Sep 12, 2023

Integrating Informatica with Data Fabric and Data Mesh Architecture

- Anita Ayyagari, Senior Solutions Architect, CSA
- Kamal Abrol, Principal Solutions Architect, CSA
- Rashmi P, Solutions Architect, CSA

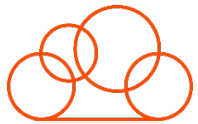
Informatica

Housekeeping Tips

- Today's Webinar is scheduled for 1 hour
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the Q&A option and we will respond at the end of the presentation
- The webinar is being recorded and will be available on our Success Portal - where you can also download the slide deck for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the post-webinar survey and provide your feedback and suggestions for upcoming topics.

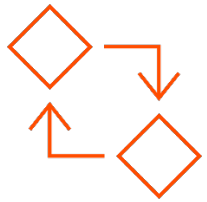
Informatica

Feature Rich Success Portal



Bootstrap trial and POC Customers

Enriched Customer Onboarding experience



Product Learning Paths and Weekly Expert Sessions

Informatica Concierge

Tailored training and content recommendations

Informatica

More Information

Success Portal

https://success.informatica.com

Communities & Support

https://network.informatica.com

Documentation

https://docs.informatica.com

University

https://www.informatica.com/in/services-and-training/informatica-university.html

Informatica

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Informatica

12/09/2023

Data Mesh Vs Data Fabric

Anita Ayyagari ,Kamal Abrol, Rashmi
Customer Success Architect Team

Informatica™

Agenda

- Overview of Data Mesh and Data Fabric
- Difference between Data Mesh and Data Fabric
- Data Mesh-Technology Needs
- IDMC platform -Integration with Data Mesh
- Data **Mesh** Federated Governance via IDMC
- Data Fabric-Technology Needs
- IDMC platform -Integration with Data Fabric
- Customer Reference use case & benefits

Overview

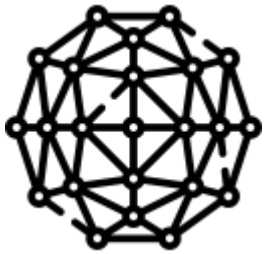
Data Mesh and Data Fabric

Data mesh and **Data fabric** architectures alike aim to abstract data management complexity. Both also deliver data with agility and scale.

Monolithic, legacy architecture and centralized data platforms thwart business agility and make it difficult to quickly adjust to the ever-changing data landscape.

Data as a service, a data management strategy aiming to leverage data as a business asset for greater business agility, has also been used. These logical architecture approaches aim to scale the delivery of data to satisfy diverse use cases.

What is Data Mesh ?



A data mesh is a decentralized data architecture where data is treated as a product and owned by teams that most intimately know and consume the data. It allows business domains—for example, marketing, sales, customer service, and more to easily access important data without transporting it to a data lake or data warehouse and without needing expert data teams to intervene.

Data Mesh is founded on four key Pillars

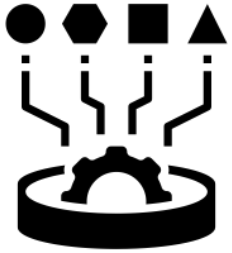
- Domain Oriented decentralized ownership
- Data as a product
- Self serve Data Infrastructure
- Federated Governance

Informatica

Why Businesses need Data Mesh ?

- **Improving data latency** – Centralized data/IT teams cannot handle all the questions of business and product owners quickly enough
- **Bridging the business and IT disconnect** - helps address the lack of domain expertise in centralized data teams, as they often don't fully understand the business context
- **Productivity issues** - Addresses lack of access to the right information at the right time
- **Data landscape complexity** – Provides a scalable data management approach that can support resilience and continuous change
- **Data Literacy and Internal culture** – Breaks down silos and eliminate the barriers to leveraging data democratically

What is Data Fabric ?



Data fabric delivers a unified, integrated, and intelligent end-to-end data platform to support new and emerging use cases. It automates all data management functions — including ingestion, transformation, orchestration, governance, security, preparation, quality and curation — enabling insights and analytics to accelerate use cases quickly

Data Fabric is founded on five key Pillars

- Augmented Data Catalog
- Knowledge graph enhanced with Semantics
- Metadata activation and recommendation engine
- Data preparation and data delivery
- Orchestration and DataOps

Informatica

Why Businesses need Data Fabric ?

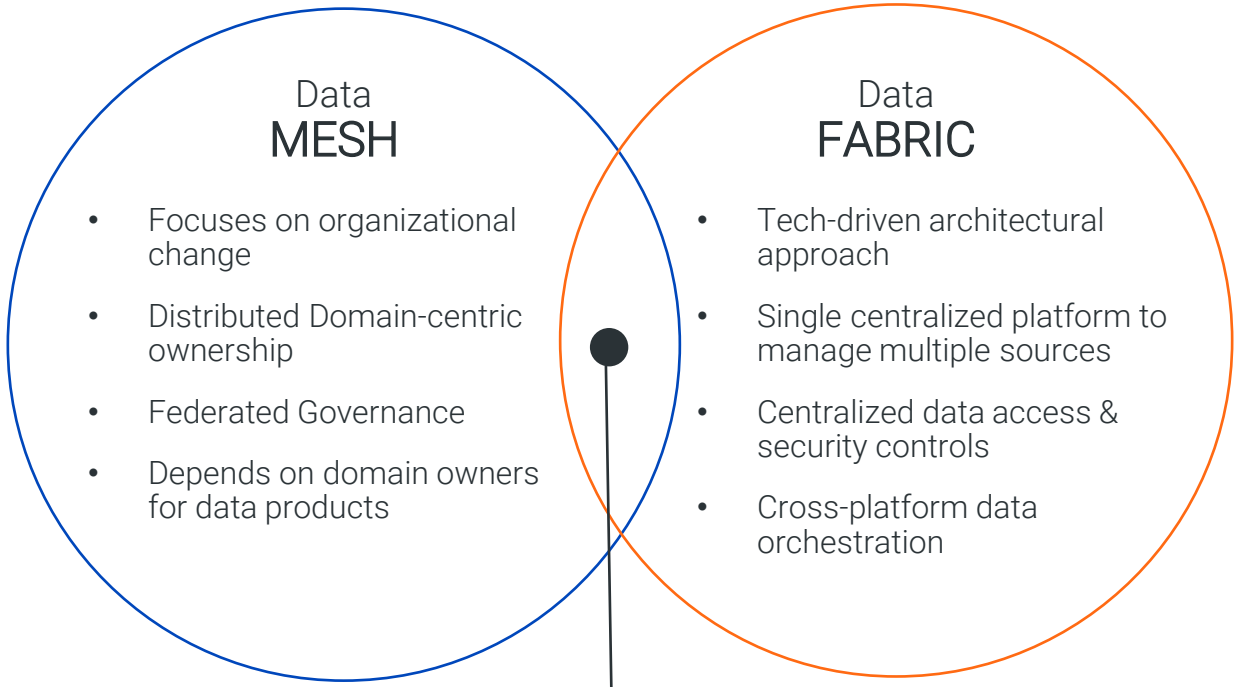
Companies are hindered in driving growth and competitive advantage because they are unable to

- **Scale to support an expanding number of business use cases in an agile manner**
- **Control spiraling costs and resource needs due to inefficiency, complexity, and lack of automation.**
- **Provide frictionless access to holistic, trusted, and timely insights**
- **Respond rapidly to volatile market conditions and competitive scenarios**

Data Mesh vs Data Fabric

What is Data Fabric? How is it different from data mesh?

A Data Fabric is an architecture concept **focused on automation** in discovering, connecting, recognizing, suggesting and delivering data assets to data consumers based on a rich enterprise metadata foundation. It **aims to connect and integrate** your organization to accelerate self-service data discovery and analytics via intelligent automation of data management tasks leveraging metadata.



Data Fabric and Data Mesh can co-exist and be simultaneously be used to expedite data use at scale

IDMC alignment with Data Mesh?

