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

ANZ
Microsoft
AWS
Schneider Electric
Aditya Birla Group
AstraZeneca
Snowflake
Google
Verizon
TVS
NatWest Group
Credit Suisse
Tech Mahindra
S Group
Transurban
Michelin
Capgemini
GM
Brown Brothers Harriman
Paysafe
US Treasury
Loblaw
UnionBank of the Philippines
Grant Thornton
Oliver Wyman
Sberbank
Agricultural Bank of China
Direct Line Group
Avast
Santander
US Trust
Guidehouse
New York Life
Autodesk
Barclays
Wipro
Westpac
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

<table>
<thead>
<tr>
<th>Information Management Track Curricula</th>
<th>Certification Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM Foundations</td>
<td>NEW on eLearning! – DCAM (Data Management Capability Assessment Model)</td>
</tr>
<tr>
<td>Data Integration</td>
<td>Certified Information Management Professional (CIMP)</td>
</tr>
<tr>
<td>Data Quality</td>
<td>Certified Data Steward (CDS)</td>
</tr>
<tr>
<td>Master Data Management</td>
<td>Data Literacy (2021)</td>
</tr>
<tr>
<td>Data Governance</td>
<td></td>
</tr>
<tr>
<td>Data Modeling &amp; