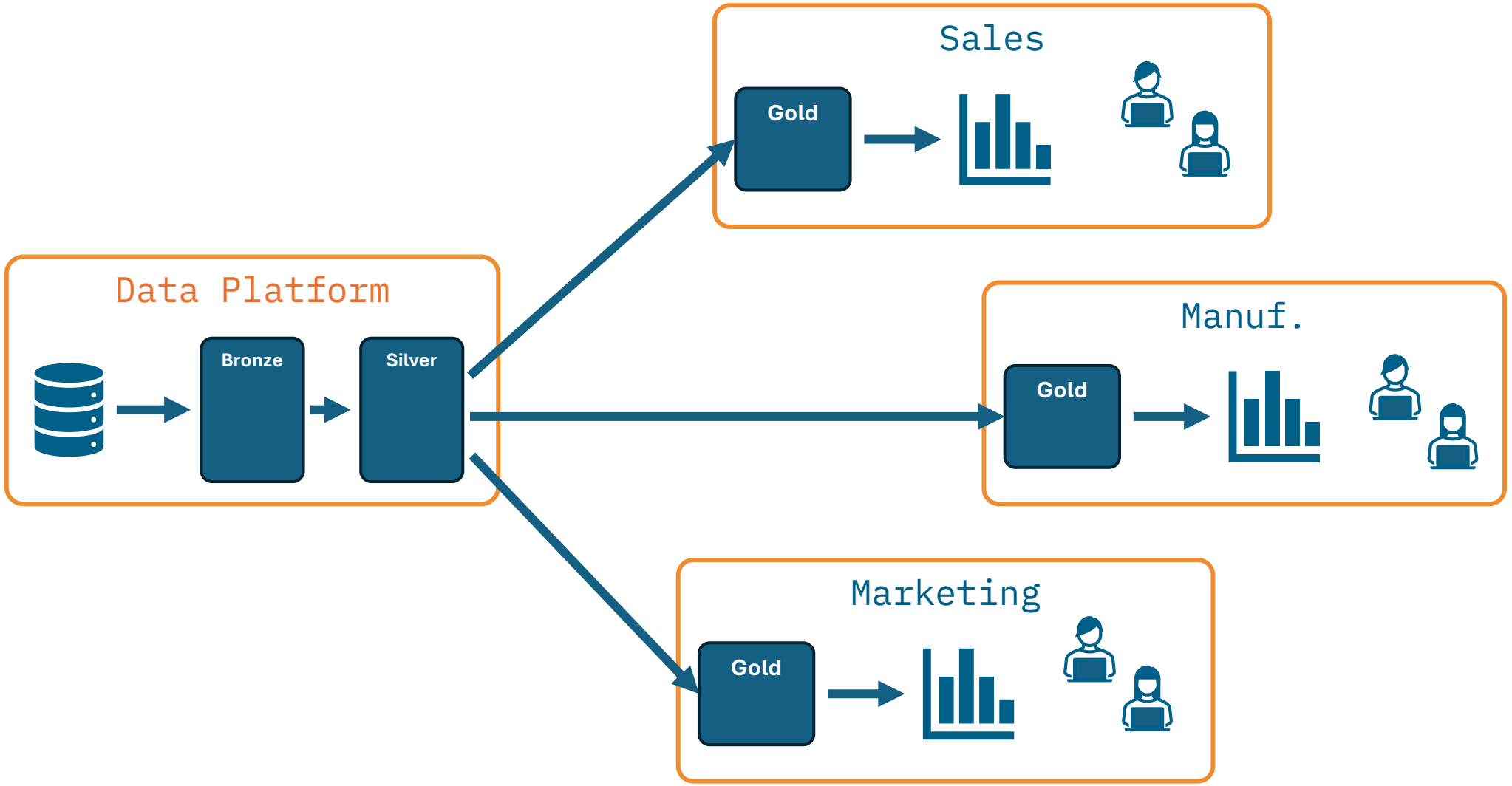
Mature practices in place already within the business