Informatica®

Agenda

- Overview of Data Mesh and Data Fabric ✔
- Difference between Data Mesh and Data Fabric ✔
- Data Mesh-**Technology Needs**
- IDMC platform -Integration with Data Mesh
- Data Mesh Federated Governance via IDMC
- Data Fabric-**Technology Needs**
- IDMC platform -Integration with Data Fabric
- Customer Reference use case & benefits

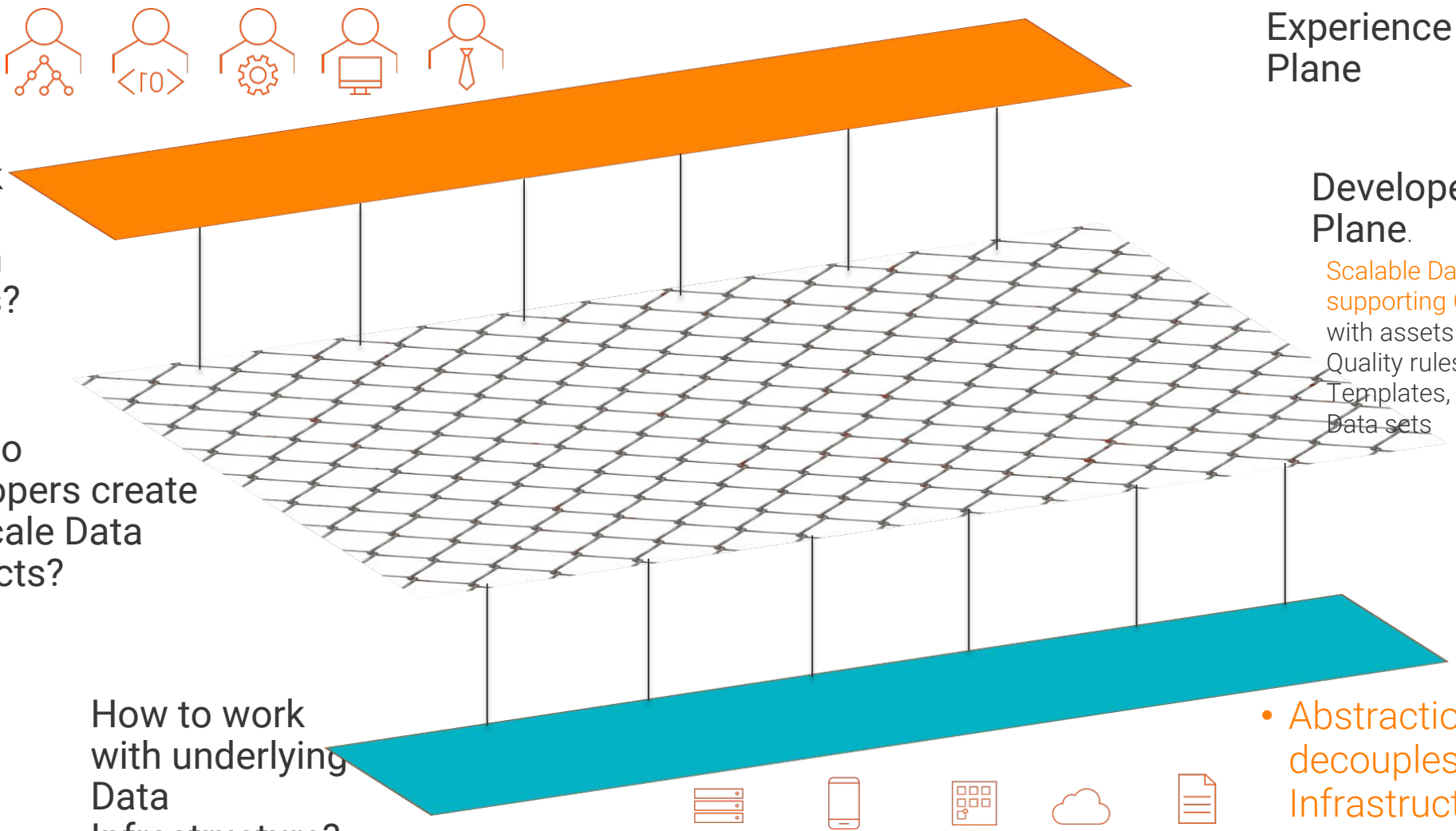
Informatica

Key Underlying Technology Needs from Data Mesh

- Trusted Data Marketplace as a StoreFront for Data Products

Mesh Experience Plane

How do Domains work together and collaborate on Data Products?



Developer Experience Plane.

Scalable Data Management supporting Cross-Domain reuse with assets such as Common Data Quality rules, Data Integration Templates, Reference & Master Data sets

How do Developers create and scale Data Products?

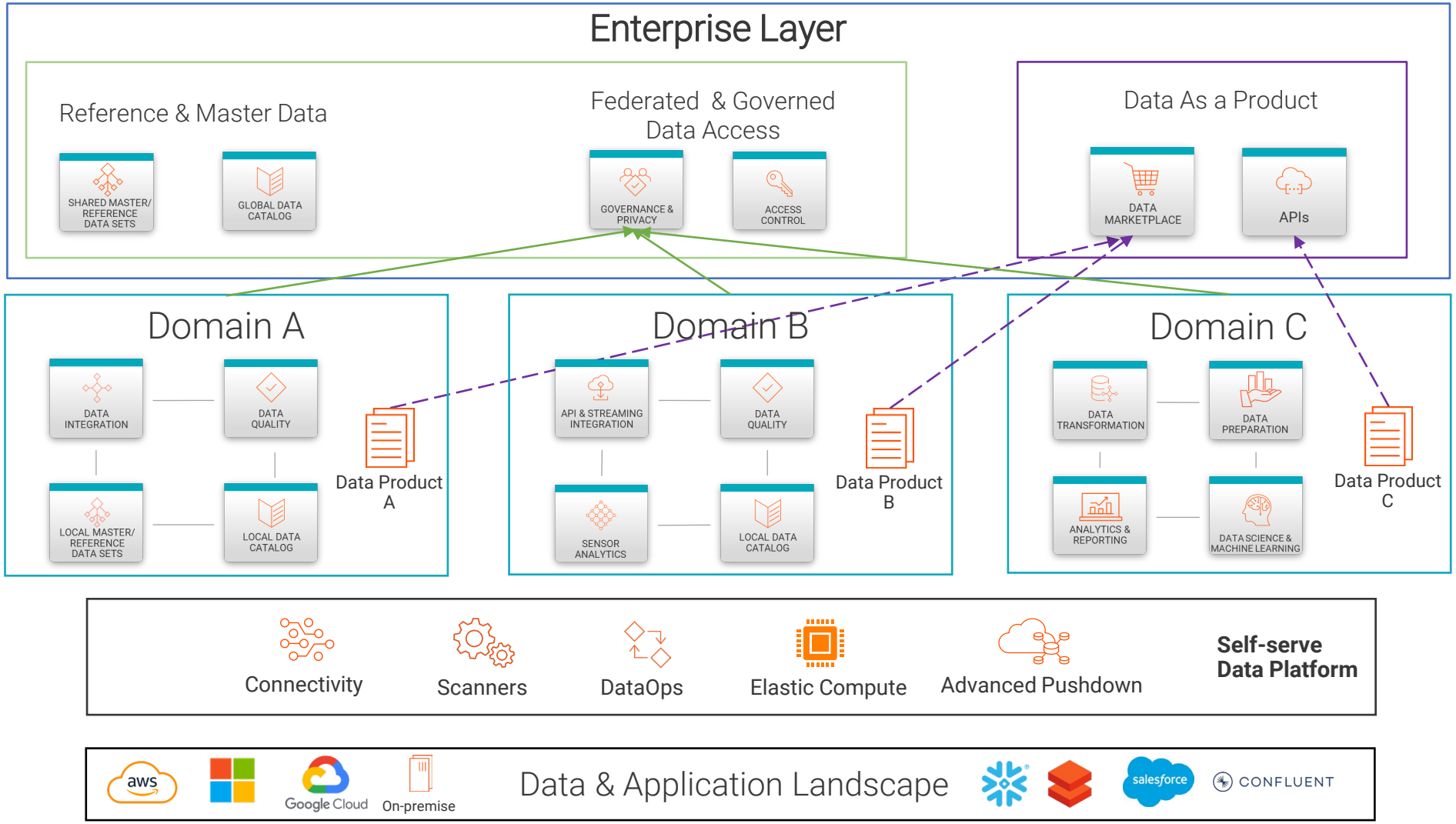
Data Infrastructure Plane

How to work with underlying Data Infrastructure?

