Data Mesh

How Data Mesh can break data barriers and promote a collaborative culture

2023

Visit website

keyrus
make data matter

We are here to solve your business challenges through modern data solutions.
Contact us at sales@keyrus.co.za or +27 87 350 8860

Content

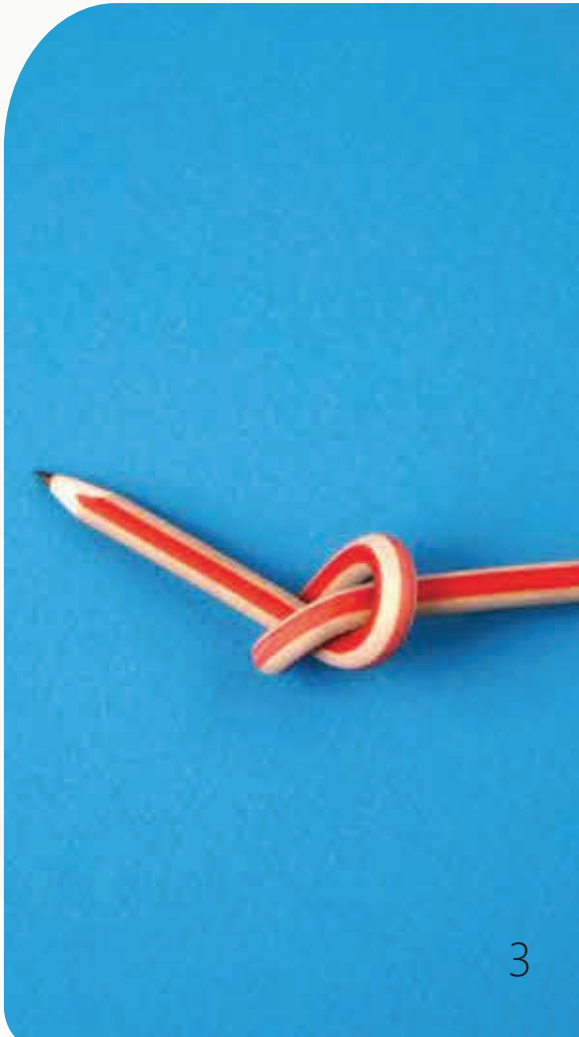
Optimal Governance vs Bottlenecks

Too often, centralized control hinders agility, as data ownership, quality, and access become bottlenecks.

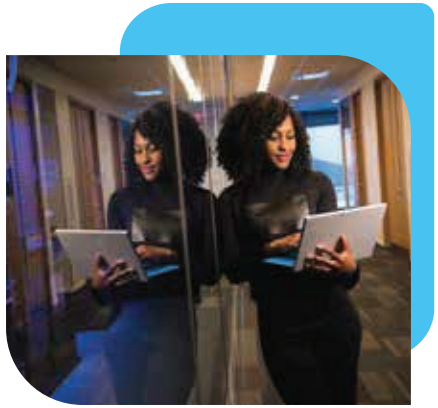
It’s not surprising than without clear ownership, it's challenging to assign accountability. This can lead to data inconsistencies and errors throughout the organisation.

Centralized data teams might not possess domain-specific knowledge necessary to ensure data accuracy and relevance. This gap can lead to suboptimal data quality and insights.

This sets the ground for Data Mesh.



A ‘data mesh’ is the latest evolution in data management.



Its evolution offer huge potential for enterprise organisations, with considerable long-term benefits for those ready and able to invest in their systems and adapt.

Zhamak Dehghani, its founder, describes a data mesh as:

“A decentralized, sociotechnical approach to share, access, and manage analytical data in complex and large-scale environments within or across organisations.”

But why is it important?

Addressing Governance challenges & **Bottlenecks**

Data Mesh solves the bottleneck that may be caused as a result of **central data warehouse / data lake implementation** by enabling business / domain teams to build and publish their own data products.

Data Mesh aims to break barriers between data producers and consumers, fostering a collaborative, data-driven, and innovative organisational culture.

Data Mesh hands over the ownership of building, managing and sharing data products to the business which is closest to the data. This enables effective data governance, data quality and data management.

It promotes federated governance by defining rules at a global/enterprise level, but giving autonomy to federated teams to implement as per their requirements

Data Mesh

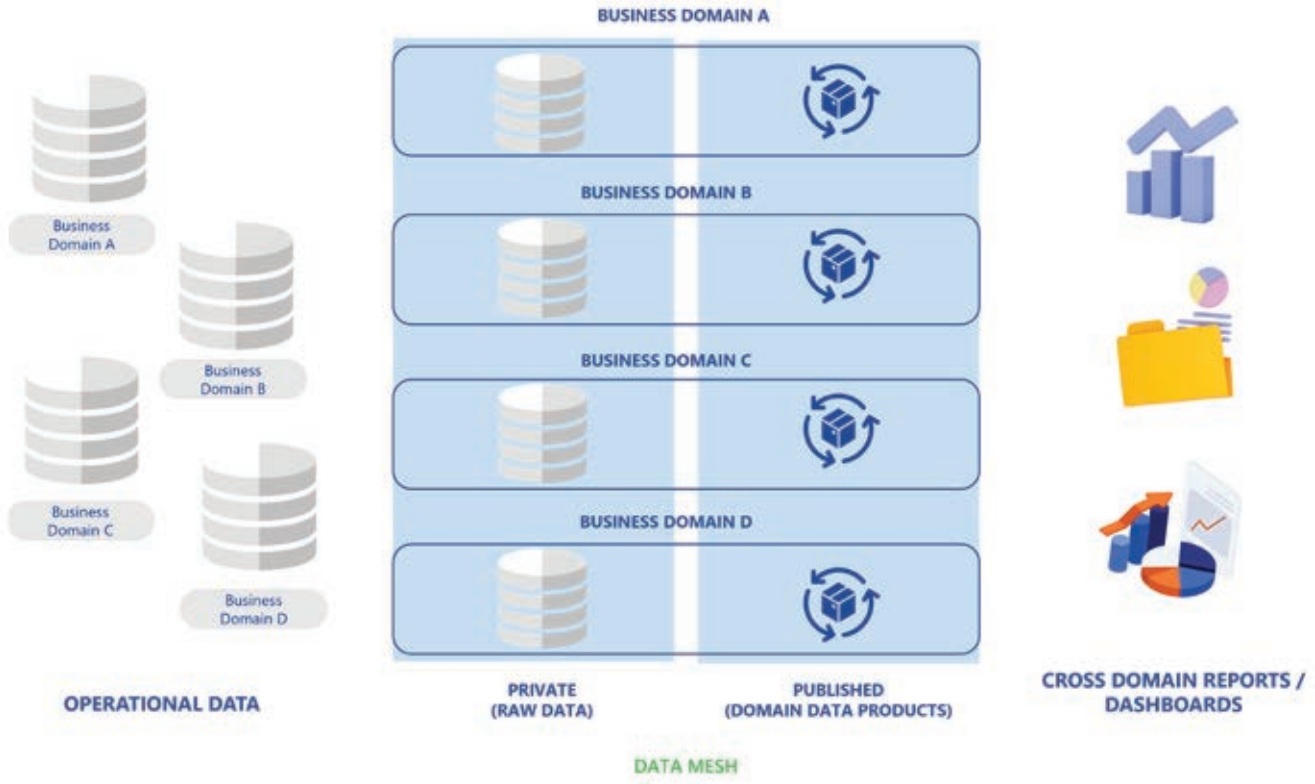
The data mesh enables business / domain teams to define, build, manage and share their data products. This is done by defining data domains and data ownership based on the enterprise data strategy and along seams of the existing organisation structure.

The business teams are enabled by providing them a self-serve data platform. The platform empowers business-teams to build, manage and share their own products with limited amount of cognitive knowledge and specialization.

Data product owners are encouraged to define data as a product with the focus on how data adds value to consumers.

The process becomes more autonomous and democratic, supported by strong governance and best practice to maintain security,

compliance, interoperability and consistency.
In this way, data mesh models have the potential to deliver significant benefits to the way data is managed and leveraged within large organisations for com-petitive advantage.



Data Mesh Characteristics

- Ownership of design, build and share of domain data products is given to the team which owns the data
- Business teams are accountable to publish the truth of their own domains
- Data mesh enablement team provides a data platform which enables multiple business teams to design, build and share their data products
- Clearly outlines principles which drive data mesh architecture and

Challenges

- Long term commitment
- Clearly outlined data strategy at an enterprise and domain level
- Certain level of data maturity to efficiently implement an enterprise level data mesh

Data Mesh
Guiding Principles

- Principle of domain ownership
- Principle of self serve data platforms
- Principle of data as a product
- Principle of federated computational

Principle of **domain ownership**

Data mesh is a decentralized approach where the responsibility of design, build and management of data products lies with the business domains closest to the data. Individual domains are res-ponsible for providing the truths of their business.

