EDMCouncil

www.edmcouncil.org

Empowering the Modern Data Platform

June 3, 2021

2:30 PM UK

9:30 AM EDT

Today's speakers

Mike Meriton
Co-Founder & COO
EDM Council

Shaun Rolls
Regional Advocate:
Continental Europe
EDM Council

Eric Bigelsen
Senior Advisor, Industry
Engagement & Project Leader, ESG
EDM Council

Today's Agenda – The Connected Modern Data Platform

1. About EDM Council
2. Cloud Data Management
3. Graph & Knowledge Graphs
4. ESG & Sustainability Data
5. Q&A

About EDM Council

EDM Council Overview

Mission: ***Elevate the practice of data and analytics management***

Principles:

We believe in a **responsible data-driven** organization and establishing **data literacy for all.**

We advocate **industry collaboration** to provide data management and analytics research, best practices, standards, training and education.

We are the **neutral, non-profit** industry forum for companies and their data & analytics professionals.

Established
2005

Worldwide
Americas, Europe, Africa, Asia, Australia

250+ Member Firms
Cross-industry, including Regulators

10,000+
Professionals

EDM Council's Areas of Advocacy

Best Practices

- DCAM & Cloud CDMC
- ESG Data
- Data ROI

Driving Standards

- Knowledge Graph
- Industry Ontologies
- Shared Lab

Training & Certification

- Virtual & eLearning Courses
- EDMWebinars & Events

Research & Benchmarking

- Global Industry Study
- Life Sciences
- Data Sharing

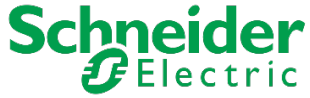
Regulatory Engagement

- Regulators participate in the agenda
- Many regulators are members

Networking

- EDMConnect Community
- DataVision, CDO Summit
- Workgroups & Forums

EDM Council – New Members



EDM Council Global Membership

Best Practices Framework

Data Management Capability Assessment Model

WHO?

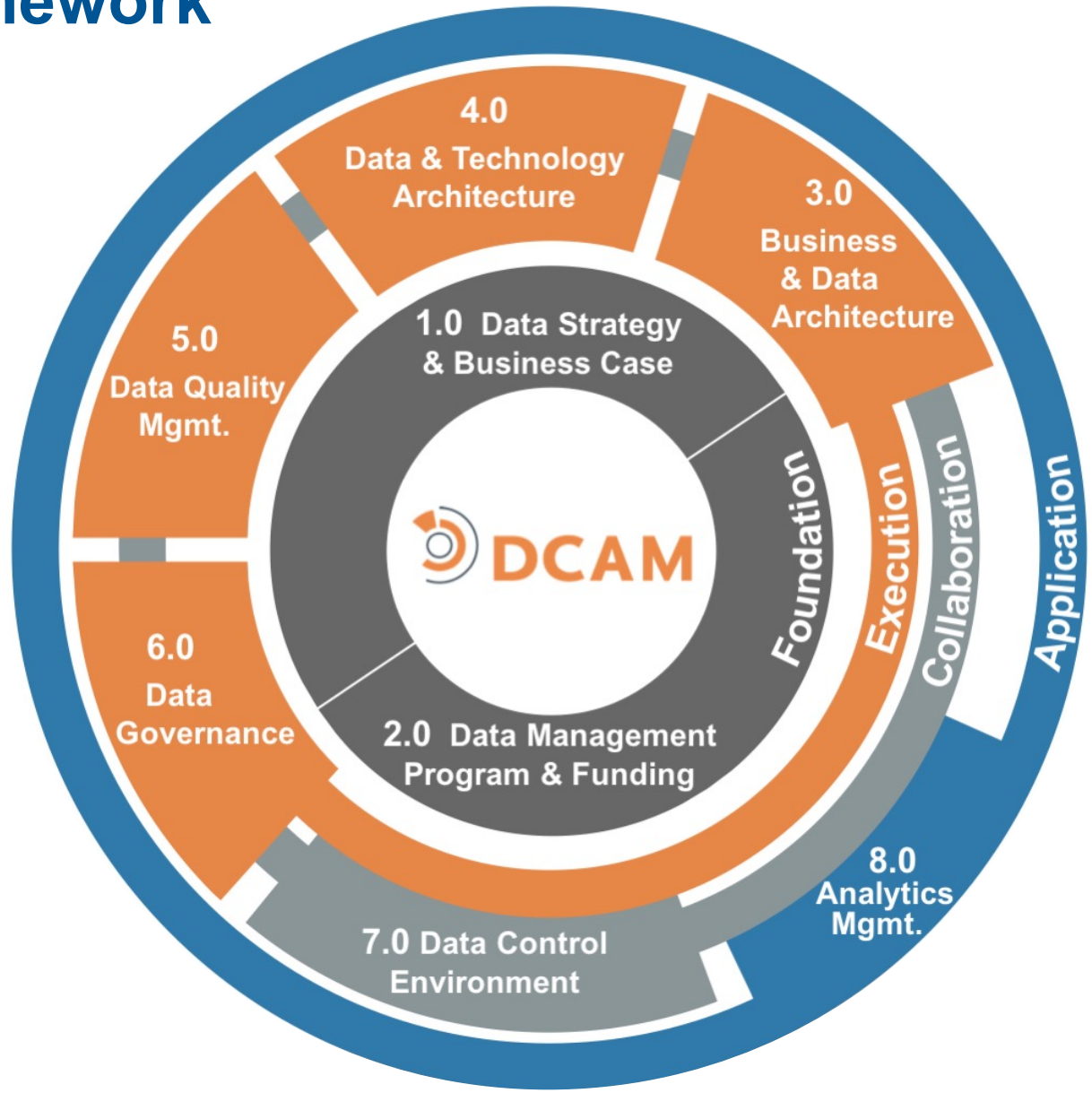
- Developed via member collaboration
- **62%** of Council members using frameworks use DCAM

WHAT'S IN DCAM?

- 8 components, 38 Capabilities, 136 Sub-capabilities
- Members flexibly apply to their organization
- Includes: Data Supply Chain, Advanced Analytics, Data Ethics and Responsible AI/ML

USED FOR:

- Program Initiation & Funding
- Team Training & Common Language
- Assessments & Benchmarking



eLearning: Training & Certification

eLearningCurve Online Data Academy

- **9 Information Management Tracks with 3 Professional Certification Programs**
- **50+ online courses and 200+ hours of education materials**

Information Management Track Curricula	
IM Foundations	Data Integration
Data Quality	Master Data Management
Data Governance	Data Modeling & Metadata
Data Stewardship	Business Intelligence & Analytics
Data Science	

Certification Programs
NEW on eLearning! – **DCAM (Data Management Capability Assessment Model)**
Certified Information Management Professional (CIMP)
Certified Data Steward (CDS)
Data Literacy (2021)

To register: EDMCouncil.org > Training > E-Learning

Cloud Data Management Capabilities (CDMC)

Cloud Data Management Capability (CDMC)

CDMC Workgroup

CDMC™
EDMCouncil
Cloud Data Management Capability

Cloud Challenges

- **Inefficiency:** data, technology, regulatory and planning challenges on nearly every cloud implementation
- **93% of firms** use 2 or more cloud providers*

CDMC Group Objectives

1. Define consistent best practices for a hybrid-cloud world
2. Align key cloud data controls to meet regulatory obligations for Sensitive Data
3. Accelerate Cloud Adoption with comprehensive framework

EDMCouncil

* Source: Flexera, 2020 State of the Cloud

CDMC: Industry Engagement

70+ Leading firms and 226 participants
actively participating since ***May 2020***

