



**Practical application of knowledge  
graphs using EDM Council's CDMC  
information model**

**Thursday, July 13  
11 AM EST**

**Knowledge Graph  
Special Interest Group  
Update**

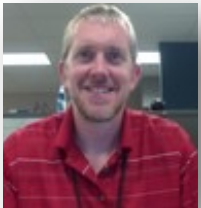
# EDM Webinar

## Knowledge Graph Special Interest Group Update 13 July 2023

*A Graph industry update with:*



**Shaun Rolls**  
Senior Advisor  
Head of Knowledge Graph SIG  
EDM Council



**Terry Hedin**  
Director, Data Architect,  
Product Manager  
LSEG



**Jim Halcomb**  
Head of Product Management  
EDM Council



**Ben Clinch**  
Head of Information Architecture  
BT Group



# Today's panel

Moderator



**Jim Halcomb**

Head of Product  
Management  
EDM Council



**Shaun Rolls**

Senior Advisor  
Head of Knowledge Graph SIG  
EDM Council



**Terry Hedin**

Director, Data Architect,  
Product Manager  
LSEG



**Ben Clinch**

Head of Information  
Architecture  
BT Group





# Poll #1



## What is your current role ?

- a. Analytics
- b. Architecture
- c. Technology
- d. Ontology/Modeling
- e. Data Governance



**Data is Everywhere**



# KNOWLEDGE GRAPHS

- WHAT ARE THEY
- HOW ARE THEY USED

# WHAT IS A KNOWLEDGE GRAPH



## Knowledge Graph

The “graph” in Knowledge Graph refers to a way of organizing and continually evolving your data that highlights relationships between data points.

**Graph** data looks like a network of interconnected points. This is in contrast to databases like Oracle or MySQL — **relational** systems — where data is stored in tables (rows, columns, cells).

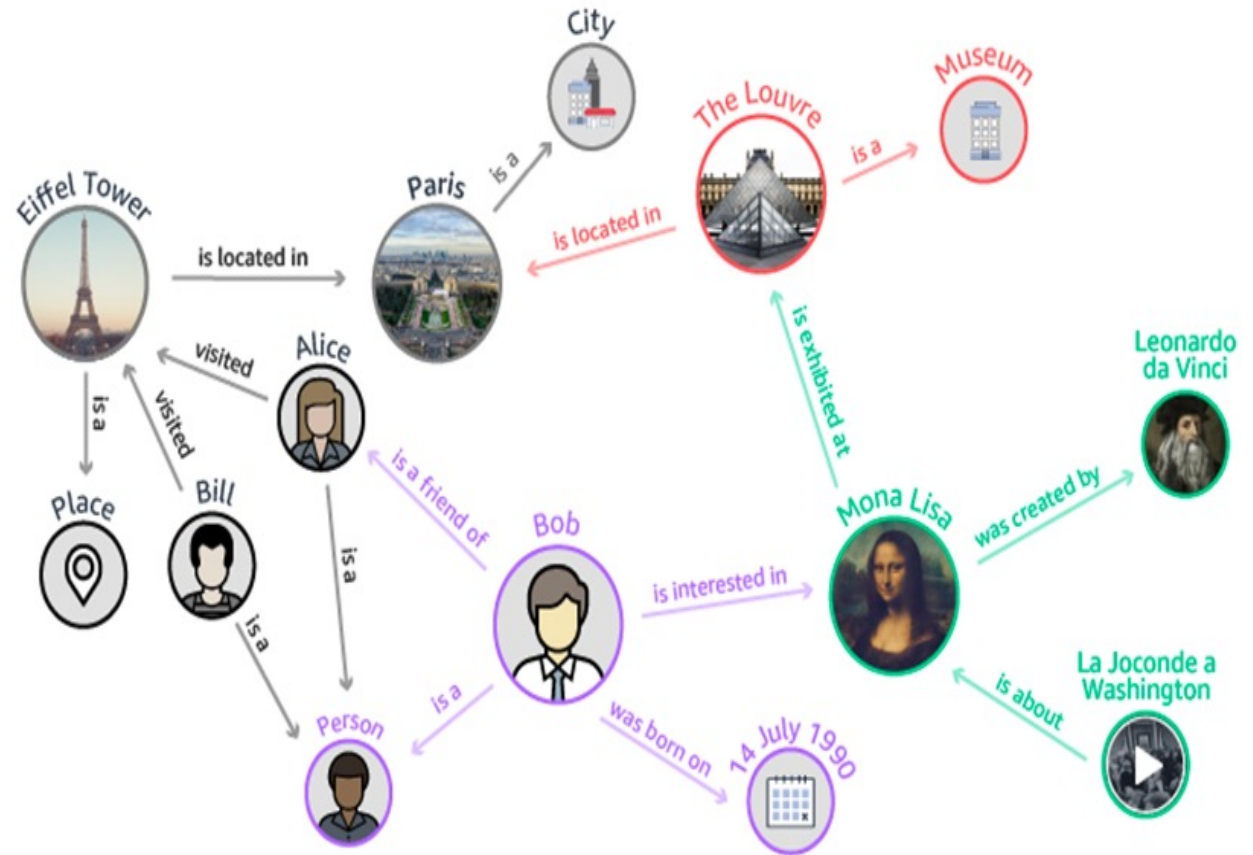
While knowledge graphs are dynamic via evolving data nodes that are continually connected/expressed.

Distributed data sets that have shared references for interoperability.



# KNOWLEDGE GRAPHS AND SEMANTIC DATA

- Knowledge graphs are structured via semantic data that focuses on the **nature of meaning**.
- The **connection** of words, concepts, symbols and language that are related to provide meaning.



# WHO USES KNOWLEDGE GRAPHS TODAY ?



Medicine , Pharma & Health



Technology



Financial Services

## Let's Define Data Fabric

It's an **emerging data management design**

for attaining ...  
that utilizes ...  
in support of ...  
regardless of ...

It is not one single tool or technology.

Flexible, reusable and augmented data integration pipelines

Knowledge graphs, semantics and ML/AI on active metadata

Faster and, in some cases, automated data access and sharing

Deployment options, use cases (operational or analytical) and/or architectural approaches