Principle of self serve **data platforms**

The platform must enable multiple domain teams to build, share and use data products in an autonomous way. The platform must support interoperability between different technologies and must not require propriety specialization of a technology vendor which is unscalable at an enterprise level.

Principle of data as **a product**

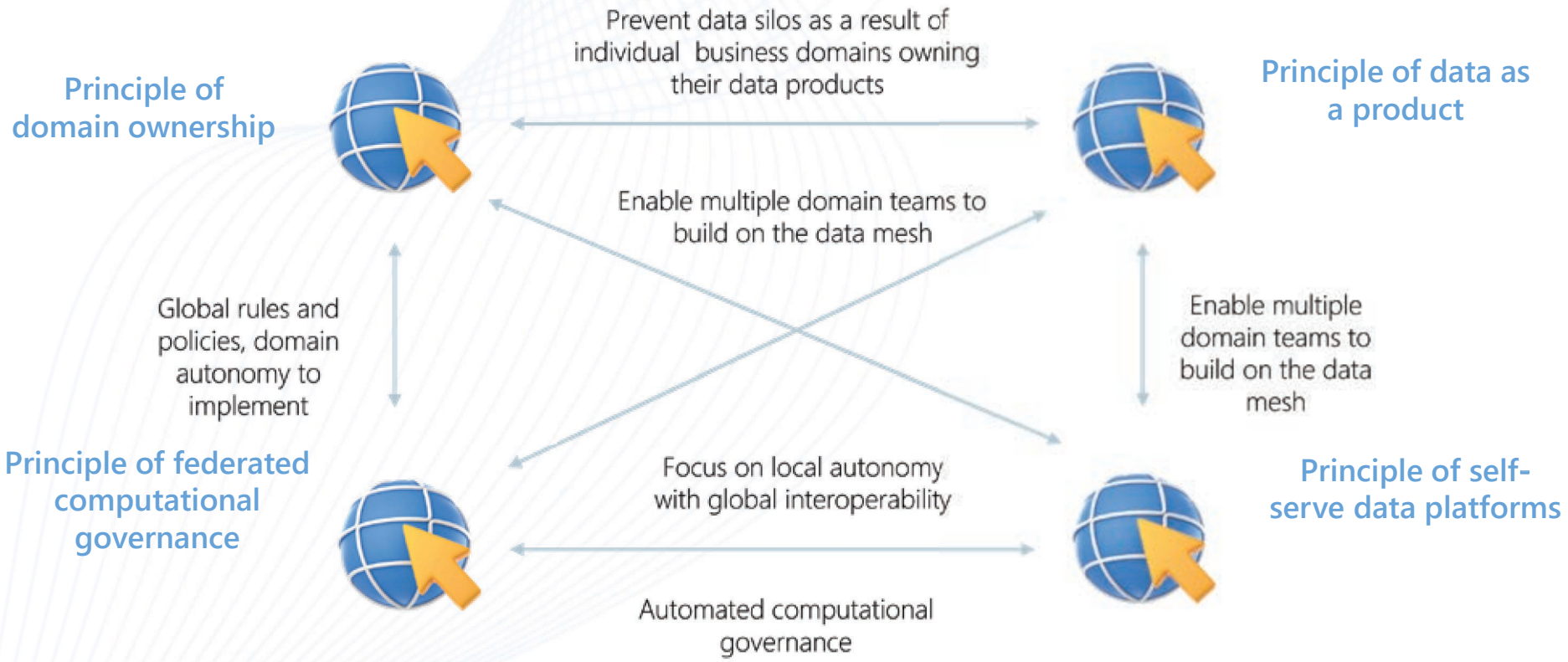
Data provided by domains is treated as a product with the aim of providing data which is valuable and usable for data consumers. This involves application of product thinking in the design of data products and defining baseline characteristics that apply to data products.

Principle of federated **computational governance**

A federated governance model creates an equilibrium between local and global optimization, giving autonomy to domains but maintaining global interoperability. Computational governance involves automation of instructions that assure data is secure, compliant and usable.