CDMC™ EDMCouncil
Cloud Data Management Capability

CDMC Workgroup

LSEG · Morgan Stanley · J.P.Morgan · CAPCO · Standard Chartered · KPMG · UBS · HSBC · citi · Goldman Sachs · PayPal · Freddie Mac (We make home possible®) · WELLS FARGO · TP ICAP · NORTHERN TRUST · BARCLAYS · Deutsche Bank · LLOYDS BANK · Tradeweb · TD Bank · BNP PARIBAS · CREDIT SUISSE · Nasdaq · SOCIETE GENERALE

Cloud & Technology Provider Certification

aws · Microsoft Azure · Google Cloud · IBM Cloud · snowflake · collibra · Informatica · BigID · data.world · PRIVITAR

Regulatory Engagement

EDMCouncil

- **US:** Federal Reserve, OCC, FDIC, NCUA, NAIC
- **Canada:** OSFI, BoC, CDIC
- **UK:** BoE, FCA, IOC
- **EU:** EBA, ECB, DORA Act
- **Germany:** BaFin
- **Swiss:** FinMA
- **Japan:** FSA
- **Australia:** APRA
- **Singapore:** MAS
- **India:** RBI, SEBI
- **Africa/Middle East** regulators
- Others

Go-to-market Support

EDMCouncil

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

2H 2021 – 1H 2022 Other Industries

EDMCouncil

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

CDMC Workgroup Publication on track for Q3 2021

EDMCouncil

Cloud Data Management Capabilities (CDMC) Key Controls

Sensitive *Data* includes classifications such as:

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

Security Controls must be enabled for sensitive data and evidence must be recorded — 4

Data Privacy Impact Assessments (PIAs) must be automatically triggered for personal data according to its jurisdiction — 4

Data Consumption Purpose is required for all Data Sharing Agreements involving sensitive data — 3

Entitlements and Access for Sensitive Data must be defaulted to creator and owner and all access must be tracked — 3

Classification must be automated for all data — 2

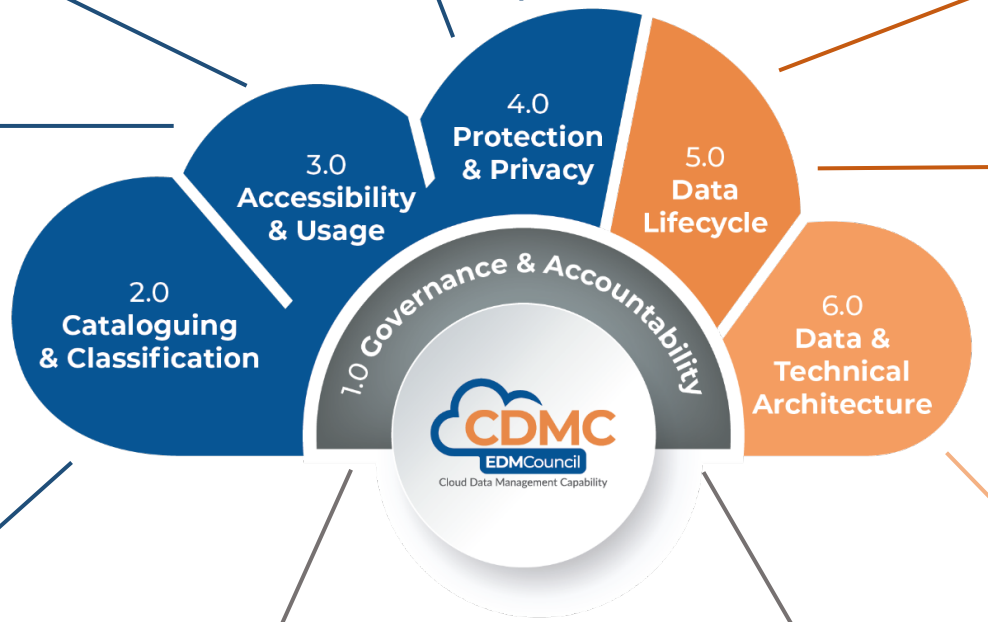
Cataloguing must be automated for all data — 2

Data Quality Metrics are required for all sensitive data — 5

***Data Retention*, Archiving and Purging** must be managed according to the schedule that applies to its classification — 5

Data Lineage information is required for all sensitive data — 6

Cost Metrics directly associated with the use, storage and movement of data must be available in the catalog — 6



The ***Authorization field*** in a data catalog must be populated for all data assets containing sensitive data — 1

The ***Ownership field*** in a data catalog must be populated for all sensitive data — 1

EDMCouncil

CDMC Capabilities: Requirements & Automated Controls

Component	Capability	Sub-Capability	CDMC Automation Checklist (DRAFT)
1. Data Accountability & Governance	**1.1 The Cloud Data Management business case** is defined and measurable	1.1.1 TBA	
	1.2 Data ownership established for both migrated & cloud-generated data	1.2.1 Data Owner roles and responsibilities are defined and agreed 1.2.2 Data ownership is established for newly cloud generated data	• Ownership field required for specified classifications (eg PII, MNPI, ISC=HR|C)
	1.3 Data sourcing and consumption are governed and supported by automation	1.3.1 Data sourcing is governed for accuracy, data duplication control and data protection 1.3.2 Data consumption is governed for accuracy, trustworthiness and usage 1.3.3 Data sources are authorized and certified 1.3.4 Data provisioning and consumption are standardised and automated	• Authorisation field required on cataloged data assets for specified classifications (eg PII, MNPI, ISC=HR|C)
	1.4 Data Sovereignty, Residency and Localisation are actively managed	1.4.1 TBA	
2. Data Cataloguing & Classification	**2.1 Data catalogues** are implemented, used and interoperable	2.1.1 Data cataloguing is defined, scoped and actively used 2.1.2 Metadata is discoverable, enriched, managed and exposed in Data Catalogues 2.1.3 Data catalogues are interoperable across multi and hybrid cloud environments	• Auto Cataloguing Supported
	2.2 Data classifications are defined and used	2.2.1 Data classifications are defined, applied and actively used	• Auto Classification Supported • Personal Information auto-discovery • ISC auto-discovery (HR|C|IUO|P) • MNPI auto-discovery • Client Information auto-discovery • Additional classifications
3. Data Accessibility & Usage	**3.1 Data entitlements** are managed, enforced and tracked	3.1.1 Data entitlement rights and obligations are captured as metadata 3.1.2 Data entitlement rights are enforced 3.1.3 Enforcement of data entitlement rights is evidenced	• Entitlements automatically restricted to creator & owner for sensitive data (ISC=HR|C|IUO) • Access tracked for all sensitive data (ISC=HR|C|IUO)
	3.2 Ethical use and purpose of data use are tracked	3.2.1 TBA	• Purpose required for all DSAs for all PII & sensitive data (ISC=HR|C|IUO)
4. Securing Data & Privacy	**4.1 Data is secured** and controls are evidenced	4.1.1 Data is encrypted at rest, in motion and while in use 4.1.2 Implementation of data security controls is evidenced 4.1.3 Data obfuscation techniques are defined, scoped and applied based on risk and utility 4.1.4 A Data Loss Prevention regime is in place	• Security controls and DLP are automatically enabled for all sensitive data (ISC=HR|C|IUO) • Security control evidence is sent to Catalog for all sensitive data
	4.2 A data privacy framework is defined and operational	4.2.1 A data privacy framework is defined and agreed 4.2.2 The data privacy framework is operational	• Recommended data privacy framework workflows (eg jurisdictionally aware PIAs) are automatically triggered for all personal data
5. Data Lifecycle Management	**5.1 Data quality** is managed	5.1.1 Data Quality Rules Management 5.1.2 Data Quality Profiling and Validation 5.1.3 Data Quality outputs publication 5.1.4 Data Quality issue management	• Data Quality metrics information required for specified classifications (eg PII, MNPI, ISC=HR|C)
	5.2 The data lifecycle is planned and managed	5.2.1 Archiving and purging of data are managed	• Data meeting specific retention schedule classifications is automatically retained, archived and purged
6. Data & Technical Architecture	**6.1 Data provenance and lineage** are understood	6.1.1 TBA	• Provenance information required for specified classifications (eg PII, MNPI, ISC=HR|C)
	6.2 Technical design principles are established and applied	6.2.1 Data back-up, resilience & recovery - (Rewording TBA) 6.2.2 Portability & exit plan - (Rewording TBA) 6.2.3 Usage & cost transparency - (Rewording TBA)	• Well architected framework evaluated metrics • Costs directly associated with managing data assets are available to view by Data Owners in the catalog