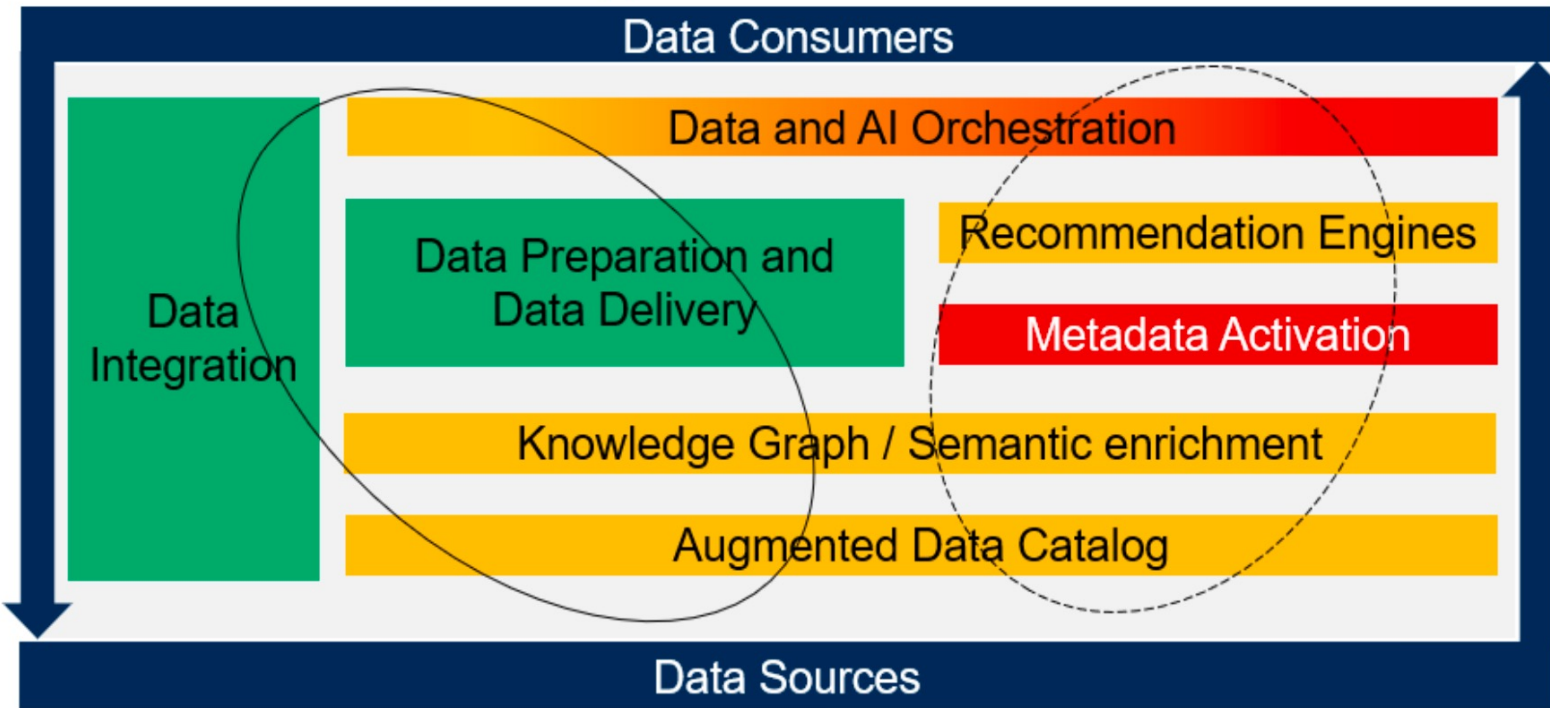




# Gartner's Evaluation of Data Fabric

## Technology Pillars of the Data Fabric Design

Market Maturity: ■ High ■ Medium ■ Low



Deployment paths:

Foundational path

Advanced path

As of June 2022:

<b>Benefit Rating</b>	Transformational
<b>Tech Maturity</b>	Emerging
<b>Market Penetration</b>	1% to 5% of Target Audience
<b>Time to Mainstream</b>	5-10 Years

Source: [Quick Answer: What Is Data Fabric Design?](#) (G00767157)

## Poll #2



# What is your current knowledge graph experience level?

- a. Explorer – curious about topic
- b. Early Adopter – starting to learn and consider applications
- c. Practitioner – applying KG in one or more projects
- d. Expert – own and maintain knowledge graph models



Cloud Data Management Capabilities

# Cloud Data Management Capabilities

## CDMC Working Group

### Cloud Challenges

- **Regulatory Mandate:** safeguard cloud computing services to protect customers' sensitive information
- **Inefficiency:** data, technology, regulatory and planning challenges on nearly every cloud implementation
- **Complexity:** 89% of firms use 2 or more cloud providers\*

### CDMC Group Objectives

1. **Best Practices** for a hybrid-cloud world
2. **Cloud Data Controls** to meet regulatory obligations for protecting Sensitive Data
3. **Accelerate Trusted Cloud Adoption** modeled after the DCAM data management industry framework

\* Source: Flexera, 2022 [State of the Cloud](#)



# CDMC Industry Engagement

100+ leading firms – 300+ SME participants – 17 GSIB's



## CDMC Working Group



## Cloud & Technology Providers



## Regulatory Engagement



- **US:** Federal Reserve, SEC, CFTC, FDIC
- **Canada:** OSFI
- **UK:** BoE, FCA, ICO
- **EU:** ECB, ESMA (pending)
- **Germany:** BaFin
- **Switzerland:** FinMA
- **Australia:** APRA
- **Singapore:** MAS
- **Israel:** Bank of Israel
- **India:** RBI, SEBI (pending)
- **Africa/Middle East:** 20+ Regulators
- Others in process...

## CDMC Adoption Support



- Training Courses
- Cloud Service Certification
- Open Source Tools
- CDMC Authorized Partner Program

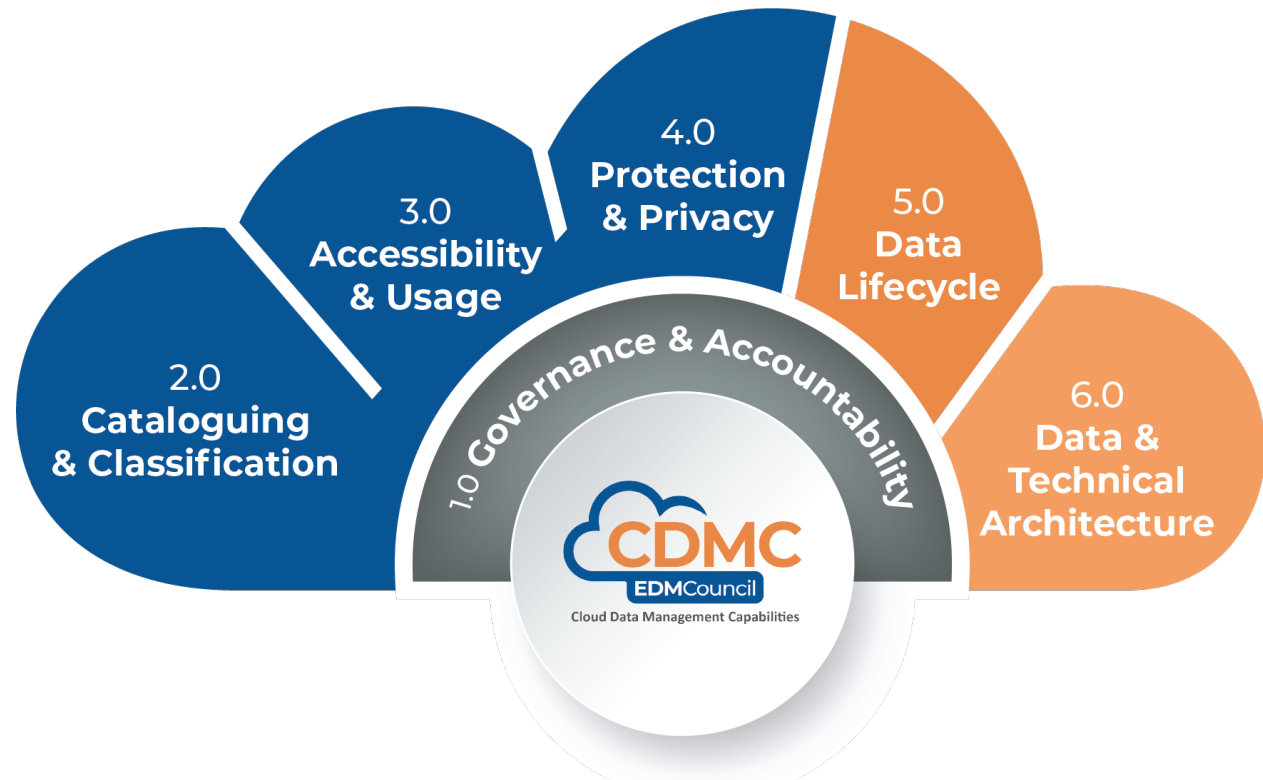
**2022**  
CDMC Framework Validated  
Cross Industry



- Life Sciences
- Telecommunications
- Manufacturing
- Retail / Services / CPG
- Consumer Tech
- Government / Defense

CDMC Working Group Delivered release v1.1  
28 September 2021

# CDMC Framework



Best practice framework for the migration, management,  
and enablement of sensitive data in cloud

**160+** Pages

**6** Components

**14** Capabilities

**37** Sub-capabilities

**14** Key Controls

Free license – Download at:  
[EDMCouncil.org](https://EDMCouncil.org) > Frameworks > CDMC

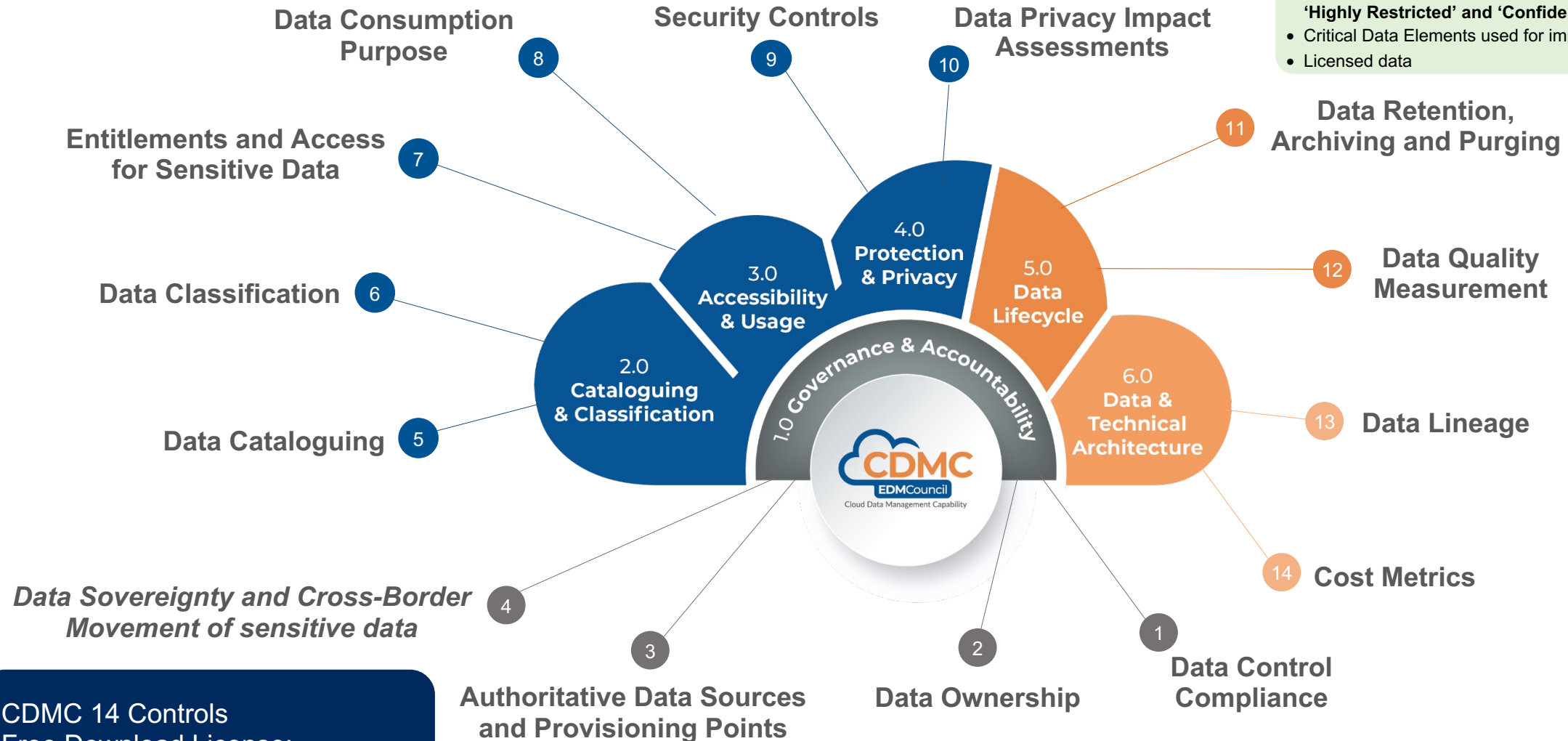


# Cloud Data Management Capabilities

## 14 Key Controls for Managing Data Risk

**Sensitive Data** includes classifications such as:

- Personal Information (PI) / Sensitive Personal Data
- Personally Identifiable Information (PII)
- Personal Health Information (PHI)
- Company or Client Identifiable Information
- Material Non-Public Information (MNPI)
- Specific Information Sensitivity Classifications (such as 'Highly Restricted' and 'Confidential')
- Critical Data Elements used for important business processes
- Licensed data



CDMC 14 Controls  
Free Download License:  
[EDMCouncil.org](https://EDMCouncil.org) > Frameworks



## Poll #3



# Why are you interested in Knowledge Graph?

- a. Analytics
- b. Fraud Detection
- c. Network Management
- d. Data Lineage
- e. Data Fabric/Catalog
- f. Data Science
- g. Other

# CDMC Information Model

## What is it?

Conceptual graph metadata model of entities and relationships required to deliver CDMC.

Simple machine-readable (RDF) vocabulary aligned to data management capabilities.