How the mesh principals **interoperate**

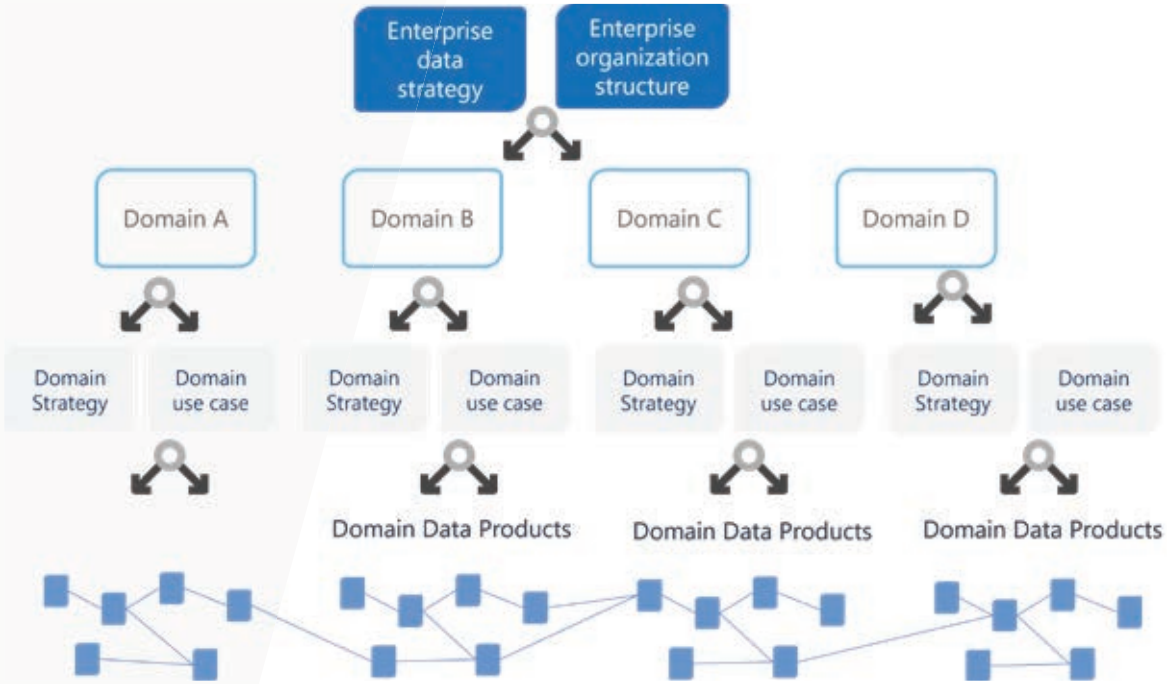
The mesh principals are defined to be collectively necessary for a strategic implementation