CDMC: Accelerating Cloud Adoption

CDMC
EDMCouncil
Cloud Data Management Capability

CDMC Industry Objective:

*Build **Trust, Confidence, and Dependability** for Cloud Adoption*

Companies

CDMC structured framework of auditable Cloud processes and controls – especially for sensitive data

Cloud Service Providers (CSP)

CDMC requirements and controls can be automated into CSP platforms which accelerates adoption and provides market confidence

Application, Technology & Data Providers

Incorporates certified CDMC capabilities and controls into services and solutions to ensure high degree of reliability and operational effectiveness

Consultants & Systems Integrators (SI)

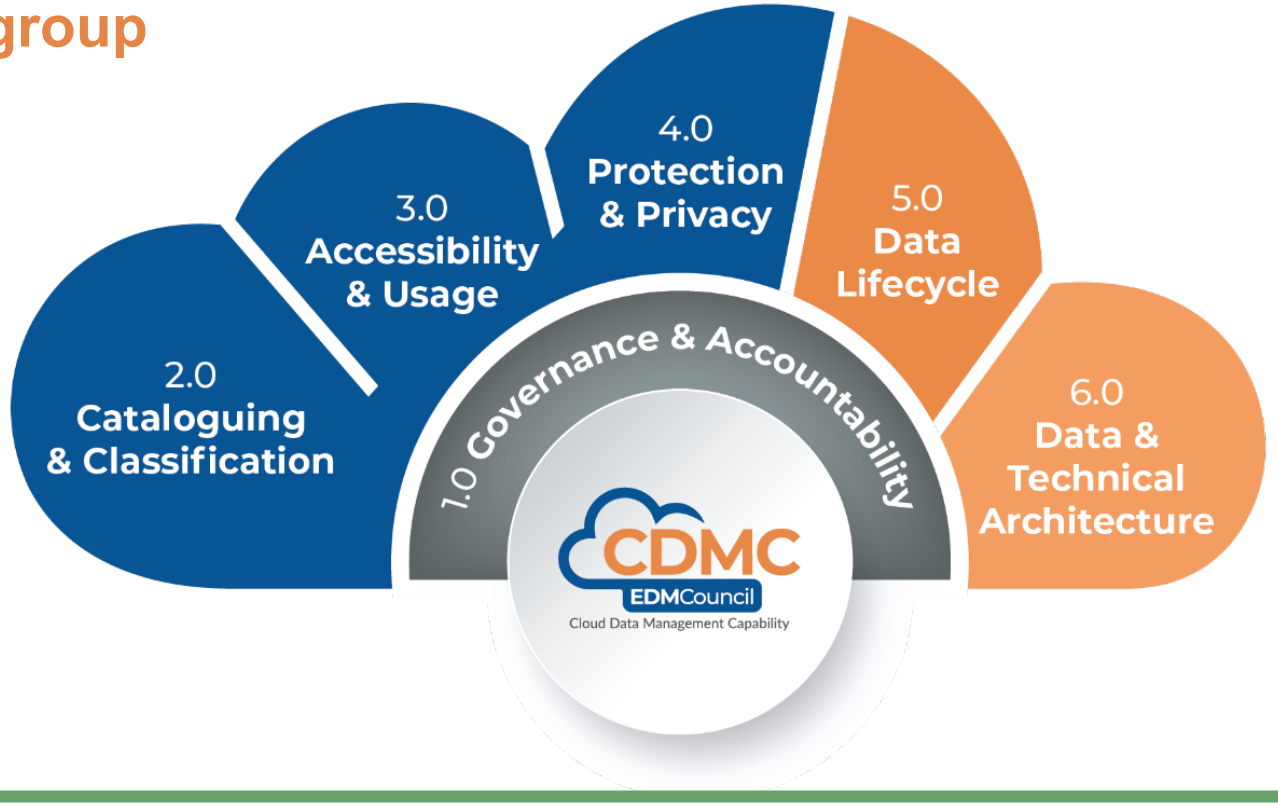
CDMC enables training & assessments, gap analysis, strategy development, and execution services for end clients adopting cloud

Regulators

CDMC provides industry guidance for auditing and validating key cloud controls, especially for sensitive data

Get involved with CDMC

CDMC Open Workgroup



Learn more or join the Open CDMC Group

www.EDMCouncil.org > Click on **Groups** > **Cloud…**

Data ROI Workgroup

Data ROI: Goal, Objectives and Expected Outcomes

Goal

Develop a Data ROI framework / template to support data and analytics organizations

---

Objectives

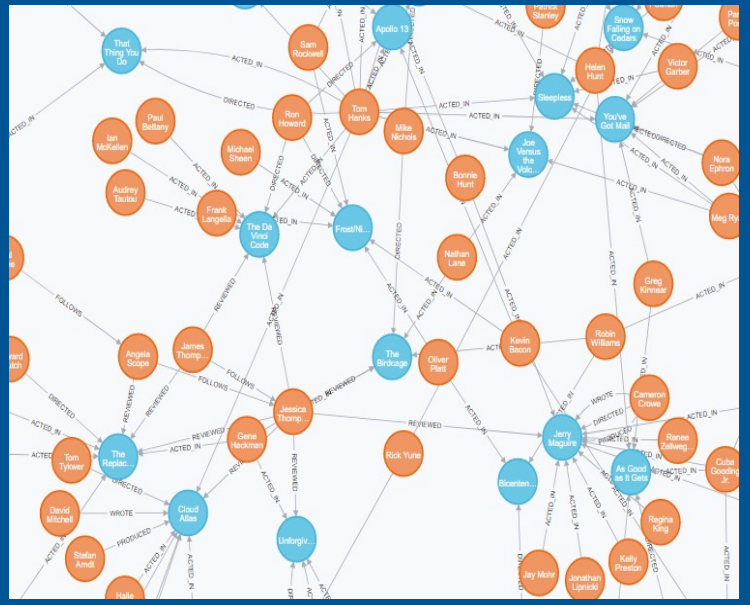
Data ROI framework to answer three proposed data use cases:

Data Program ROI	Data Projects ROI	Data as a Balance Sheet Asset

Get involved: www.EDMCouncil.org – Go to Groups > Data ROI

Graph & Knowledge Graph Architecture

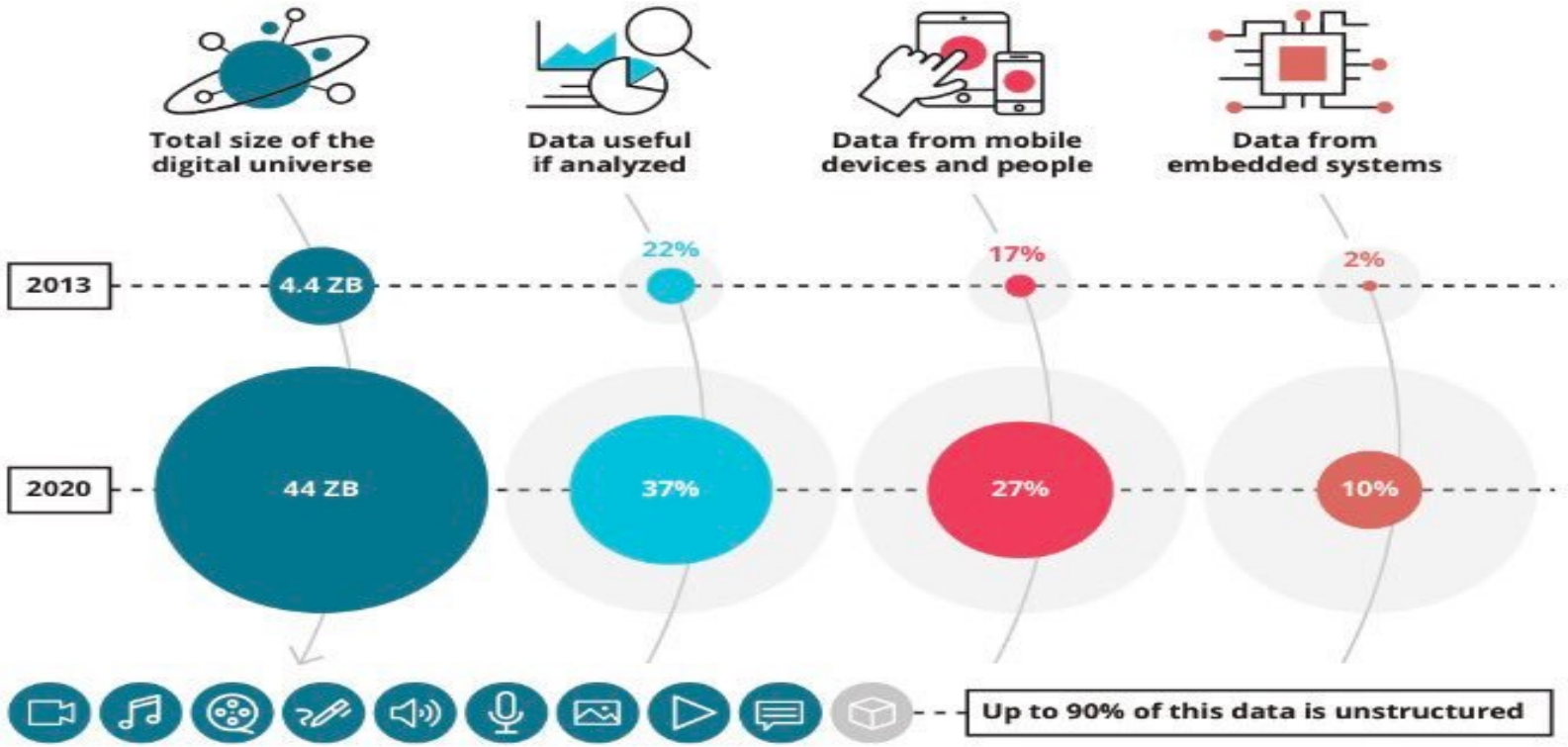
Open Knowledge Graph



EXPLOSION OF DATA VOLUME & COMPLEXITY

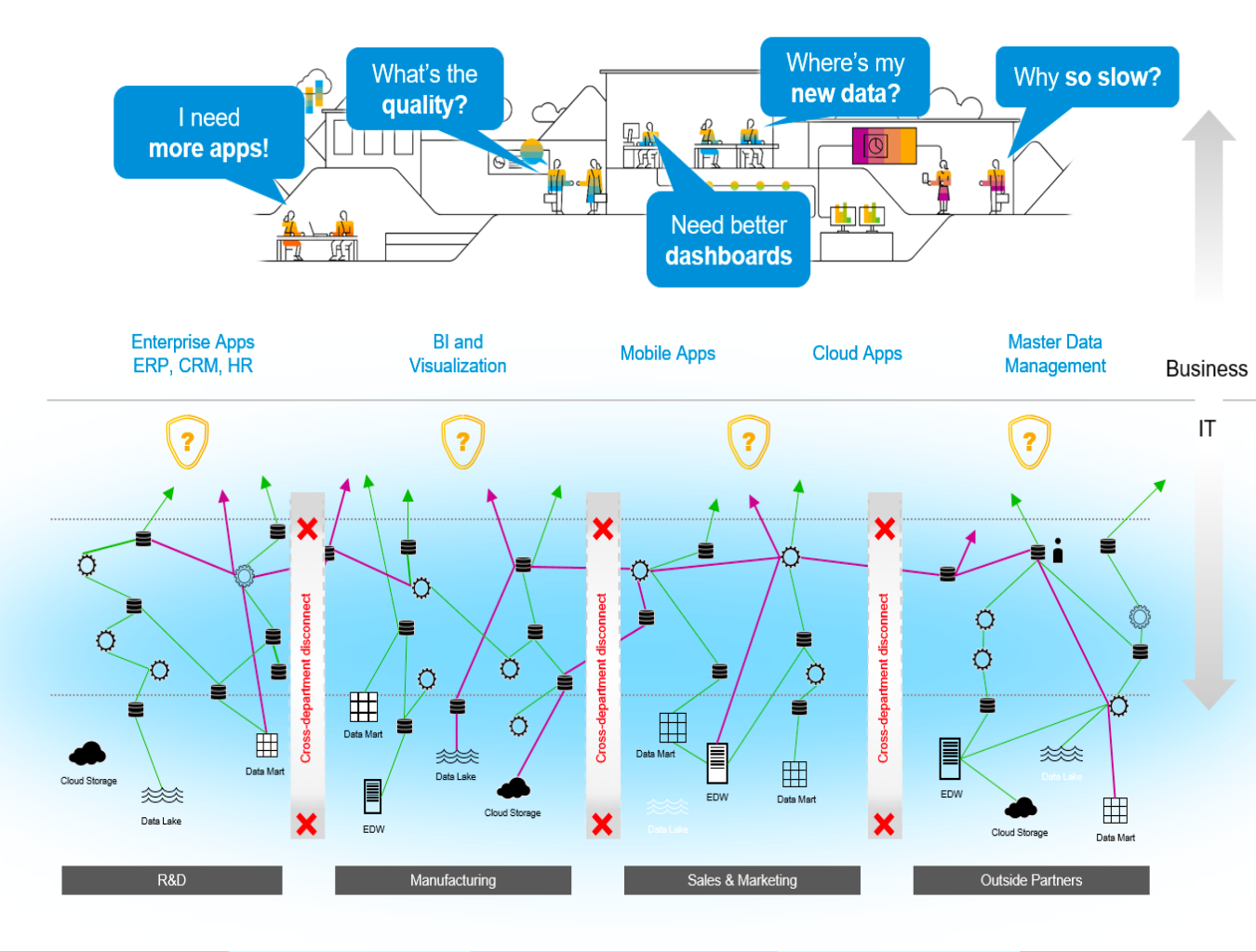
Figure 1. The expanding digital universe, 2013–2020

In 2020, the digital universe is expected to reach 44 zettabytes. One zettabyte is equal to one billion terabytes. Data valuable for enterprises, especially unstructured data from the Internet of Things and nontraditional sources, is projected to increase in absolute and relative sizes.