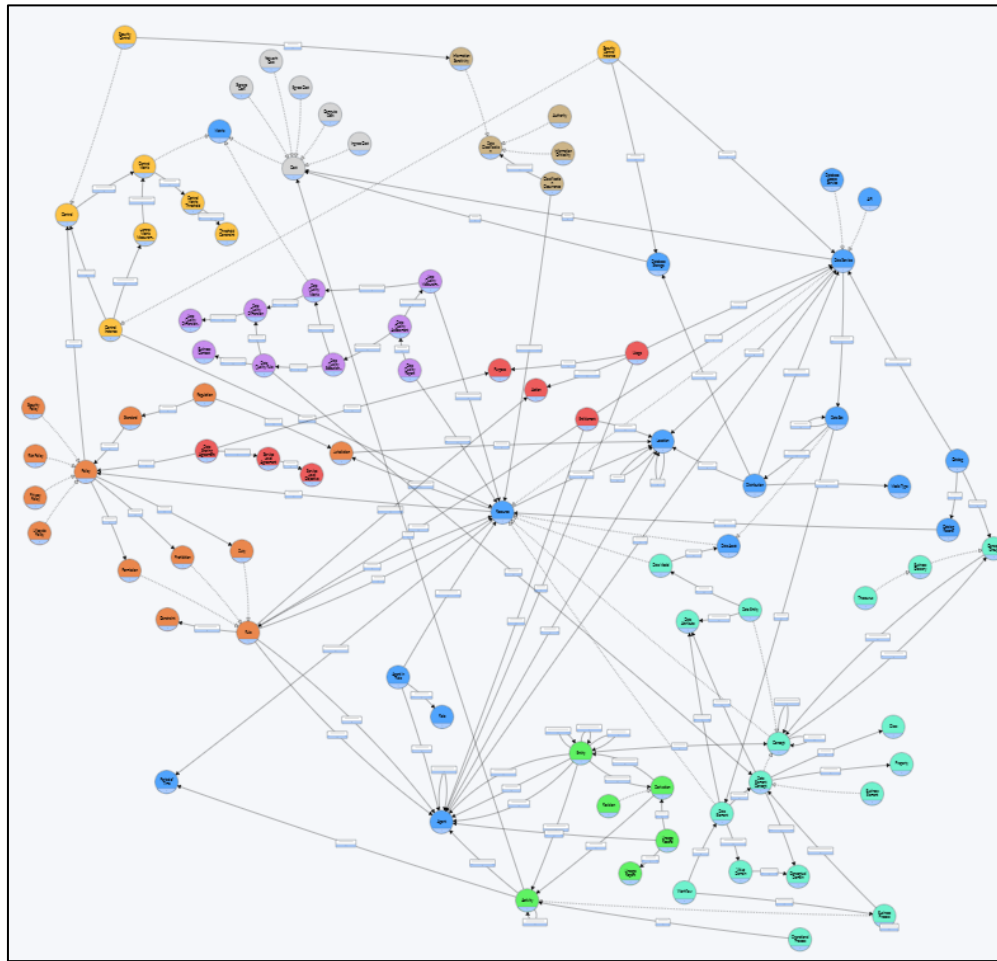
Developed by millions of geniuses

# CDMC Information Model

## Objectives

- ✔ Provide a **reference information metadata model** for designing information systems in line with the CDMC Framework
- ✔ Provide a **formal vocabulary** and point of connection between the capabilities of the CDMC framework
- ✔ Create **interoperability** between existing open standards to harmonize data sharing
- ✔ Allow practitioners to **evidence data management controls** using a consistent model across technologies
- ✔ Demonstrate how the information model **supports regulatory review**

# CDMC Information Model



## CDMC Information Model Controls Tests Mappings

Version 1.1  
October 2022

[edmcouncil.org/page/CDMC](https://edmcouncil.org/page/CDMC)



Copyright © 2022 EDM Council Inc. All rights reserved.  
Possession and application subject to CDMC™ Terms of Use.

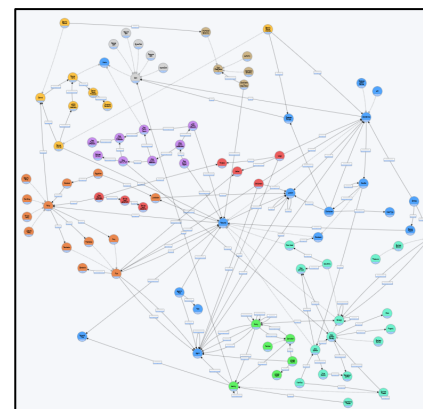
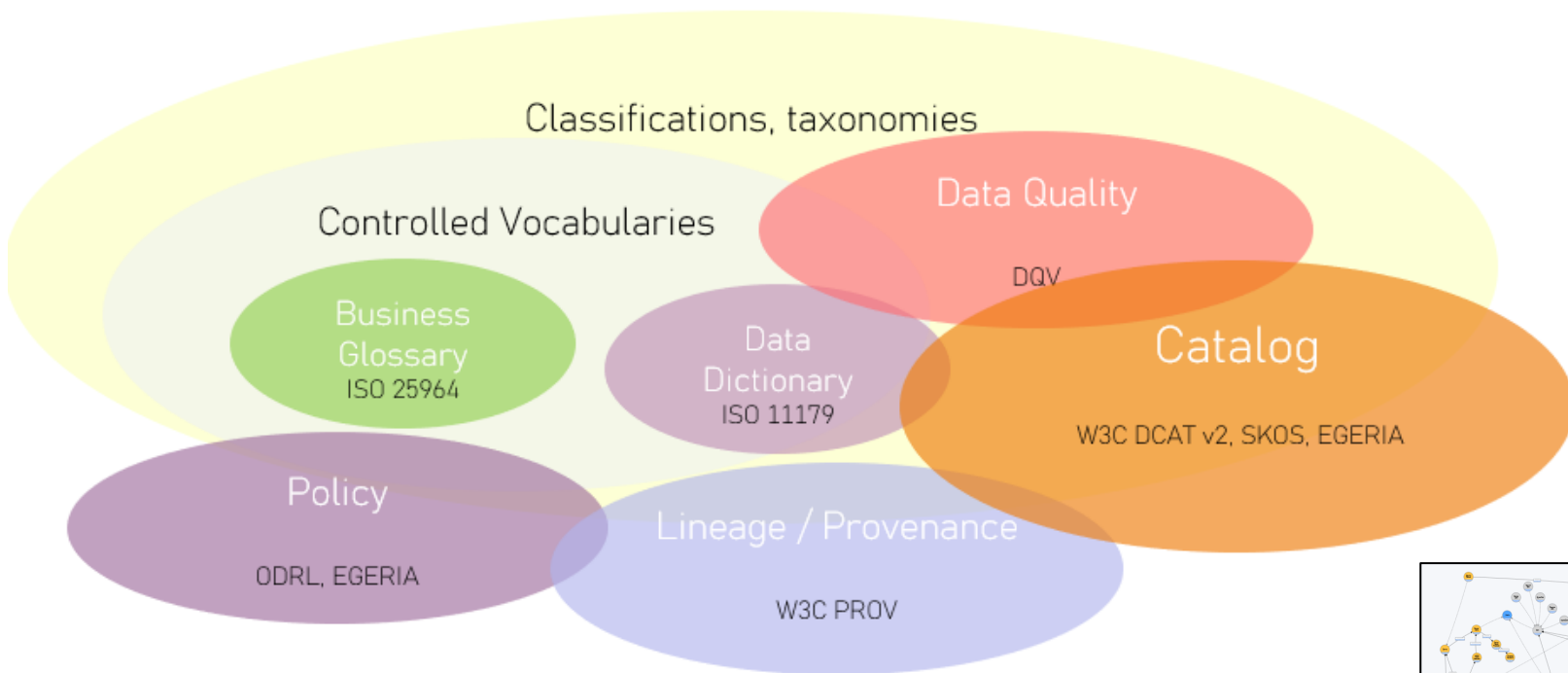






# Information Model Domains

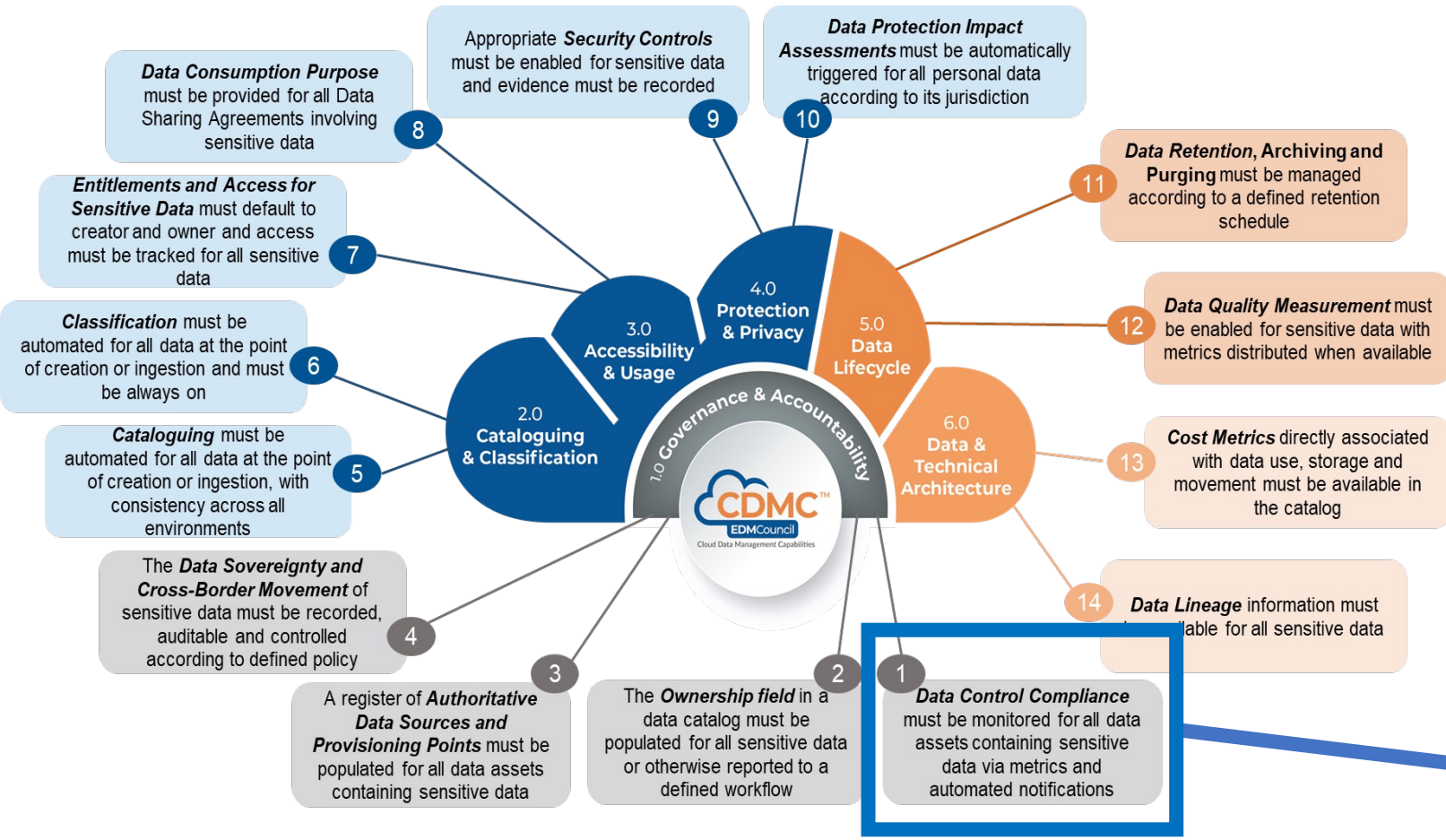
CDMC provides a top-level vocabulary based on existing and open standards



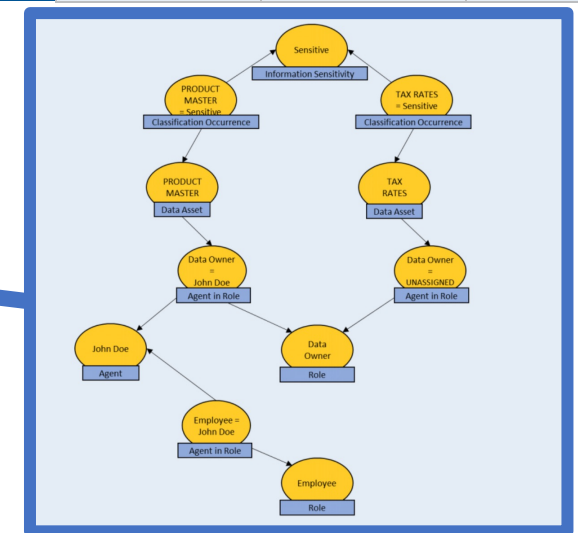
Standards usage / awareness	...
CSVW	
DCMI	
DQV	
FOAF	
PROV	
DCAM CDE	
OPA / Rego	
DCAT	
POD	
ODRL	
SDMX	
Egeria	
XACML	
ISO-11179	
ISO-2382	
ISO-25964	
+ Add another card	

Influence from 16 open vocabularies

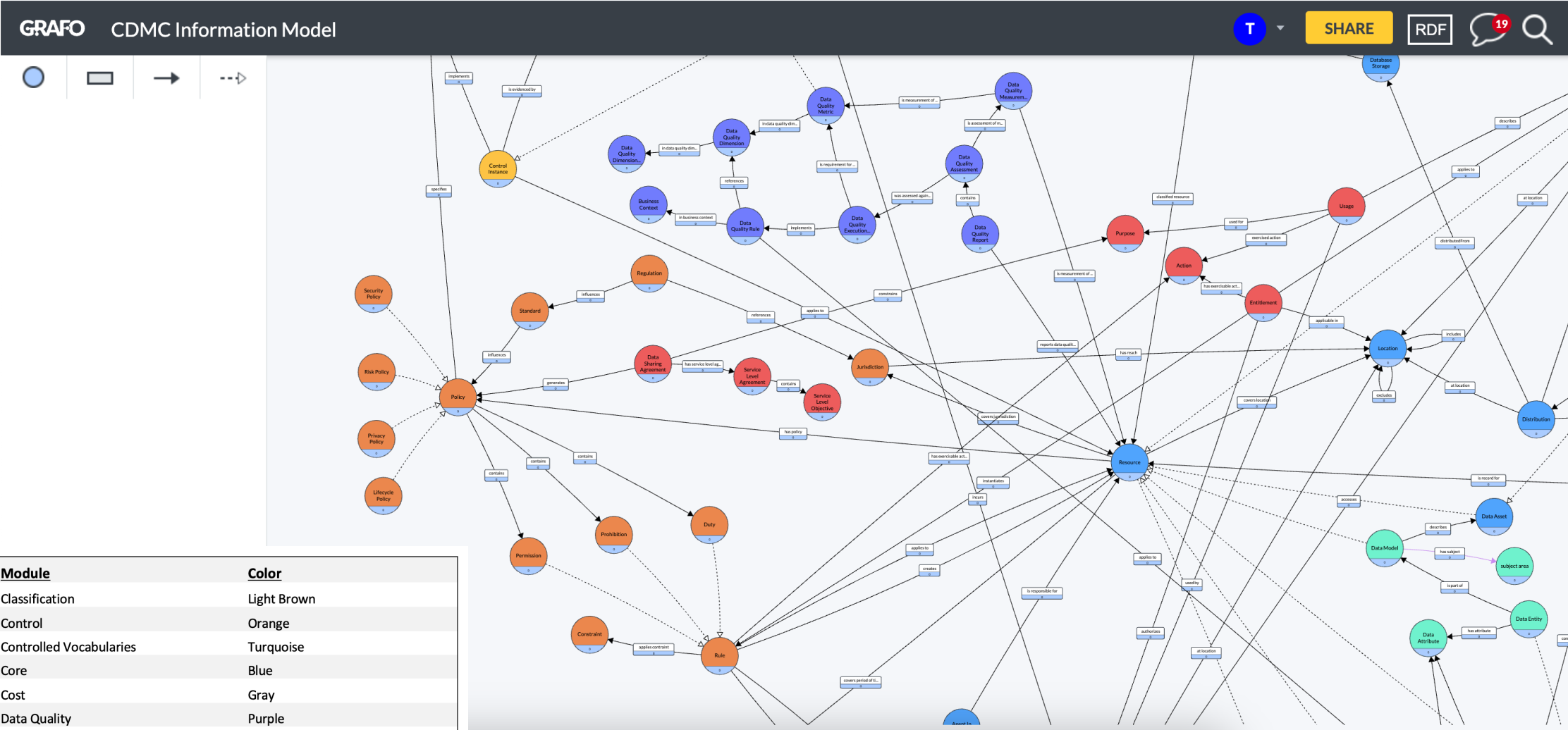
# How the Information Model Support the Key Controls