- Abstraction Layer which decouples underlying Data Infrastructure powered with DataOps enabling Agility

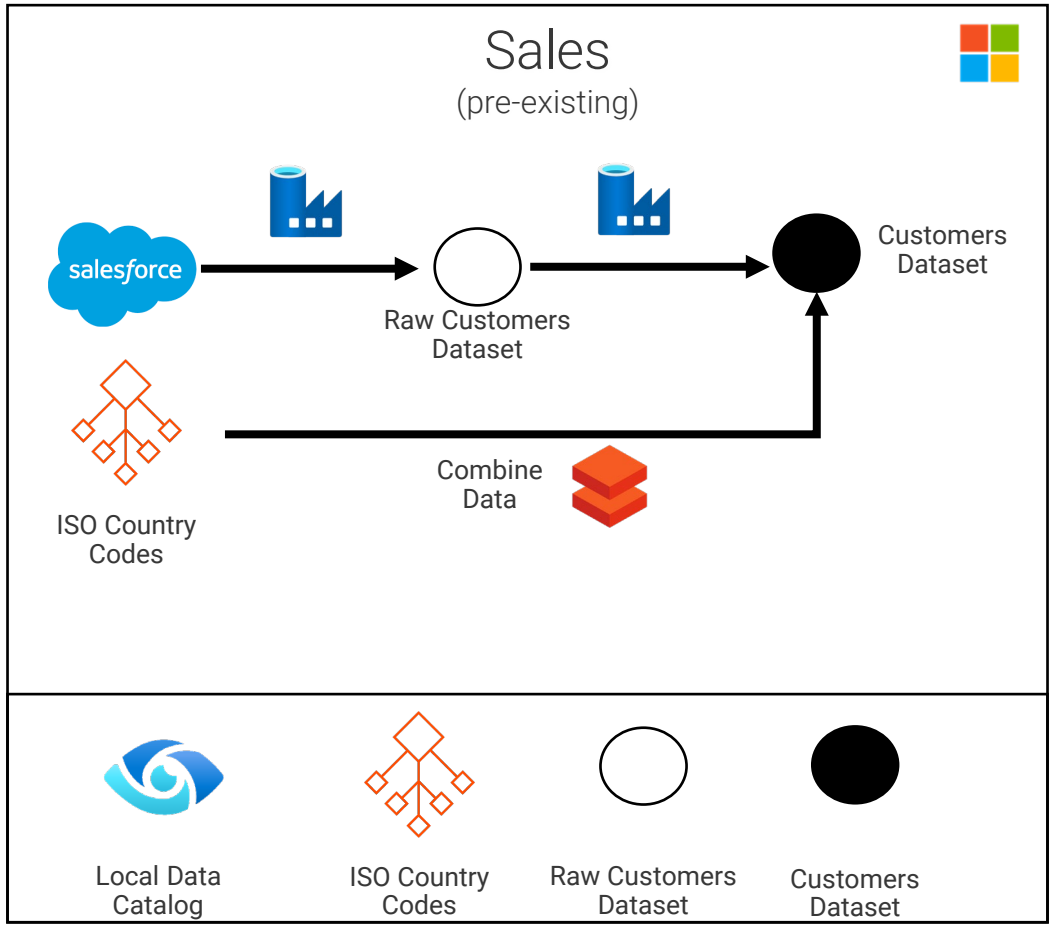
Informatica

Data Mesh Architecture with Informatica



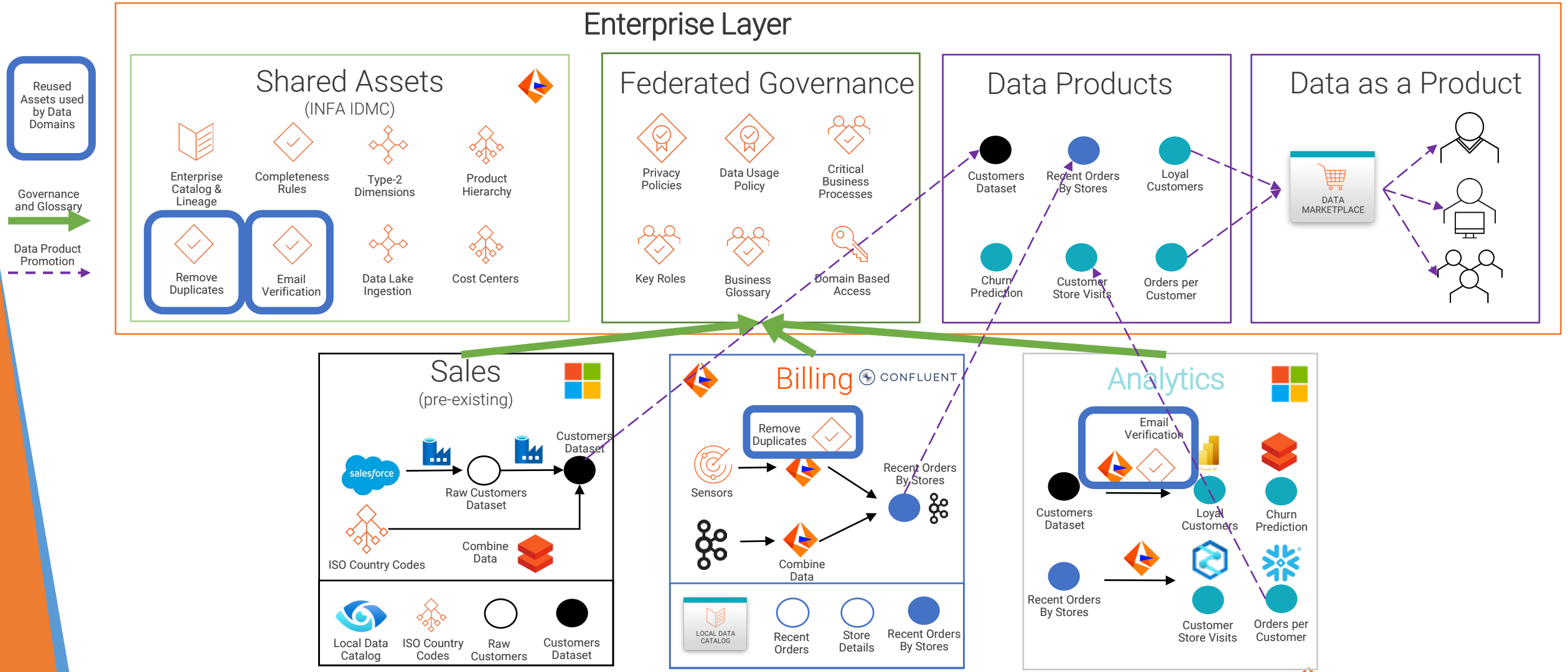
Informatica

Data Mesh Architecture with Informatica – Example



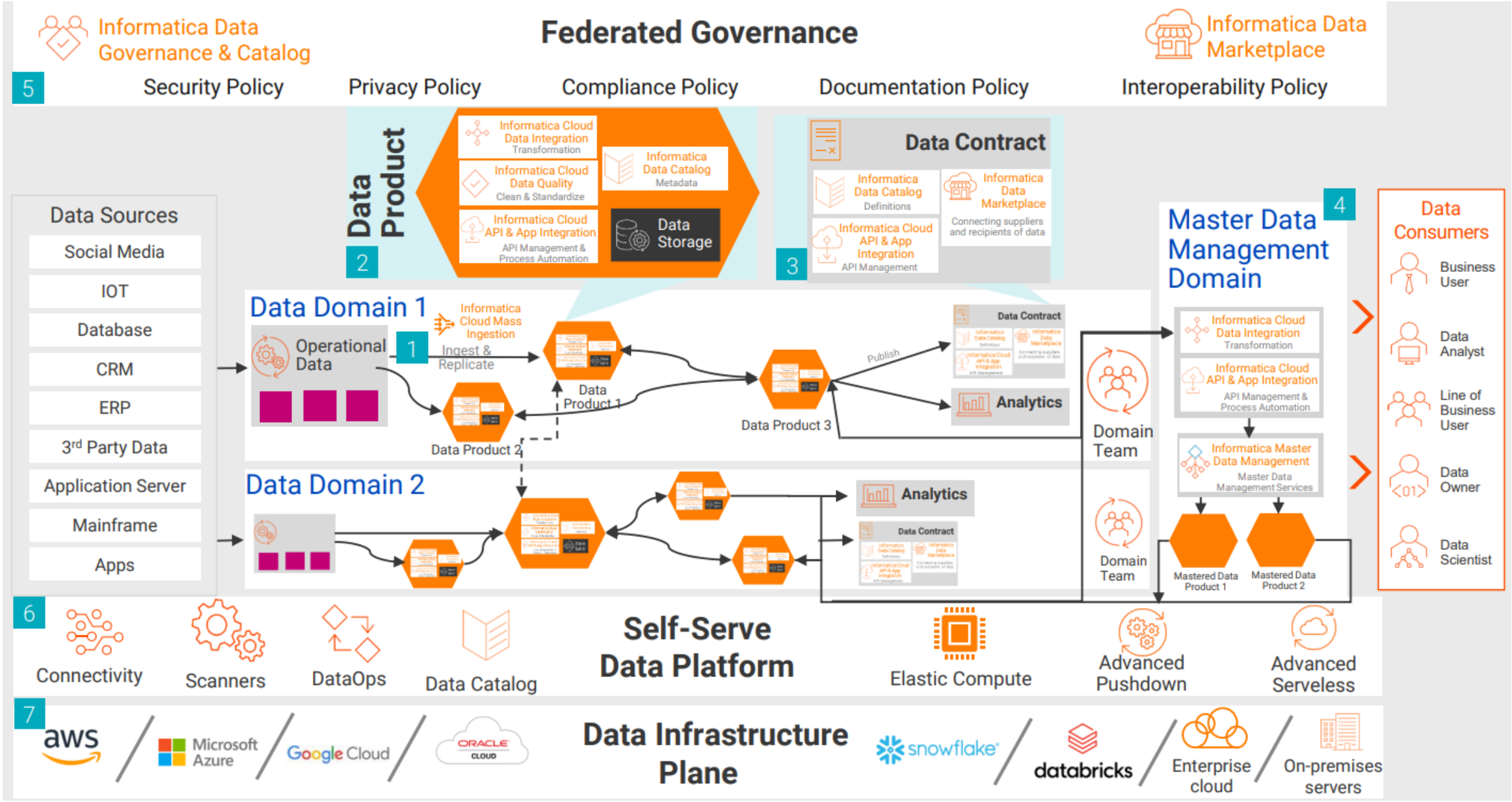
Informatica

Data Mesh Architecture with Informatica – Example



Informatica

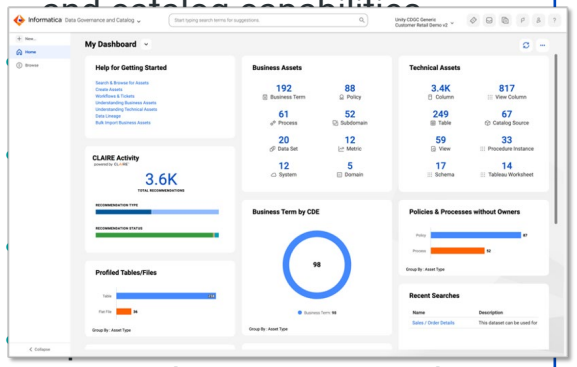
Data Mesh **Federated Governance** -Reference Architecture



Key IDMC Services Empowering Data Mesh

Data Governance and Catalog

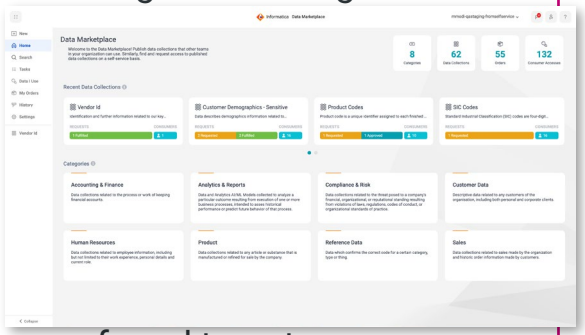
- Experience unified governance and catalog capabilities



powered automation and intelligence

Data Marketplace

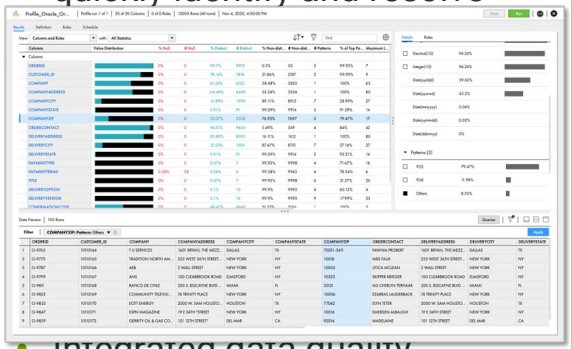
- Launch a storefront with order management and governed



preferred targets
- Track fulfillment with full visibility

Data Quality and Observability

- Intelligent Observability to quickly identify and resolve



- Integrated data quality scorecards for trust-based governance

Master Data Management

- Connect data across the enterprise for a 360-degree

- AI-powered matching algorithms to manage relationships

IDMC platform alignment with Data Fabric

Informatica

Agenda

- Overview of Data Mesh and Data Fabric ✔
- Difference between Data Mesh and Data Fabric ✔
- Data Mesh-Technology Needs ✔
- IDMC platform -Integration with Data Mesh ✔
- Data Mesh Federated Governance via IDMC ✔
- Data Fabric- Needs
- IDMC platform -Integration with Data Fabric
- Customer Reference use case & benefits

Informatica

Typical Data Challenges

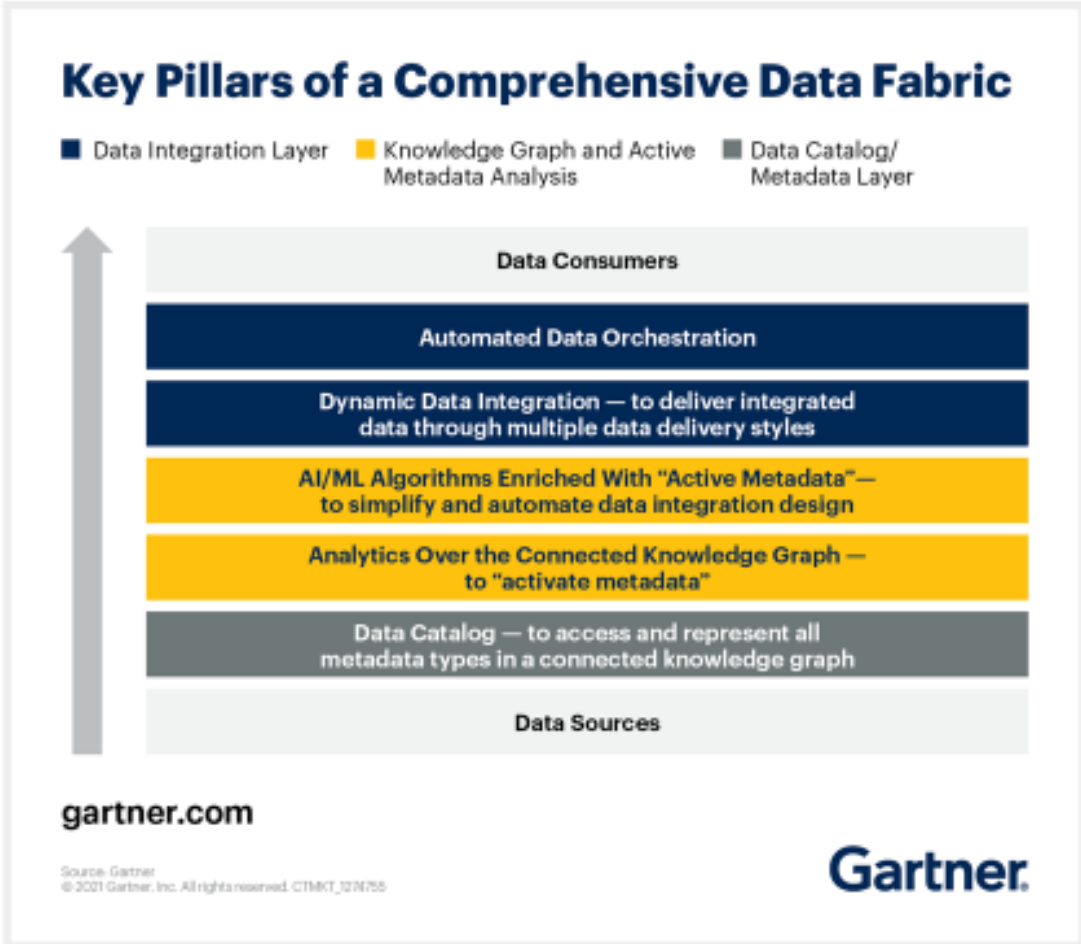
- I need data but I can't find it
  - No search capabilities; who do I go to in order to find what I'm looking for?
- I found data but how do I know I can trust is?
  - Where did the data come from?
  - How can I trust the quality of the data?
- I'm part of I.T. and I got a request to provision data, I need to create a data set quickly and efficiently
  - Hand coding takes too much time
- It's very time consuming and cumbersome to try to manually discover issues with the data and see if there is any sensitive data that needs to be protected/curated.
- Right access to the right people

Data Fabric Solution Benefits

- Easy to use self-service Cloud Data Marketplace with Amazon like shopping experience
  - Embedded policies for terms of use
  - Data quality metrics for assets
  - U.I. Driven interface for provisioning data (request and fulfillment)
- Data Governance & Catalog with Lineage
  - Knowledge Graph views showing data come from
  - Integrated AI/ML for automating classifications of assets in the Catalog
  - Integrated AI/ML for automating business glossary terms
- Robust, Scalable Cloud Integration Platform
  - Integrated Data Profiling to quickly identify any anomalies with the data leveraging AI/ML recommendations
  - Integrated Data Quality to resolve and issues with pre-built reusable, DQ Bundles to jump start projects
  - Integrated Data Masking to secure any sensitive data
  - Flexible architecture including serverless, elastic and advanced push down to Cloud Data Warehouse

Informatica

Data Fabric Key Pillars



Informatica

DATA CONSUMERS

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

Intelligent Data Management Cloud

DISCOVER & UNDERSTAND	ACCCESS & INTEGRATE	CONNECT & AUTOMATE	CLEANSE & TRUST	MASTER & RELATE	GOVERN & PROTECT	SHARE & DEMOCRATIZE
DATA CATALOG	DATA INTEGRATION	API & APP INTEGRATION	DATA QUALITY	MDM & 360 APPLICATIONS	GOVERNANCE & PRIVACY	DATA MARKETPLACE