Source: Deloitte, 2017

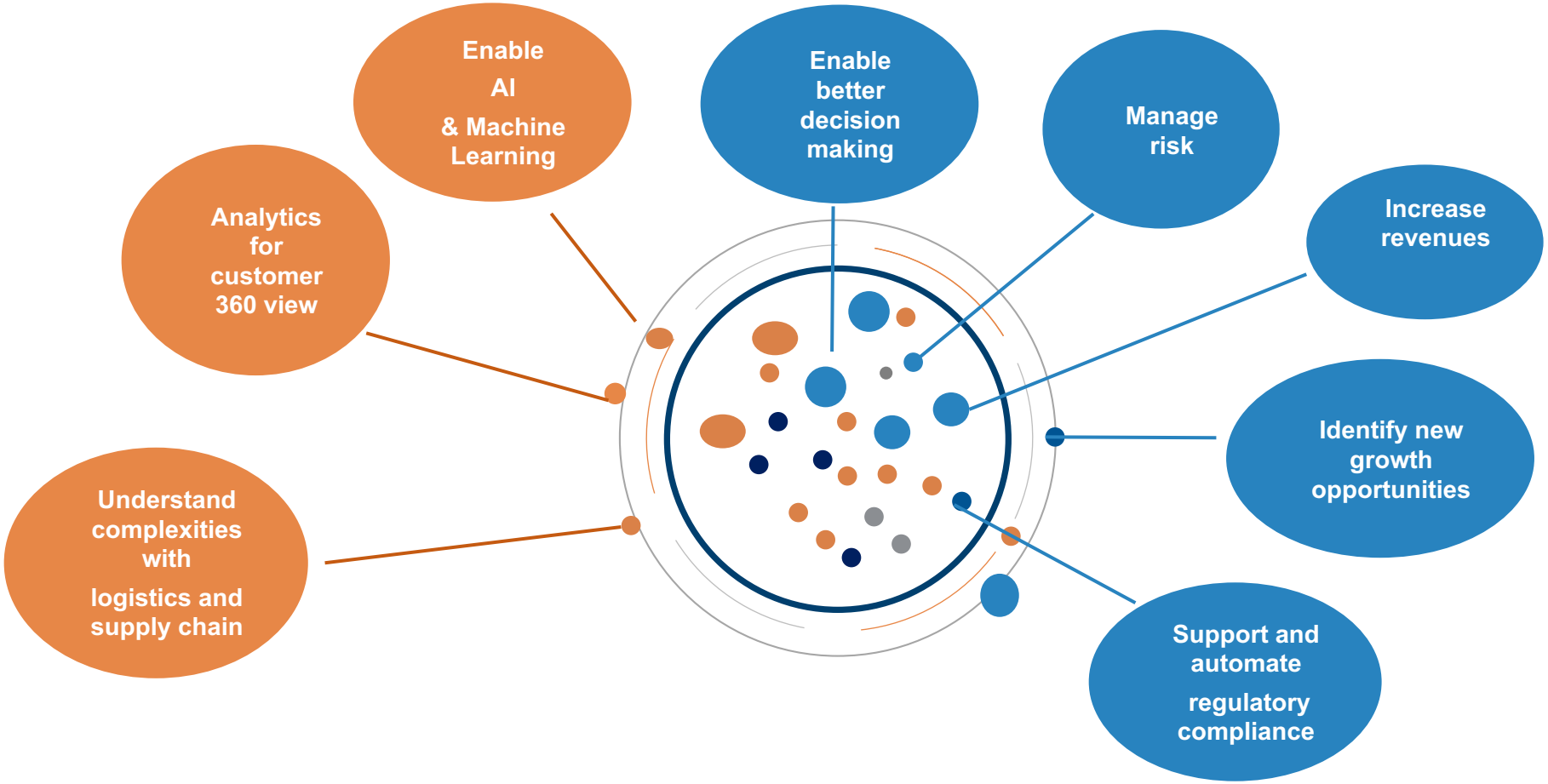
COMPLEX DATA ANALYSIS – NEEDS NEW DATA SOLUTIONS



- Business people are asking increasingly **complex questions** across structured and unstructured data.
- Complex analysis often requires **disributing and blending** data from multiple sources, multiple business units and increasingly expanding new data sources.
- **Analyzing this at scale** is not practical using traditonal relational data bases and query tools such as SQL.
- **Graph** data expresses your data into connected **relationships** across a network of **interconnected data points**.

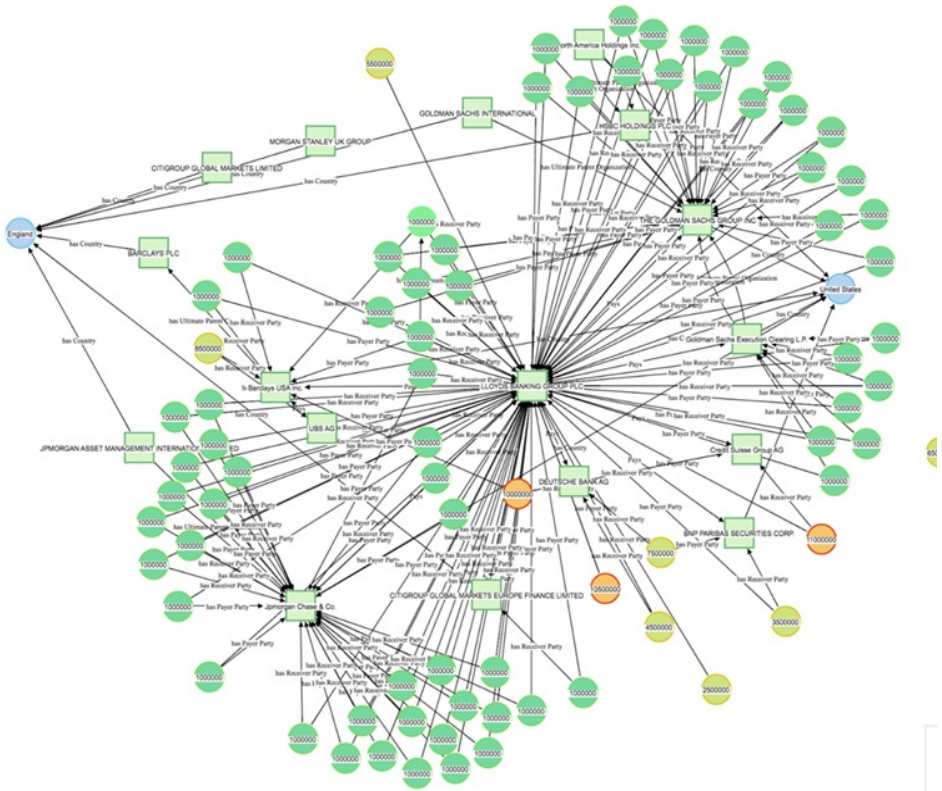
Source: Gartner.com

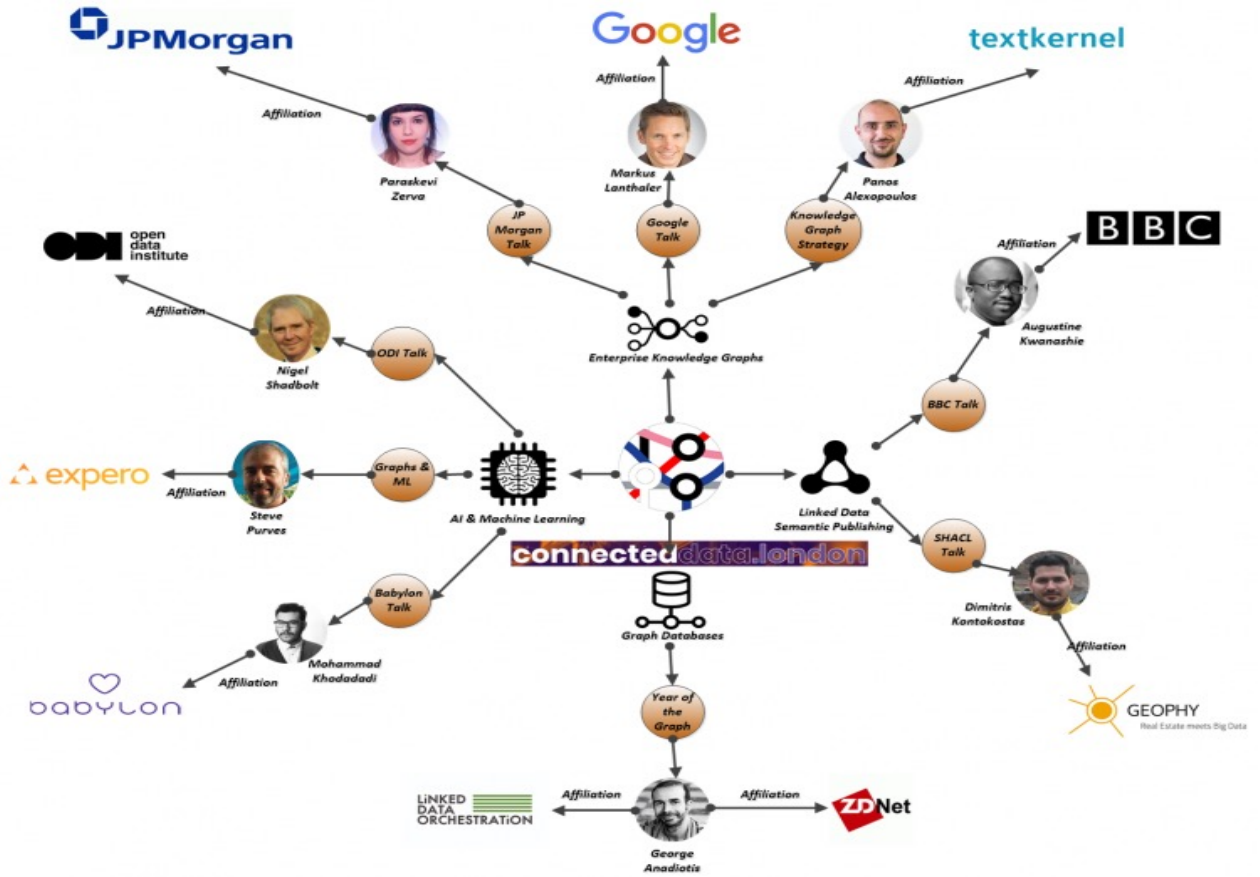
WHY KNOWLEDGE GRAPHS?



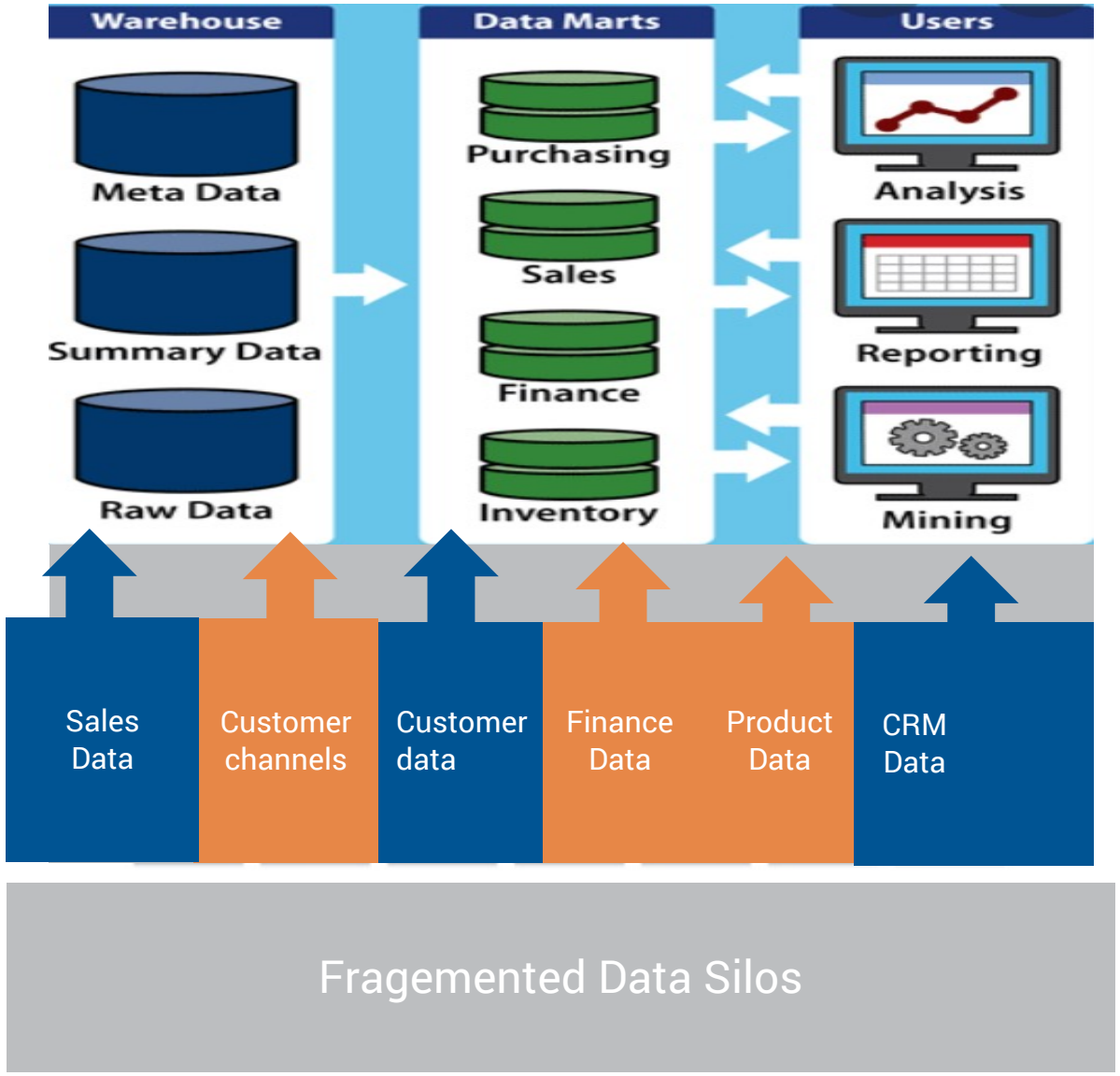
Gartner lists Graph powered augmented analytics among the top 10 trends, with 30% of the global organizations making Graph powered decisions by 2023 (Source: Gartner Top 10 Trends in Data and Analytics for 2020)

Graph View = Understand Patterns & Data Relationships





Silo Data Solutions = Fragmented Data Views & Poor Insights

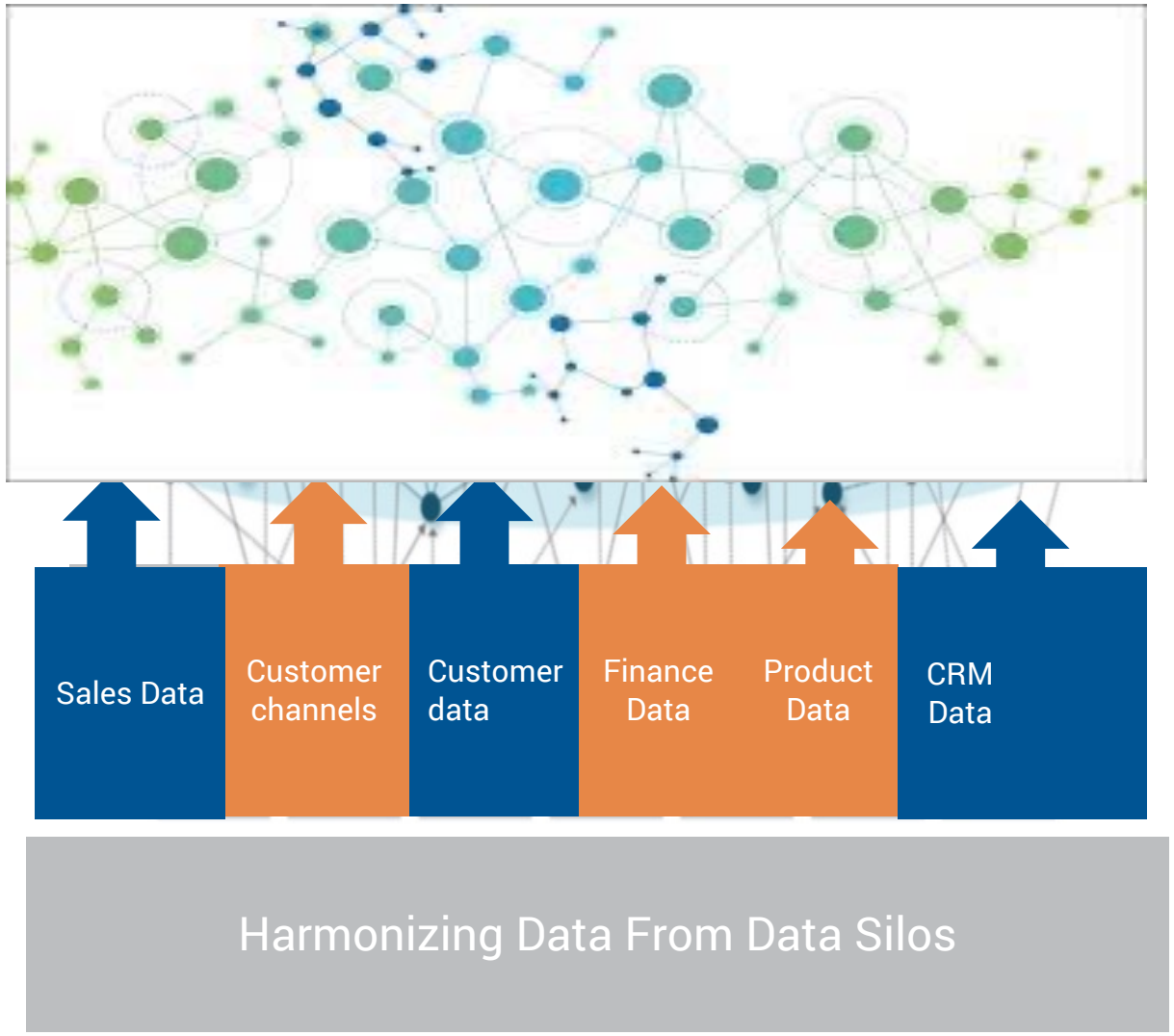


Poor Data Quality

Silo Data Management

Low Return on Analytics

GRAPH Solutions = Holistic Data Views & Enriched Data Insights



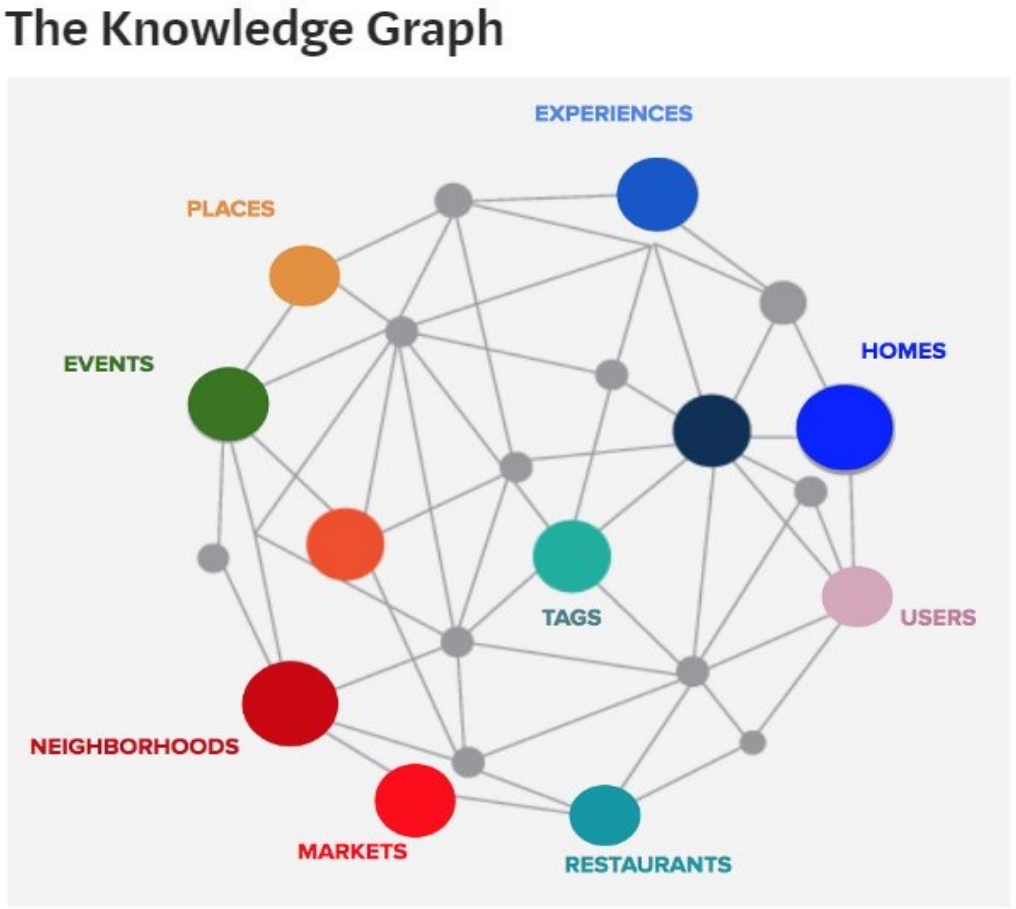
High Data Quality

Integrated Data Management

High Return on Analytics

Is Your Modern Data Platform Using Graph?

Graph Data Solutions Used By Leading Organisations Today



A visualization of the Knowledge Graph

KNOWLEDGE GRAPHS

- Don’t just watch….get involved!

EDM Council & Graph – Learn & Graph Innovation Lab

EDM COUNCIL MEMBERS

UBS · UnionBank · element22 · aws · AEGON · Bank of America · Deutsche Bank · RBC · BANK OF

Get involved: www.edmcouncil.org – Email: srolls@edmcouncil.org

ESG Workgroup

ESG Workgroup

What is ESG and why is it important?