CONTROL PROCEDURE AND TEST SPECIFICATION 1: DATA CONTROL COMPLIANCE		
<b>Component</b>	<b>1.0 Governance &amp; Accountability</b>	
<b>Capability</b>	<b>1.1 Cloud data management business cases are defined and managed</b>	
<b>Control Description</b>	Data Control Compliance must be monitored for all data assets containing sensitive data via metrics and automated notifications. The metrics must be calculated from the extent of implementation of the CDMC Key Controls specified in subsequent sections.	
<b>Control Test (highlighting key terms)</b>	<ol style="list-style-type: none"> <li>1. Confirm <u>data asset</u> controls are continuously operational and tracked for sensitive assets for all other CDMC controls.</li> <li>2. Generate report/dashboard that presents the <u>metrics</u> for <u>sensitive data</u> controls.</li> <li>3. Confirm the ability to set an <u>escalation threshold</u> for each metric.</li> <li>4. Test the automated reporting against threshold limits:               <ol style="list-style-type: none"> <li>4.1 Impact the effectiveness of a <u>Key Control</u> in a manner that causes a metric threshold to be outside threshold limits.</li> <li>4.2 Confirm the automated reporting of the move outside threshold limits.</li> </ol> </li> </ol>	
<b>Term Mapping</b>	<b>Control Test Term</b>	<b>Information Model Term</b>
	Data Asset	Data Asset Resource
	Metric	Control Metric / Control Metric Measurement
	Sensitive Data	Data Asset / Classification Occurrence / Information Sensitivity
	Escalation Threshold	Control Metric Threshold
	Key Control	Control
		Notes
		"Sensitive Data" as used in the document should be implemented as defined Information Sensitivity classifications that, if applied to a Data Asset, indicate that the Data Asset should be treated as sensitive



# Information Model : Live Demo



# Snowflake, AWS & Microsoft Announce CDMC Automated Data Risk Control Framework



**SNOWFLAKE ACHIEVES FIRST CLOUD DATA MANAGEMENT CAPABILITIES (CDMC) ASSESSMENT**

SEP 14, 2021

September 2021



December 2022



March 20, 2023 • 5 min read

Microsoft achieves first native Cloud Data Management Capabilities certification

Mike Flasko Vice President and General Manager, Data Governance

March 2023





# Google Cloud becomes the next CDMC Cloud Certified Solution



Security & Identity

## How to migrate sensitive data with confidence using Google Cloud's CDMC-certified architecture

June 26, 2023



ANNOUNCEMENTS



## EDM Council Designates Google Cloud as a Certified Cloud Solution

Mon, Jun 26, 2023

Google Cloud clients migrating sensitive data to the cloud can do so with greater confidence now that Google Cloud has been certified against EDMC's CDMC control framework

New York, NY – June 26, 2023 – EDM Council, the leading global non-profit trade association for advocating data management and analytics, has today announced that Google Cloud's new solution architecture, which combines Google Cloud's BigQuery and Datanplex, has been independently

Recent Announcements [See all](#)



Aldefi becomes the newest member to join the EDM Council

# BT becomes the first telecommunications organization to complete CDMC assessment



Video Whitepaper CDO Awards Event M

**CDO TRENDS**  
Digital & Data Insights for Business Leaders

A.I. Digital Strategy Data Science Digital Security NextGenConnectivity D

DATA SCIENCE | DATA MANAGEMENT

## This New Framework Slashes Data Management Redundancies

By CDO Trends editors on June 19, 2023

[f](#) [in](#) [t](#)

An isometric illustration of a data management architecture. It features several blue and red 3D blocks of varying heights, some with data charts and graphs on their surfaces. A central red cube is the largest and most prominent. A network of lines connects various points across the blocks. Small human figures are shown interacting with the blocks, suggesting a collaborative data management process.

EDMCouncil

### CDMC panel

				
<b>Aravind Jagannathan</b> Vice President Freddie Mac	<b>Bala Ayyar</b> Chief Data Officer - Americas Société Générale	<b>Wes Anderson</b> VP, Enterprise Data Management TD Bank	<b>Ben Clinch</b> Principal Enterprise Architect BT Group	<b>Keeley Miles</b> Global Head of Data Management Ford Credit
				

## Poll #4



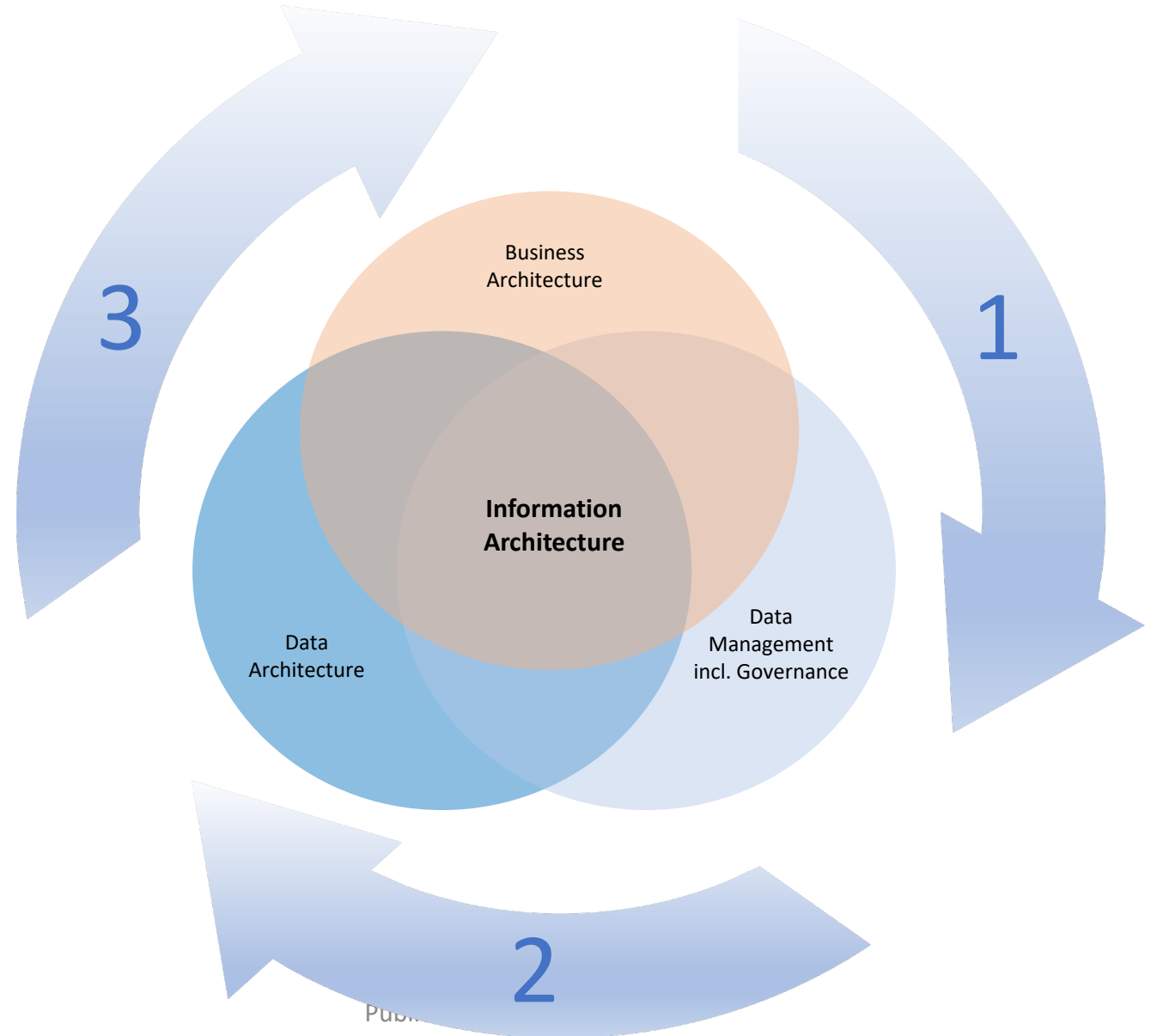
# What are your current blockers to knowledge graph adoption in your organization?

- a. Lack of expertise
- b. No clear path to adoption
- c. Unclear business benefits / ROI
- d. Insufficient funding
- e. Lack of stakeholder buy-in
- f. Confusion regarding tooling and integration

# The Data Fabric and Data Mesh movement

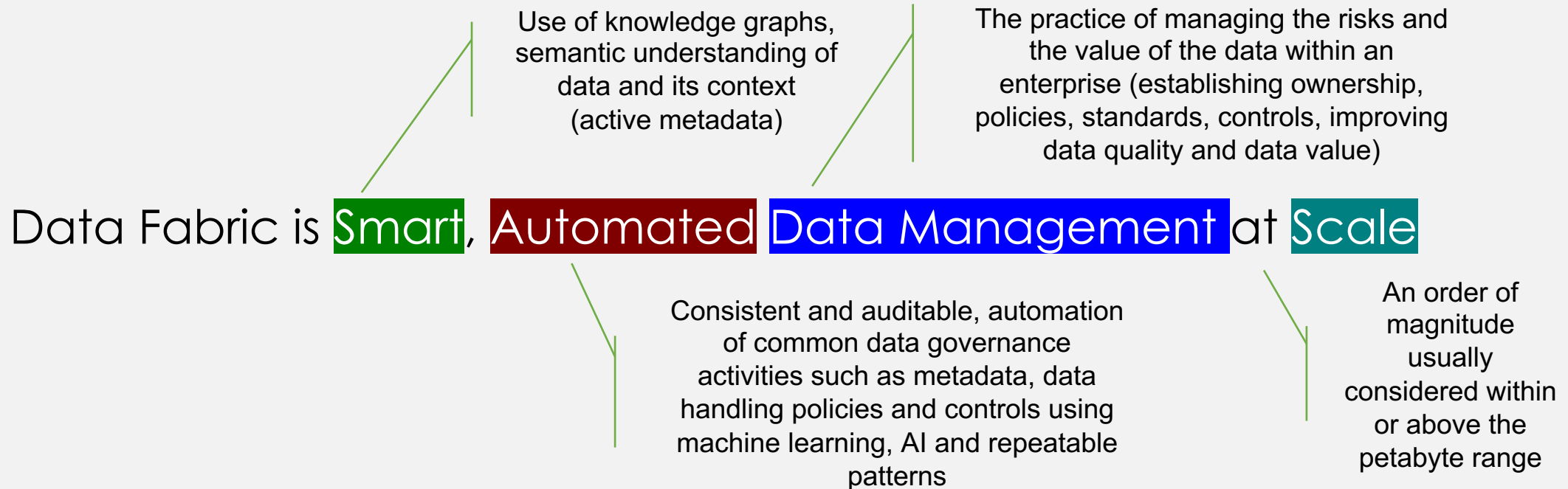
## What it means for collaboration

“Data Fabric” and “Data Mesh” as terms, point towards Information Architecture thinking bringing together disciplines across the business





# Data Fabric: A practical definition for BT



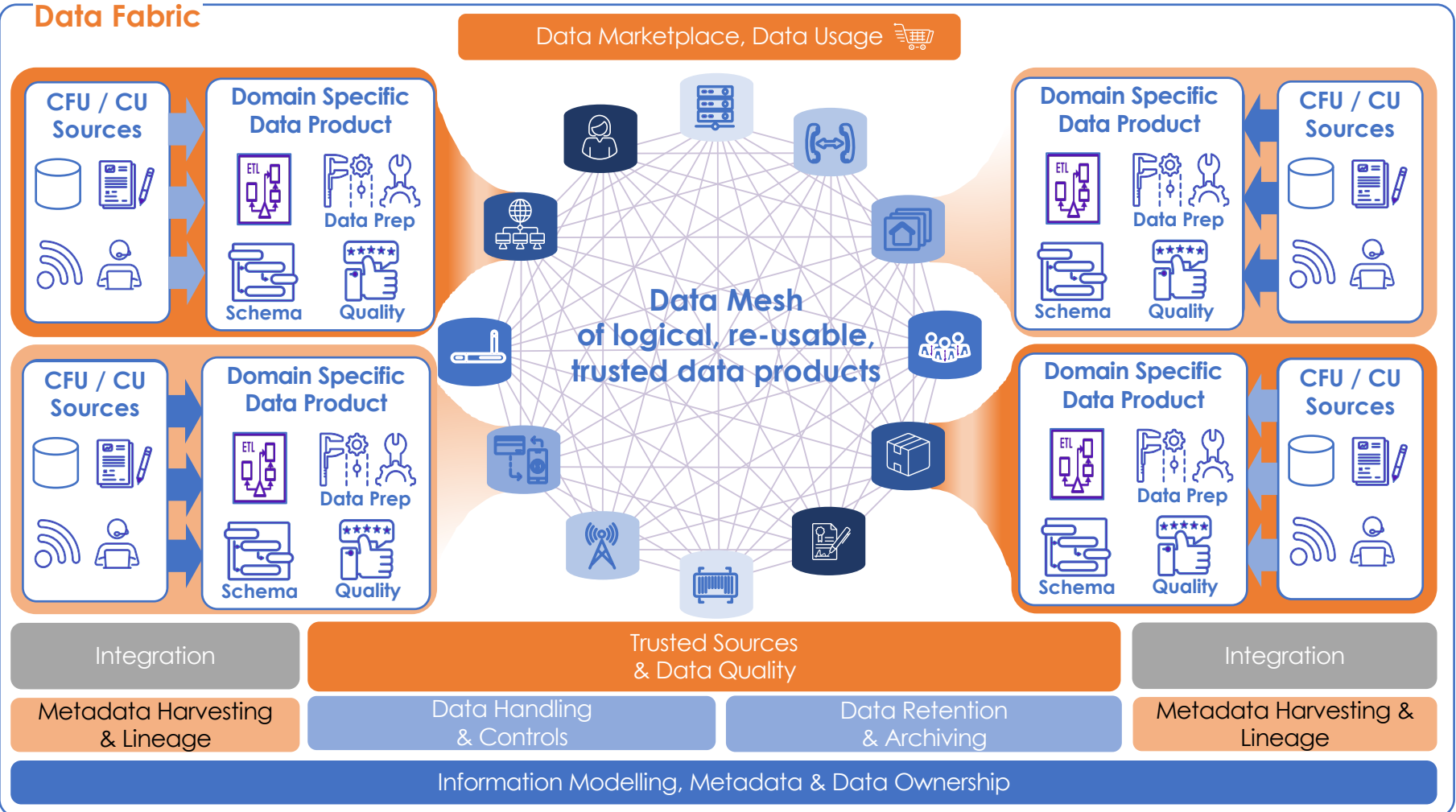
Point to Note: Data Fabric doesn't replace the need for data management; it automates it

# Wrapping the Fabric around the Mesh

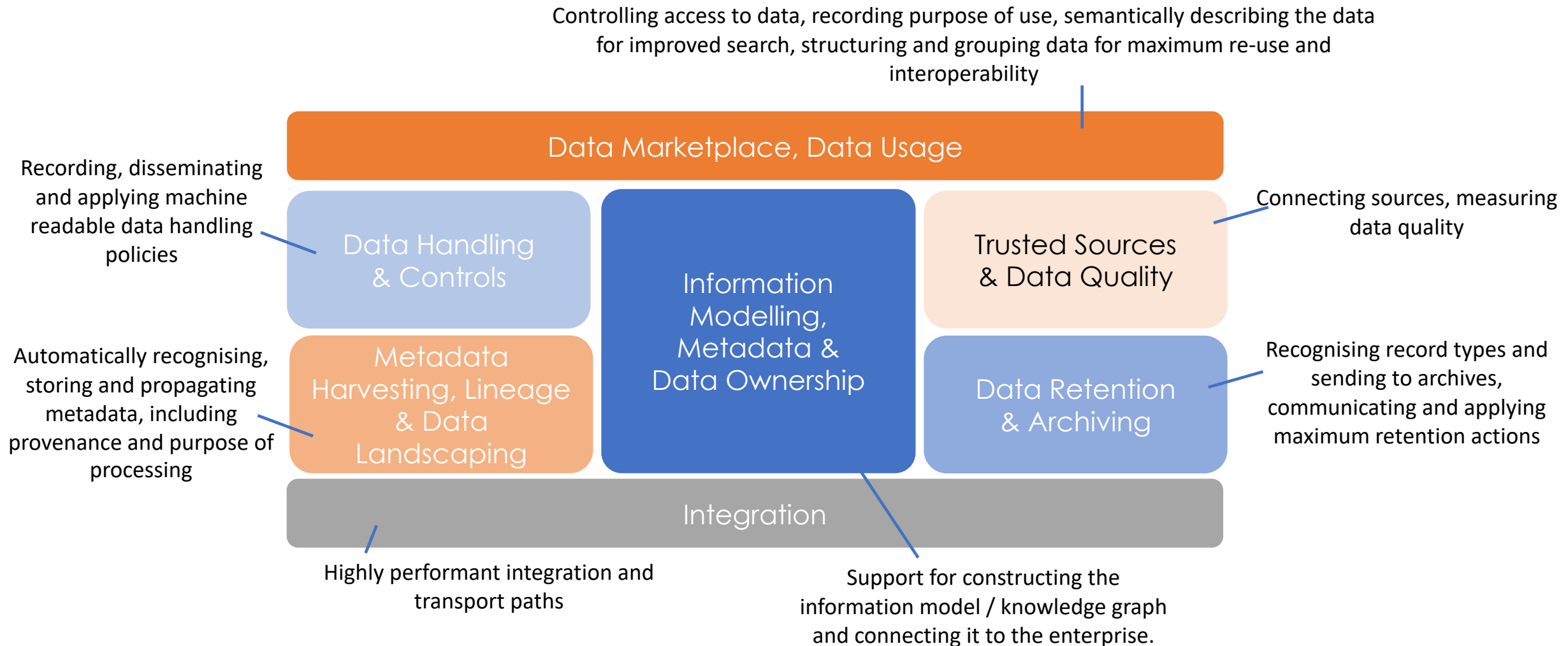
The data fabric wraps sustainable, automated data governance around the data mesh.

The fabric provides policies, standards and tools to make federated data governance consistent and work for all of BT.

## Analytic & Operational Planes



# Information Architecture Conceptual View: A simplified view

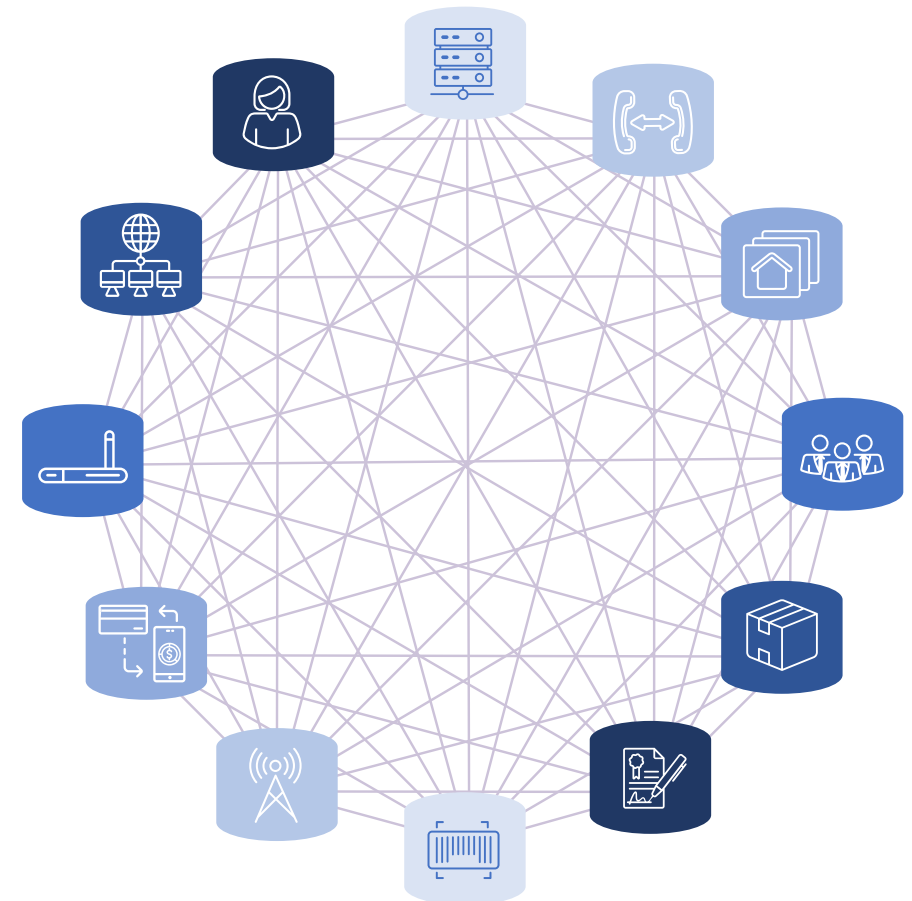


# Vision for the Information Model

*“The Information Model is for our data what the organisation chart is for our people.*

*In a digital company data is one of our most valuable assets.*

*The information model needs to be as pervasive in our everyday thinking and activities as the org chart.”*



The Information Model (a knowledge graph) provides Data Fabric with a structure and a rich set of semantic meaning that can be used to inform entity resolution and automate controls while simplifying the way we manage data.



# Information Model Usage: A Worked Example

Instance Data Terms	Information Model Term	Definition
	Classification Occurrence	instance of applying a Data Classification to a Resource
	Control	defined way to ensure control of data assets
	Control Instance	implementation of a Control for a Data Asset
	Control Metric	definition of a measurement that determines the effectiveness of a Control
	Control Metric Measurement	value of a Control Metric at a point in time
	Control Metric Threshold	value against which a Control Metric Measurement is compared to assess the effectiveness of a Control
	Data Asset	collection of data owned by an organization that is considered to have intrinsic value
	Information Sensitivity	classification of the information within a data asset that indicates the level of control and protection that must be applied to the asset due to the nature of the data and its sensitivity or importance to the organization or to the subject of the data
	Threshold Constraint	relationship between a Control Metric Measurement and the corresponding Control Metric Threshold that determines the effectiveness of a Control

# CDMC Information Model

## Summary

- ✓ How a distributed architecture benefits from common structures and standards (a.k.a. the information model)
- ✓ Four modes in which you can use the information model and the comparative benefits:

**Reference** – Authoritative interpretation of the metadata requirements for CDMC compliance

**Mapping** – Scalable implementation to produce a catalog of catalogs (Key Control 1)

**Digital twin** – Smart query across data sets that is consistent with the business concepts

**Interoperability** – Integration with vendor solutions and harmonize data sharing across supply chain



Cloud Data Management Capabilities

# Cloud Data Management Capabilities

**Thank You!**

## **Opportunities for Industry Collaboration**

- **New Knowledge Graph Course – Q4 2023**
- **Volunteer to help w/enhancement of the Information Model extensions to DCAM v3**
- **Volunteer to test the CDMC Information Model for existing use cases to improve or enhance the model**

- **CDMC Information Model Graph Viewer**



- **CDMC Information Model Test Mapping**