CLAIRE™
AI-Powered Metadata Intelligence & Automation

Connectivity
Metadata System of Record

DATA SOURCES

workday. | servicenow | salesforce | ORACLE CLOUD | SAP

SaaS Apps Sources + Mainframe | Applications | Databases

On-premises Sources + IoT | Machine Data | Logs

Real-time / Streaming Sources

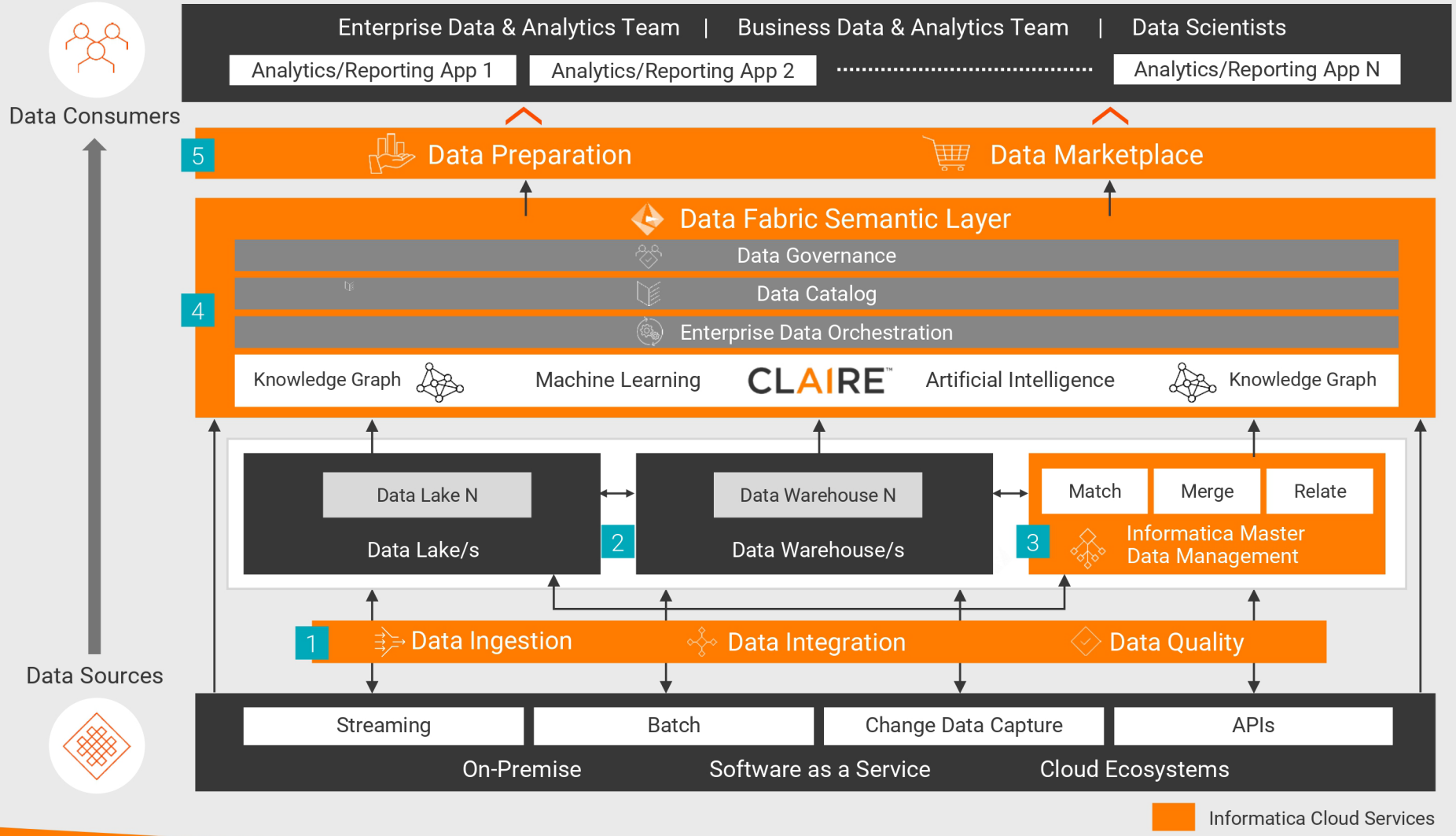
Informatica

Informatica

Reference Architecture

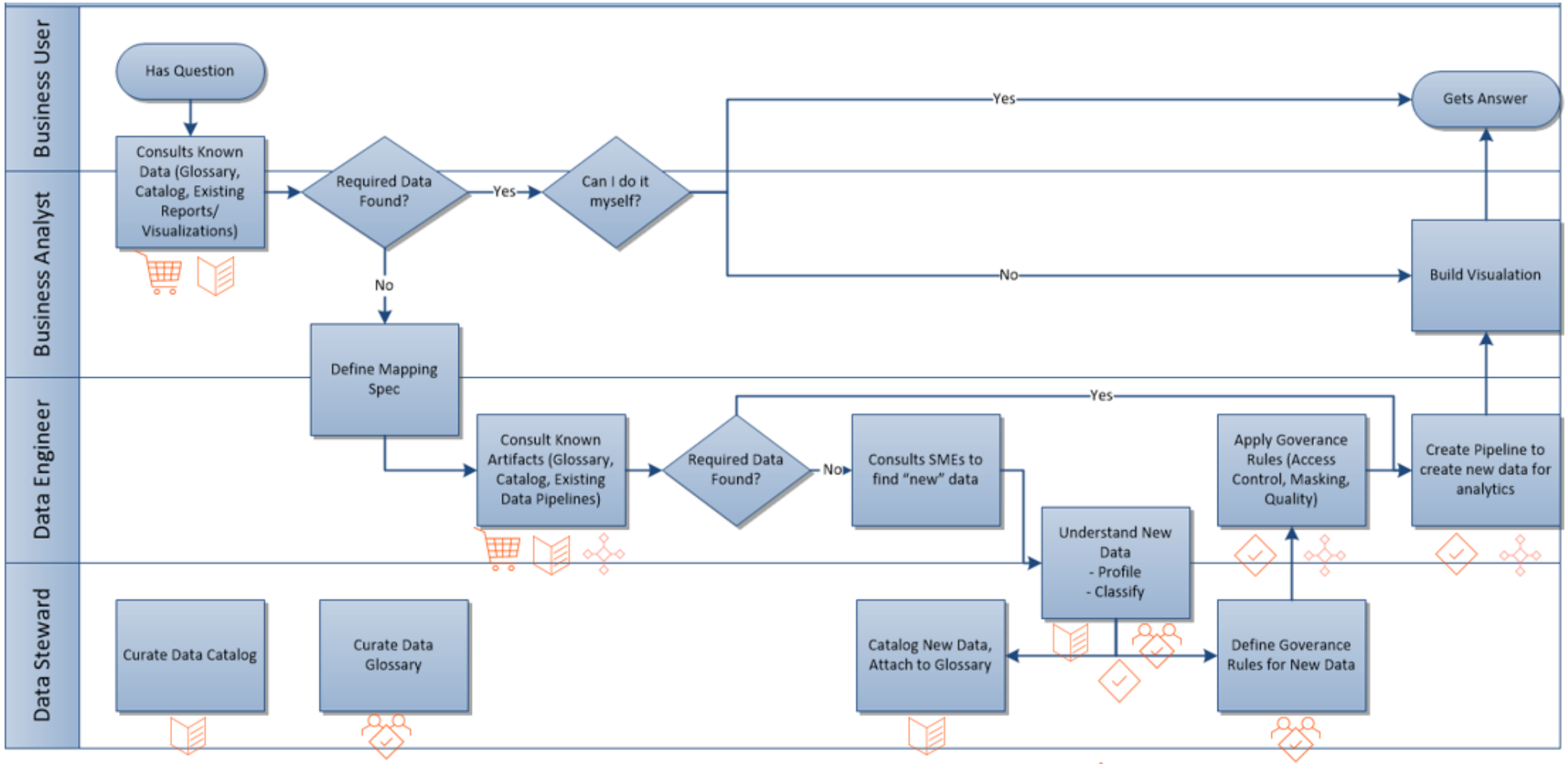
Intelligent Data Fabric

Data fabric integrates and connects entire organization's data intelligently and efficiently by abstracting underlying complexity. It minimizes disruption by enabling a highly adaptable data management strategy with augmented data integration and management.



1. **Data Ingestion, Data Integration and Data Quality**—Ingest, standardize and cleanse any data, at any speed, from on-premises, software as a service (SaaS) or cloud ecosystems using scalable streaming, batch, change data capture and APIs with comprehensive and high-performance connectivity.
2. **Data Lake/s, Data Warehouse/s and other analytic data stores**—Store structured and unstructured data using fit-for-purpose data management technologies (RDBMS, Spark, Cloud Storage Objects, NoSQL DBs).
3. **Master Data Management**—Harmonize common enterprise data that is fragmented across domains. Match duplicate data, merge common data into a "golden master" and relate this shared, high-value data to other relevant data. Make this "golden data" accessible across the fabric.
4. **Semantic Layer**—Support the "metadata system of record" across all data stores. ML/AI automates the capture and augmentation of metadata from disparate data sources and populates a knowledge graph to document linkages between data and business. The data catalog provides a semantically searchable store of this metadata, including data lineage, data profiling results and tribal knowledge, to facilitate data discovery and understanding. The data governance provides context for the technical understanding, business relevance and usage and access of the data. Enterprise data orchestration coordinates these semantic support processes and data delivery.
5. **Data Preparation and Data Marketplace**—Governed, self-service provisioning of analytic data for data consumers. Data preparation provides a user-friendly interface to gather, combine, structure and organize this data, while data marketplace creates a "shopping for data" experience to find and deliver data.