Defining the data **mesh strategy**

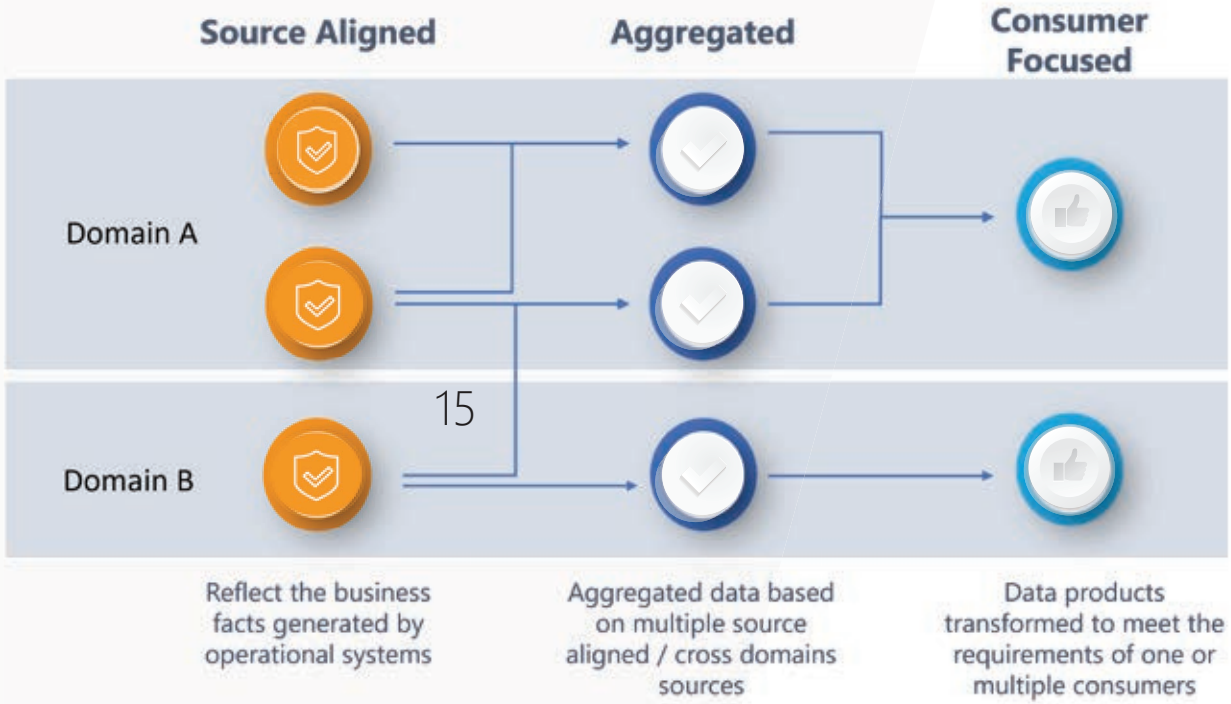
Domain definition and ownership

The data mesh is based on a top-down approach where the enterprise data strategy and the organisation structure drive the logical domain definition and the domain strategy and use cases define the data products a domain will build, manage and share.



Domain data products

Domain are responsible for publishing products which represent the truth of their business and also products which are valuable for end consumers. This requires creation of aggre-gated and consumer focused cross domain products



Data Product
Design Principals

1. Discoverable

Enable data consumers to explore, search and find required data sets and gain confidence in using them

2. Understandable

Enable data consumers to explore, search and find required data sets and gain confidence in using them

3. Trustworthy

Enable data consumers to explore, search and find required data sets and gain confidence in using them

4. Valuable

Enable data consumers to explore, search and find required data sets and gain confidence in using them

5. Interoperable

Enable data consumers to explore, search and find required data sets and gain confidence in using them

6. Accessible

Enable data consumers to explore, search and find required data sets and gain confidence in

7. Addressable

Enable data consumers to explore, search and find required data sets and gain confidence in using them

8. Secure

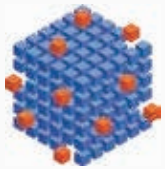
Enable data consumers to explore, search and find required data sets and gain confidence in using them

The Data Management function and Data Governance committee will ensure the data product desigin, build and operating model will align to the following data product design principals

Principle of self serve data platforms

Enable autonomous domain teams

The objective of the self serve plat form is to enable multiple autonomous domain teams to be able to build, manage and share their domain data product



Ability to build interoperable data products

The platform should be able to interoperate. This principle propagates a mutli-platform architecture which is able to interoperate via APIs

Low cognitive load on domain teams

The mesh should be designed for the generalist majority and not for a specialist minority. A high degree of specialization is an anti pattern and will be detrimental to the long term scale up of the mesh

Support end to end data management capability

The self serve data platform should support end to end data management capabili-ties. Enablinging domain teams to build, deploy, share and govern their data products

Ability to scale up to support multiple teams

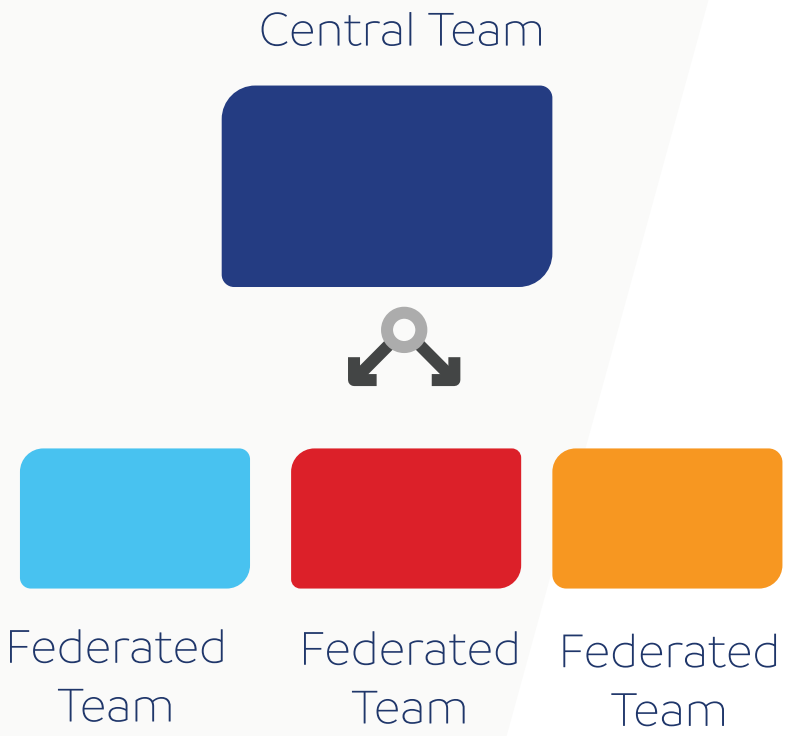
The platform must have the ability to scale as the number of domain teams building products on the mesh increases. This ability to scale should not impact platform capabilities and data products to interope-rate