No automation, main data platform is a data warehouse

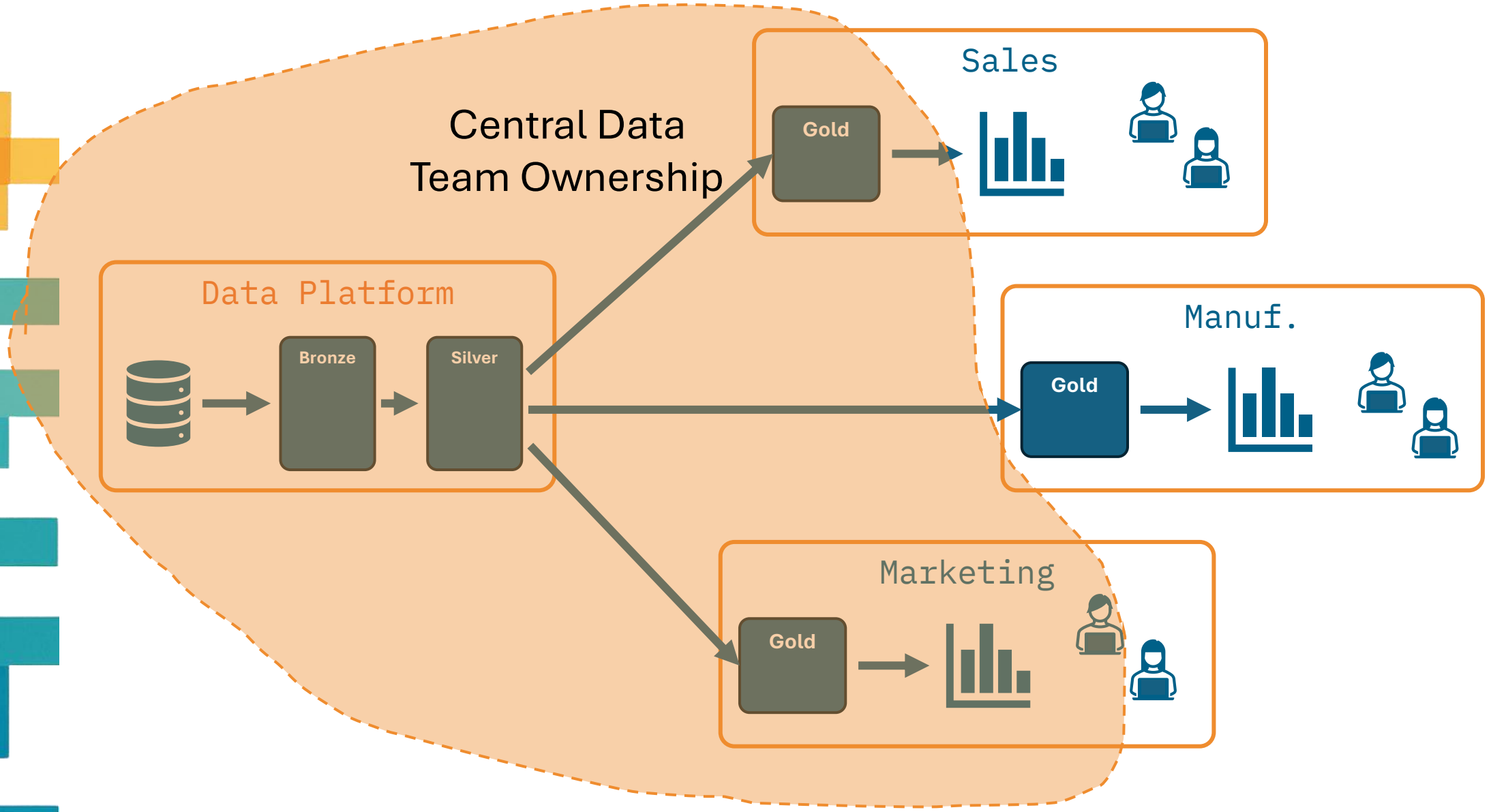
Adopted Data Lakehouse or similar architecture

What does it look like?