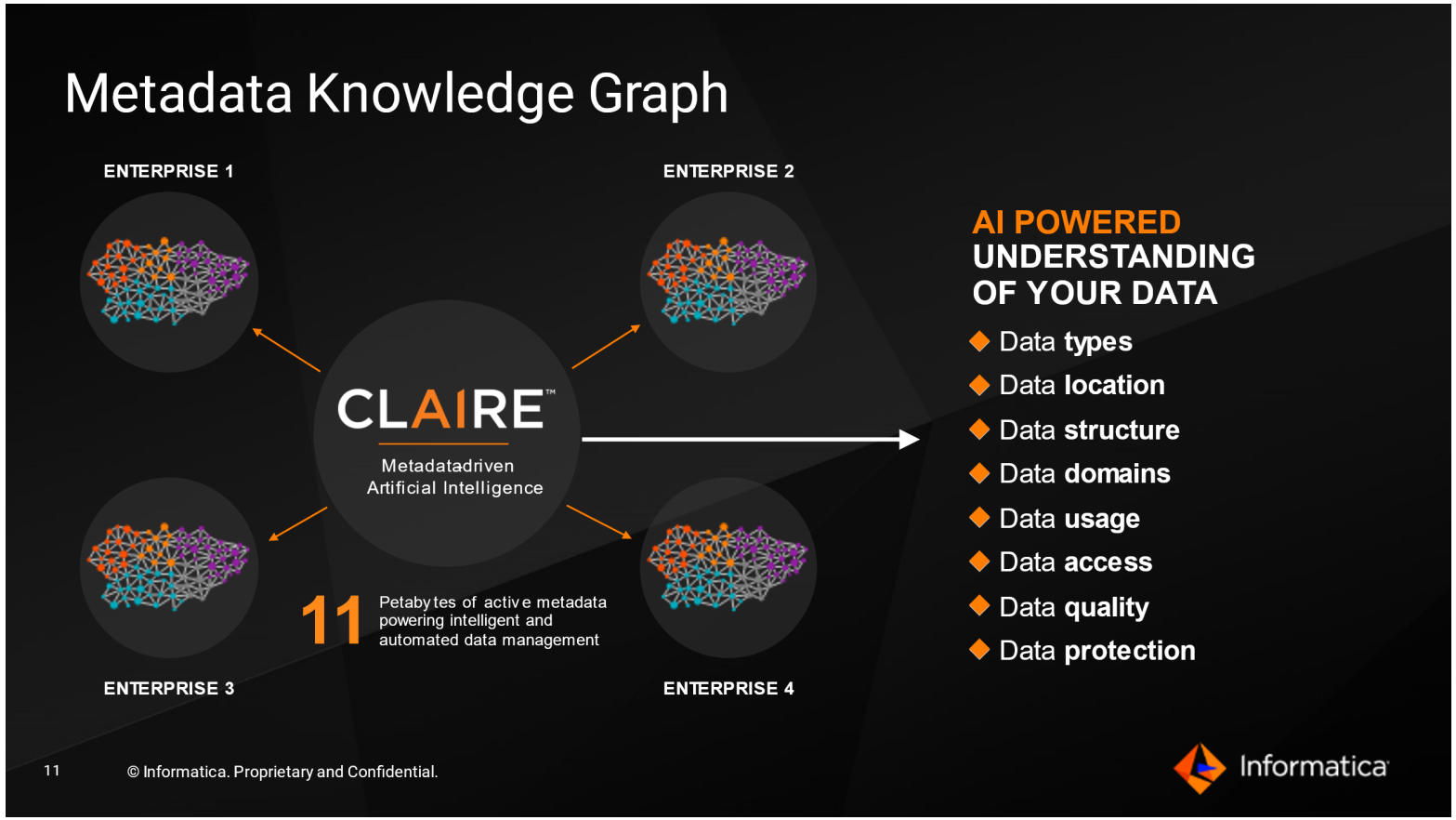
IDMC Enables Data Fabric Workflow



DATA MARKETPLACE | DATA CATALOG | GOVERNANCE & PRIVACY | DATA INTEGRATION | DATA QUALITY

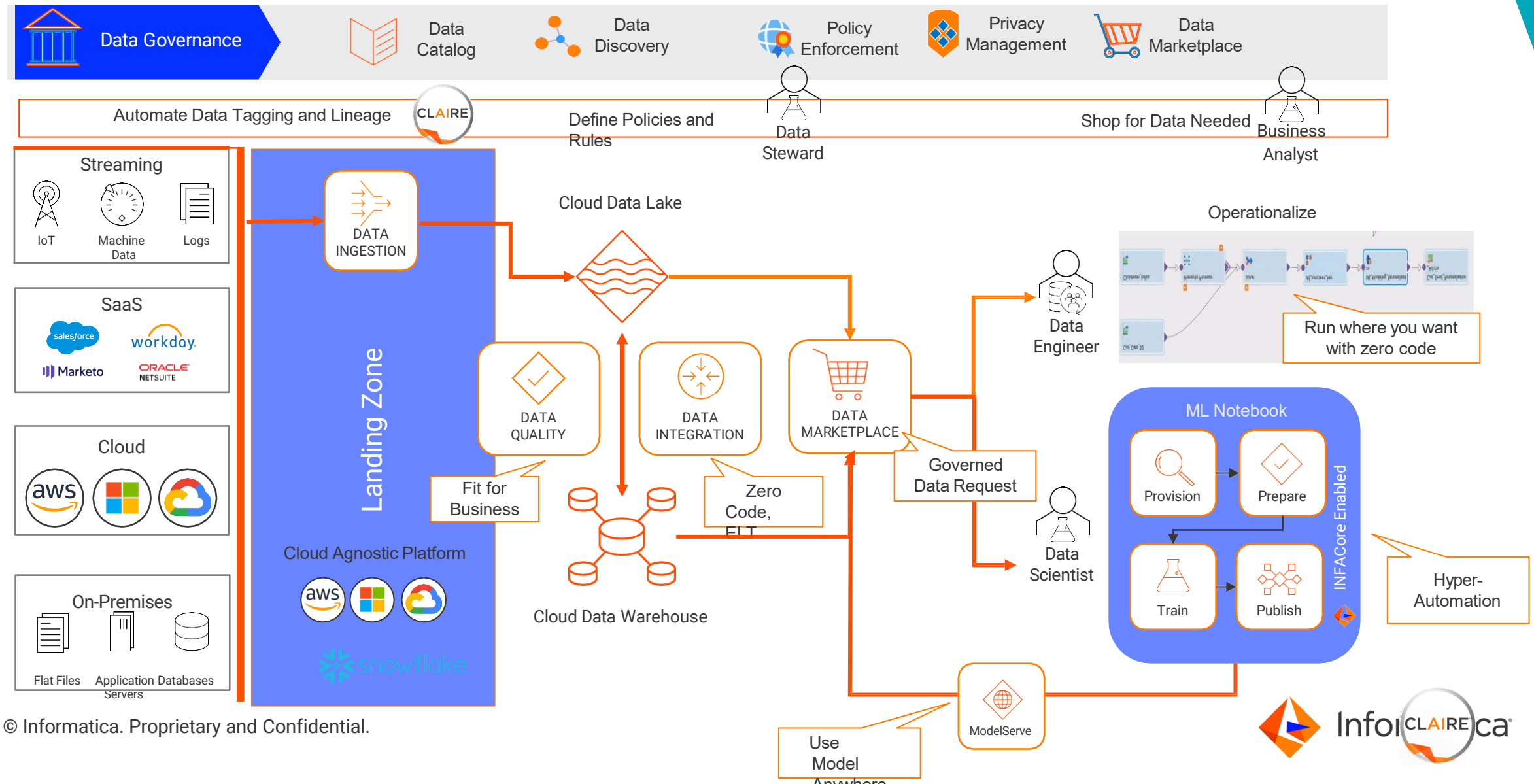
Informatica

Knowledge Graph Sample



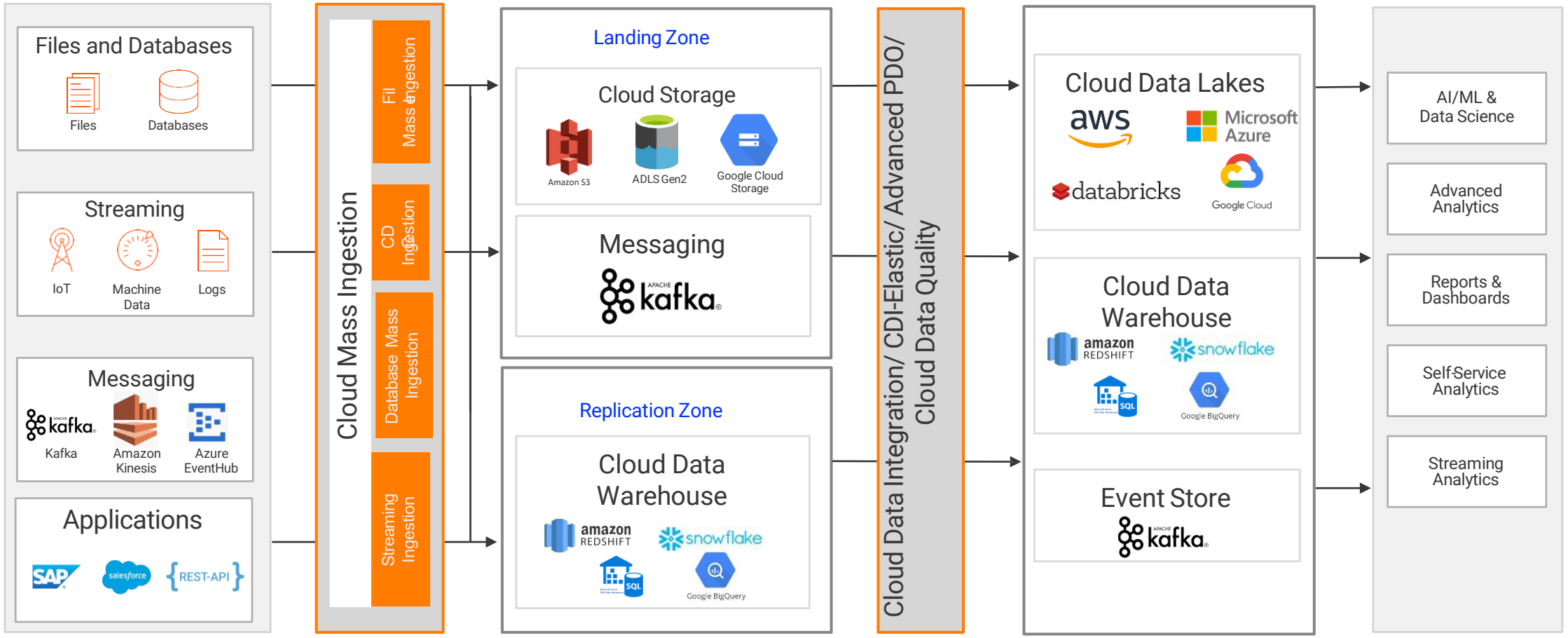
Informatica

Informatica Data Fabric Modern Reference Architecture



Mass Ingestion

Easy 4-step wizard for data engineers to mass ingest files, databases, CDC and streaming
Automatic change data capture (CDC) to ingest data into cloud data warehouses for quick decision making
Ingest tens of thousands of databases along with CDC, millions of files & streaming events



Informatica

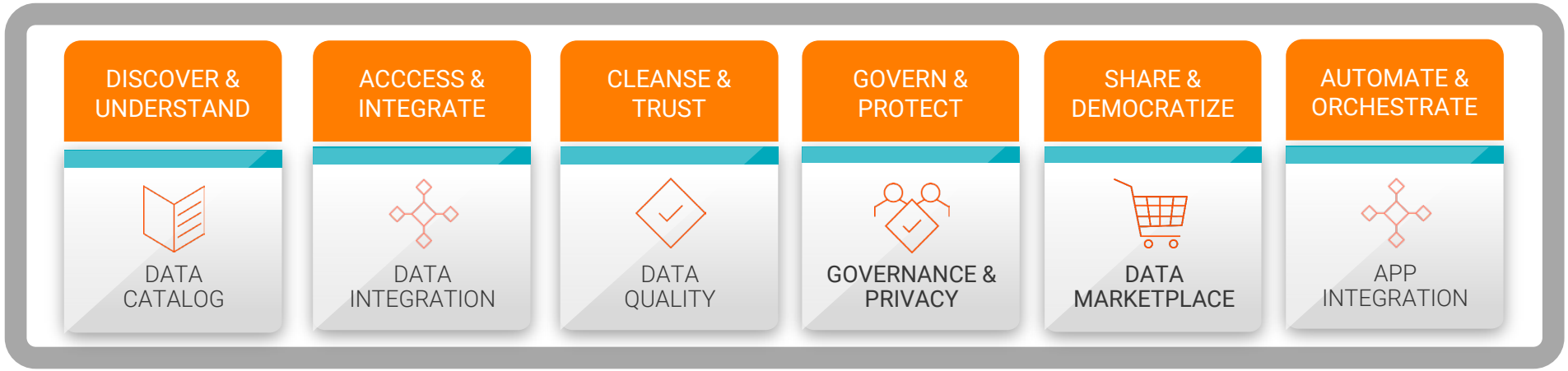
What are the business benefits of Informatica's solution?

With an Intelligent Data Fabric, you can

- Integrate and connect all your organization's data to enable frictionless data sharing for improved business outcomes.
- Accelerate self-service data discovery and analytics by making trusted data accessible faster to all applications and data consumers.
- Reduce data management costs and efforts via intelligent automation, optimization, and augmentation of data integration and management tasks.
- Respond to emerging business needs rapidly with a unified view of business data via continuous integration and analysis of diverse, siloed data assets and their business-relevant relationships. Accelerate data governance and protection with AI-powered automation of data

Choose Your Path

Go Big or Go Small: Choose the services that best suits your DF Use Case



- Choose the services that best fit your DF Use Cases
- Fit for purpose implementation plans are then aligned with the chosen services providing the following key best practice deliverables
  - Summary and Detailed Project plan
  - Non-Prod and Prod Optimize Reference Architecture and Specifications
  - End2End Process Flows
  - Optimized scanning configurations, approaches, and methodologies
  - Data Capture Templates
  - No Code/Low Code Integration/Automation/Orchestration

Informatica

Data Fabric and Data Mesh Co-exist?