Multiplane data platform architecture

We are addressing the data sharing and data governance requirements and the driving principle for our recommendation on implementation is the multiplane architecture concept which loosely classifies the platform services based on the scope of operation.

Federated Governance

The governance function is responsible to ensure that high quality, consistent data is provided to consumers in a timely manner, at the same time making sure the data is compliant to policies and secure



Define global policies rules and standards

- Define data goverance operating model and technology
- Define data privacy and security policies and requirements
- Define global compliance and regulatory policies
- Defne data quality standards, operating mdoel and technology

Local autonomy to implement

- Federated teams have the autonomy to implement governance practices defined by the central team
- They have the flexibility to implement any specific local requirements
- Federated teams will have their own data governance function which will include data owners, data stewards and data custodians

Computational Governance

Computational governance involves the automated computation of data which drives effec-tive governance of the data, this is done by creating a data governance sidecar in the data quantum defined by the data mesh.

Data product — Computational Governance

USAGE

Track usage statistics on who is consuming the data along with frequency and volume of consump-tion

OBSERVABILITY

Track usage statistics on who is consuming the data along with frequency and volume of consump-tion

SECURITY

Track usage statistics on who is consuming the data along with frequency and volume of consump-tion

COST OF STORAGE / COMPUTE

Automated and interac-tive dashboards to mea-sure and track the cost of storage and cost of compute for a data platform

DATA QUALITY

Integrate DQ into the data product build work-flow with the ability to identify, manage and clean bad data

COMPLIANCE AND POLICIES

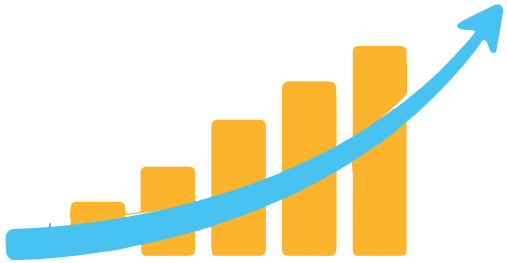
Automate application of compliance and regulato-ry policies

Defining the data mesh **execution plan**

The data mesh implementation journey

A data mesh implementation is a long-term organisational commitment.

The implementation journey follows a standard S curve.



Explore

Mature data teams kick off publishing data products on the mesh to satisfy the majority uses cases for data domains. This phase proves the value of the mesh and paves the path for expansion

Expand

The wider organisations kicks off publishing products on the enterprise data mesh. The self-serve platform is mature and supports rapid scaling and publishing of data products

Scale

A data mesh implementation is a long-term organisational commitment.

The implementation journey follows a standard S curve.

The Keyrus Team

Let's talk about how we can **make your data matter**

Adam Walker
General Manager

Scan here

Steven Hunt
Strategic Sales Manager

Scan here

Nkosinathi Xulu
Head of Data Engineering

Scan here

Craig Andrew
Head of Data Analytics

Scan here

The Keyrus Approach

We understand the importance of promoting data democratization, thus, we enable users at all levels of the organisation to access and leverage data independently. By doing so, we empower businesses with agility and autonomy, allowing different teams to take ownership of their data domains and drive decision-making with autonomy.

keyrus

make data matter

Thank you!

Visit website

We are here to solve your business challenges through modern data solutions.
Contact us at sales@keyrus.co.za or +27 87 350 8860