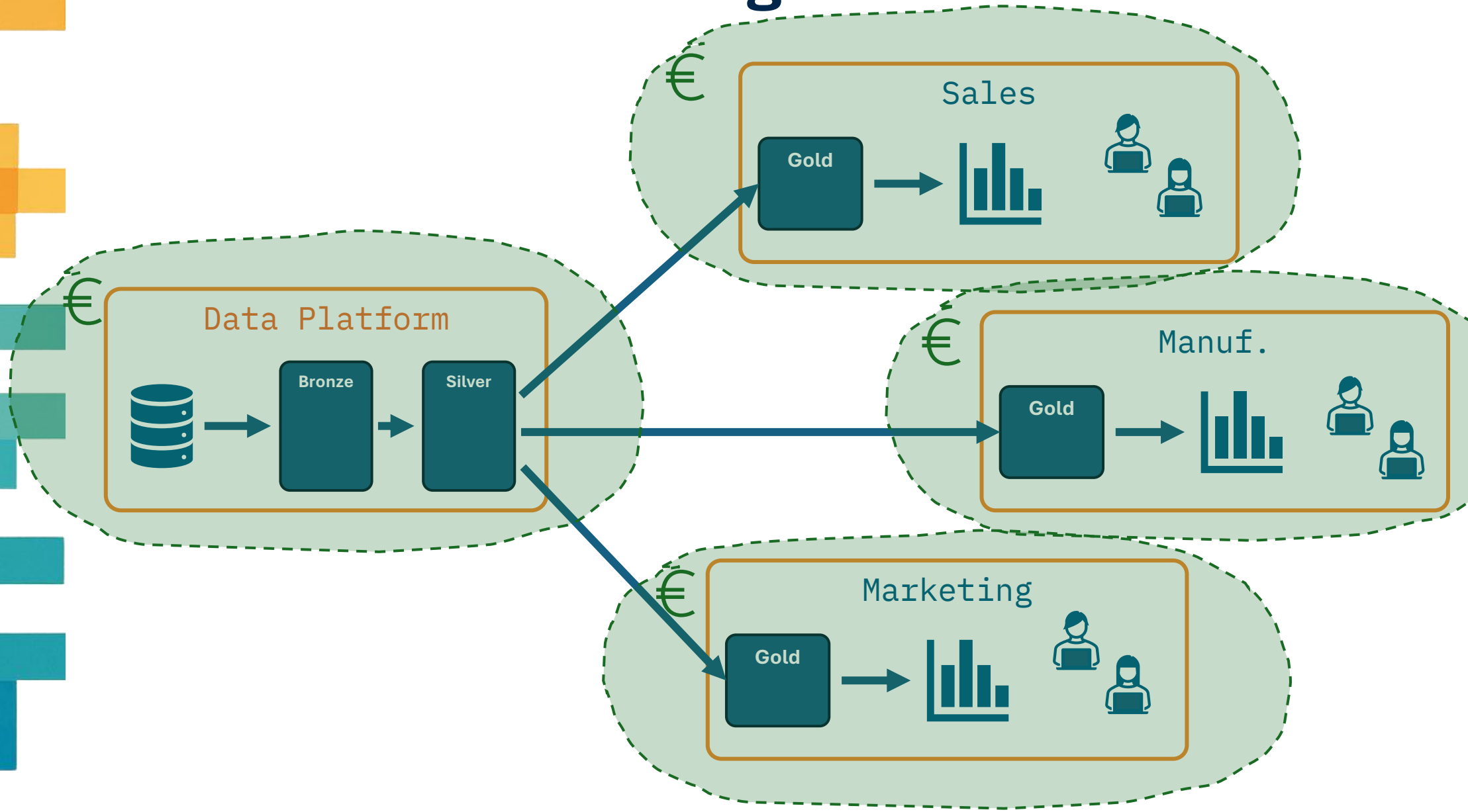
What we're working towards



What we're working towards

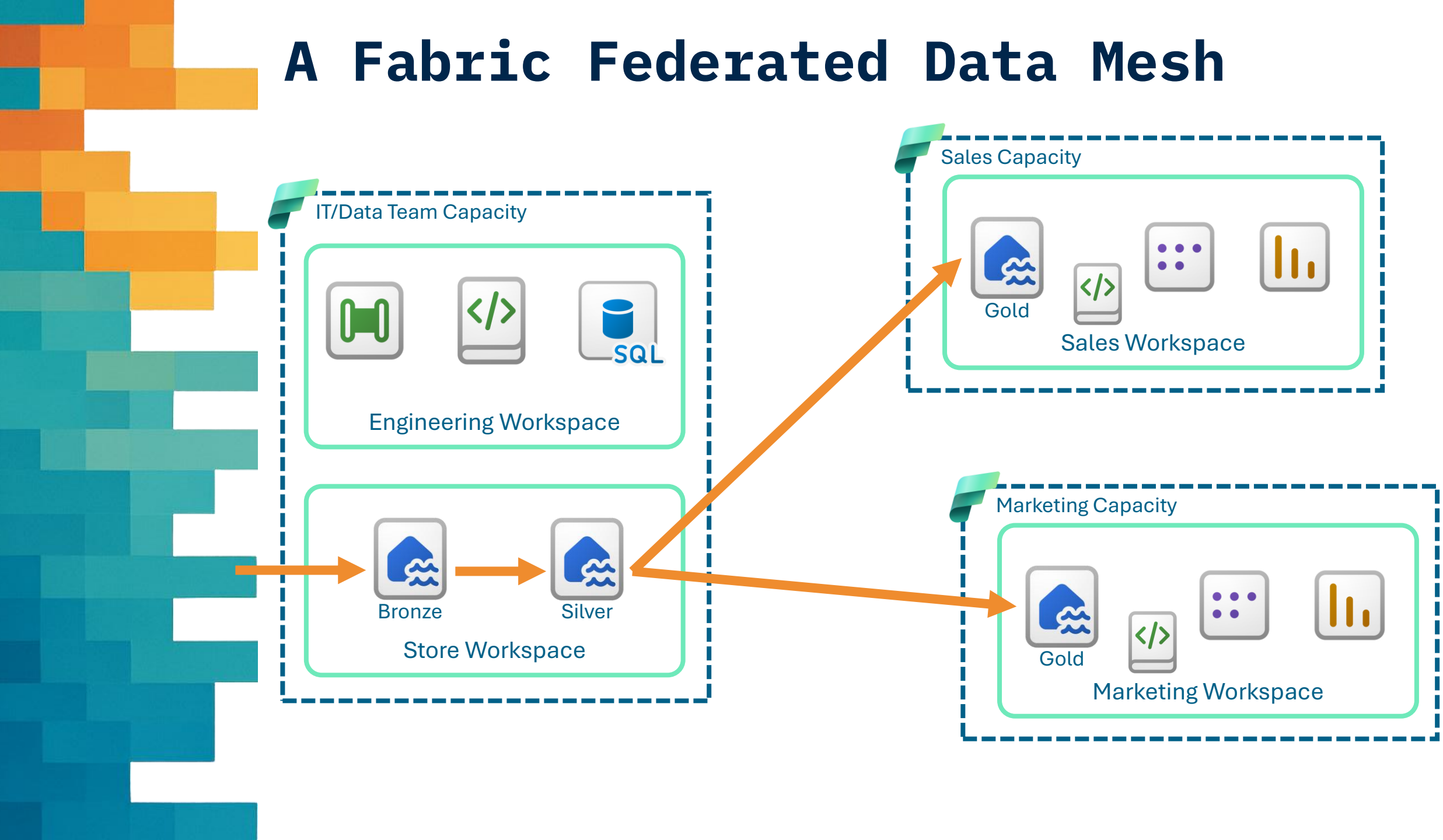


What we're working towards



The Fabric version

A Fabric Federated Data Mesh

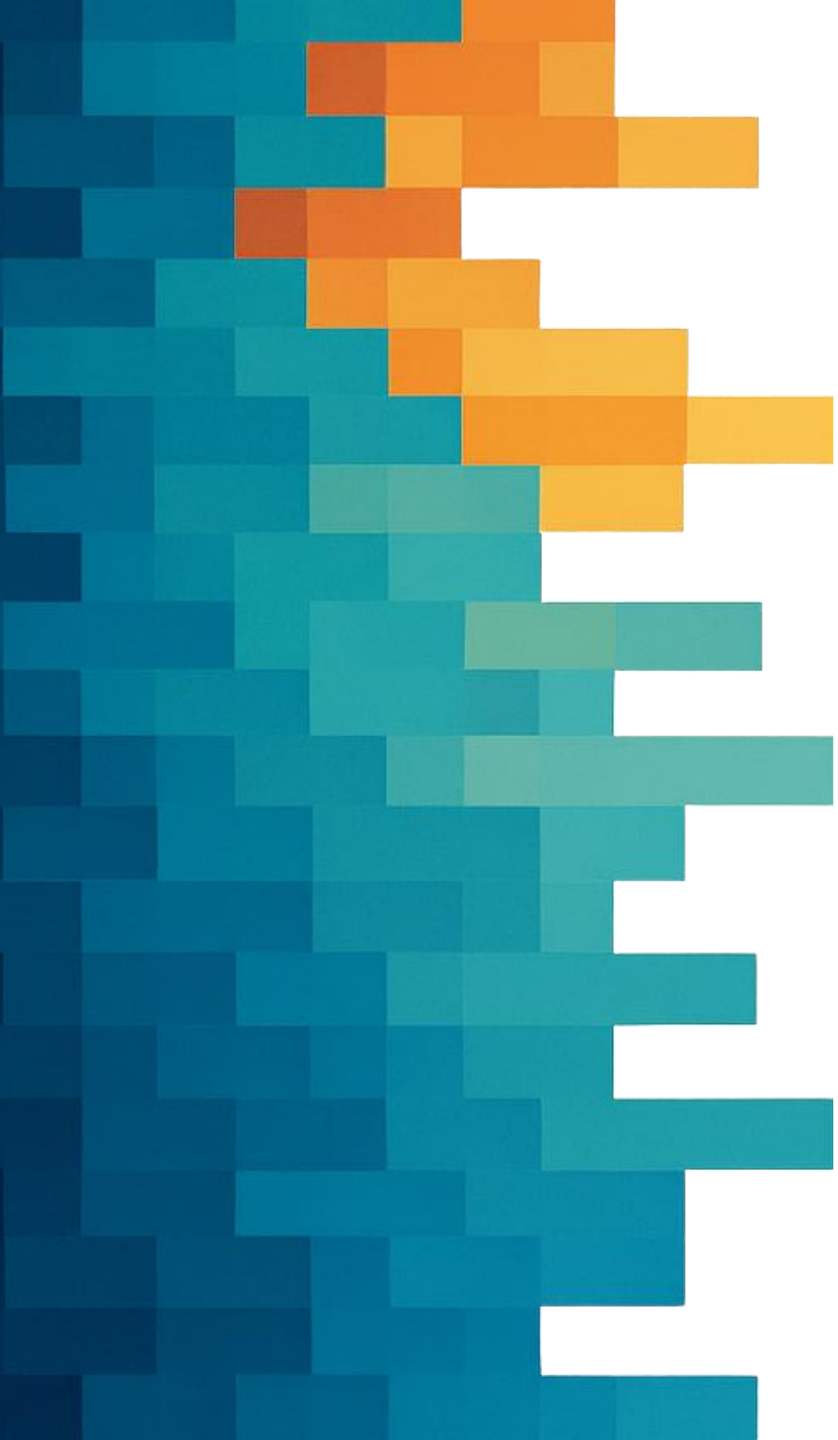


Take aways

Anti-Patterns

- “Fake federation” → still central team doing everything
- “Full decentralisation” → no standards, data chaos
- Tool-led architecture → letting the platform dictate the model
- Ignoring cost ownership in federated setups





1. Create your Champion
2. Shape Your Approach
3. Assess Maturity, Provide Support, Automate
4. Control into Guardrails

Evolve Intentionally

Thank You

Centralise or Federate - How to Scale your Lakehouse



https://craigporteous.com

https://github.com/cporteou

https://www.youtube.com/@cnporteous