Three ways a data fabric enables the implementation of a data mesh:

1.Provides data owners data products creation capabilities like cataloging data assets, transforming assets into products and following federated governance policies

2.Enable data owners and data consumers to use the data products in various ways such as publishing data products to the catalog, searching and find data products, and querying or visualizing data products leveraging data virtualization or using APIs.

3.Use insights from data fabric metadata to automate tasks by learning from patterns as part of the data product creation process or as part of the process of monitoring data product

Use Case- Hybrid Architecture

Data Product Design: Hybrid Data Fabric & Data Mesh Framework

Types of Data Assets to be built on Cloud

- **Operational Data** (near real-time)
- **Analytical Data** (historical)

Centralized store

Technology Support

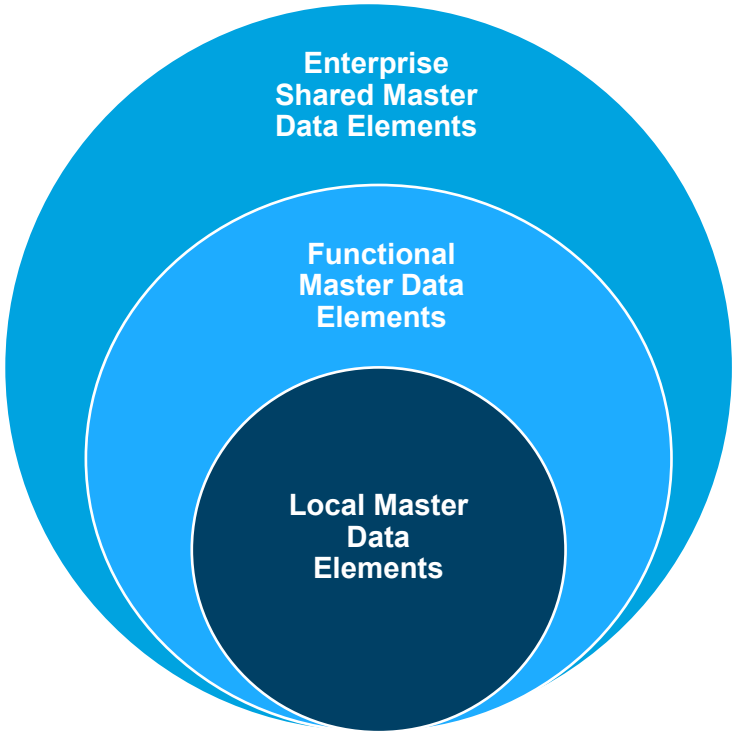
- Data Warehouse
- Data Marts
- Master Data Management
- Reference Data Management

Principles

- Domain-oriented decentralized data ownership
- Speed, ease of use, and collaboration across teams
- Discoverability, Security, Explorability, Understandability, Trustworthiness
- Self-serve infrastructure
- Federated computational governance

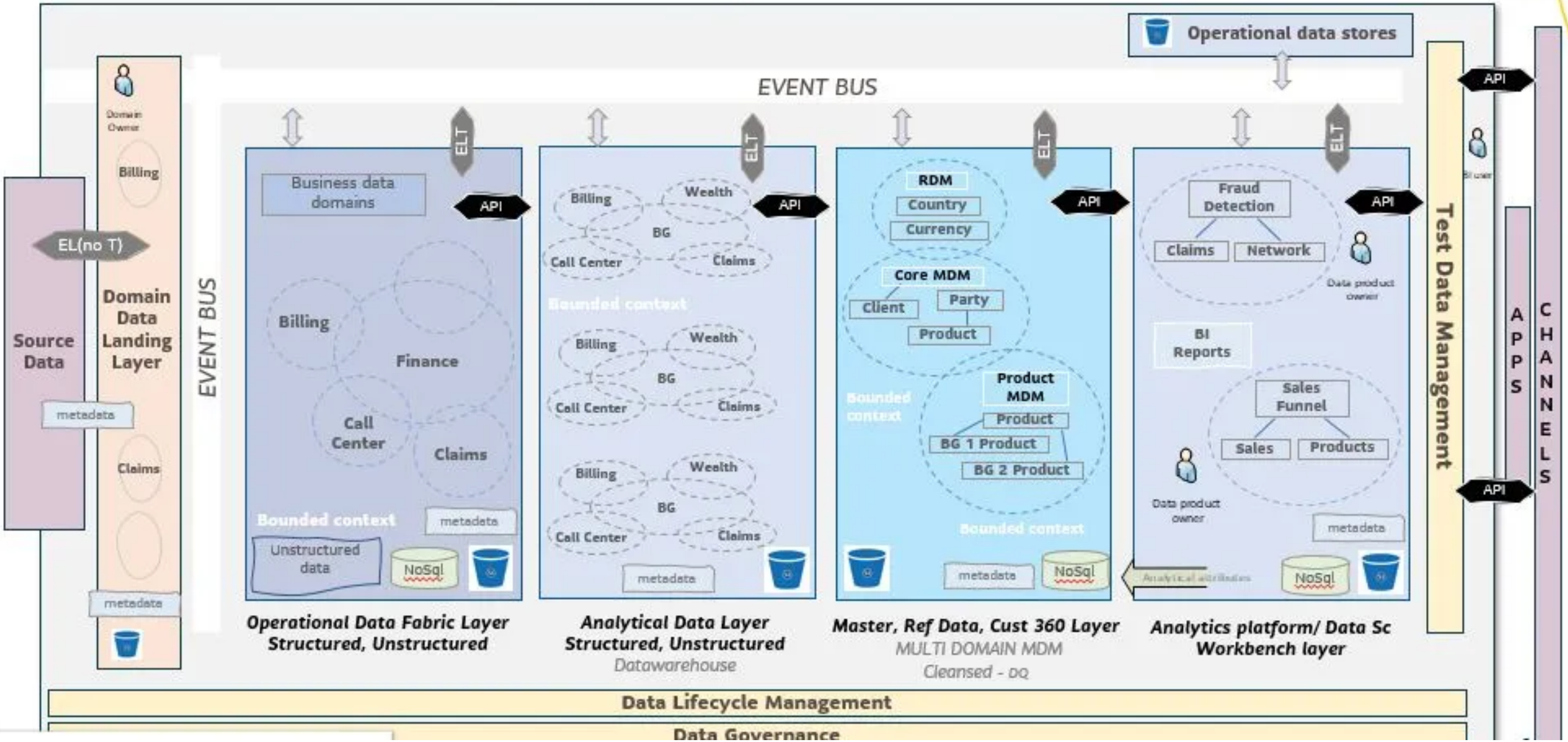
DATA PRODUCTS =  **DATA FABRIC TOOLS/TECHNIQUES** +  **DATA MESH PRINCIPLES**

Three tier Data Governance Strategy- developed to define appropriate ownership and stewardship of data in Hybrid Architecture



Categorization Type		Definition
Enterprise Shared Master Data Elements		• Data element(s) that originate in a specific application, consumed by multiple business functions • Data elements that are highly shared and need to be governed, mastered and managed by **enterprise business data domain owner** in Enterprise MDM
Data Elements to Support Functional Business Processes	Functional "Shared" Data Elements	• Governed, managed and updated by the functional MDM hubs • Shared within the same function
	Functionally "Localized" Data Elements	• Master data elements usage is restricted to a specific affiliate, country or manufacturing site to meet local requirements

Data Platform Design with Hybrid Data Fabric and Data Mesh



CSA Team

References

- Part 2 —Build a data product using Data Fabric and Data Mesh approach to learn the difference | by Dr. Shweta Shah | Medium

References

Additional Resources:

PPT Slides: Data Mesh- Reference Architecture & Customer Examples

Links listed in Additional Reading PPT Slide in last section:

- Data Mesh Principles and Logical Architecture
- How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh
- Introduction to Data Mesh

Informatica Specific:

- Informatica EMEA Architect Workshop
- Informatica EMEA Data Mesh Workshop with Snowflake & Wipro
- Mythbusters Session: Data Mesh
- Associating Data Governance with Data Mesh and Data Fabric
- Learnify :https://informatica.lessonly.com/path/265653-data-mesh/lesson/1177081-data-mesh-key-principles/complete

Informatica