- ESG – *Environmental, Social and corporate Governance* – **impacts ALL companies**
- ESG assets growing to **$53 trillion by 2025,** a THIRD of global Assets Under Management (AUM).*

ESG Data Challenges

Lack of consistent industry best practices, transparency, evolving standards and regulations, data availability, confusion across the ESG data supply chain

Goal

Cross-industry collaboration, provide data management best practices, educate data professionals, engage with regulators and standards organizations

Industry Benefits

- ✓ Coalesce the industry around a defined set of **Best practices and Benefits**
- ✓ Provide **ESG Data Transparency** to all stakeholders in the ESG supply chain

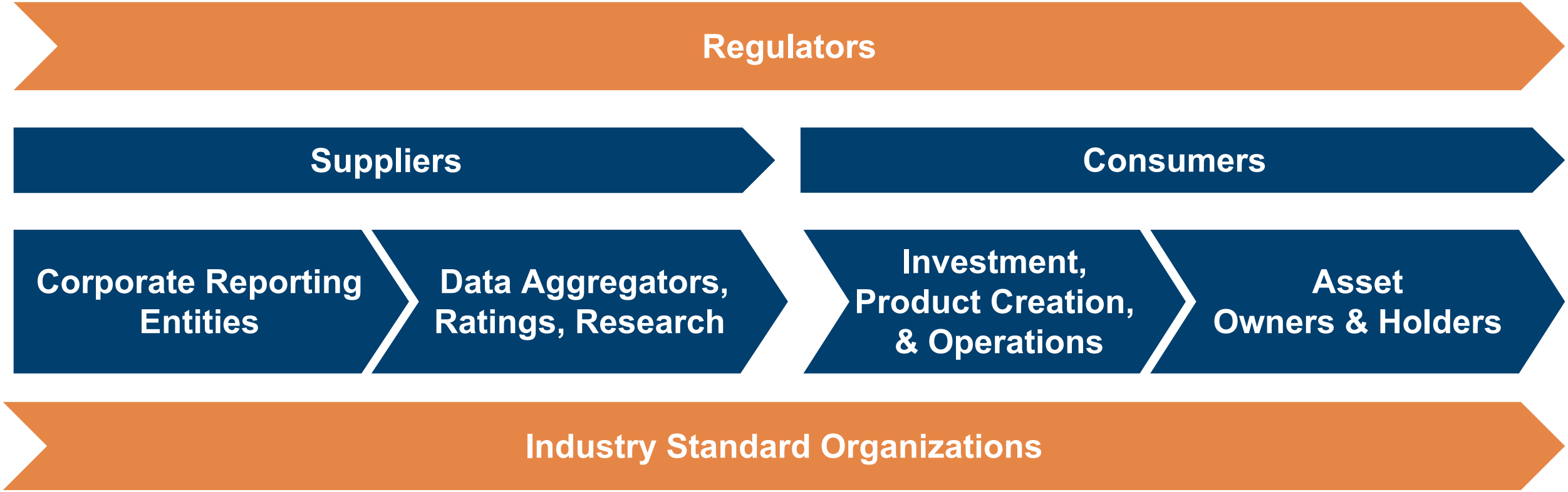
* Source Bloomberg

ESG Data Workgroup Participants

- 80+ Companies
- 120+ Professionals

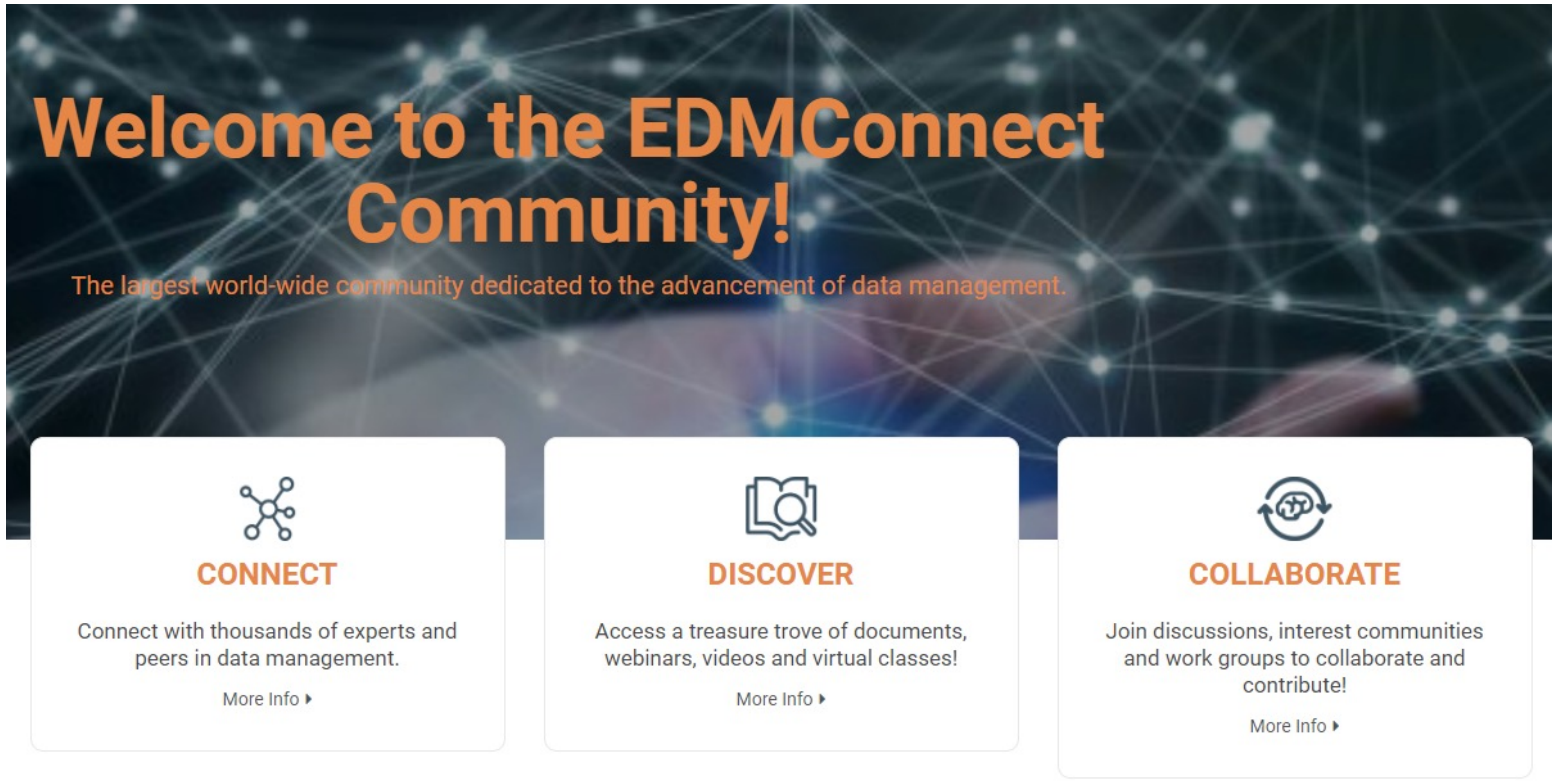
Get involved: www.edmcouncil.org – Go to Groups > ESG Data

ESG Data Supply Chain



EDMConnect – Connect & Collaborate

https://edmcouncil.org/



REGISTER NOW: edmcouncil.org/NewMemReg

EDMCouncil

Get involved with EDM Council!

LEARN MORE

- Visit edmcouncil.org – Go to “Groups”
- Contact us at info@edmcouncil.org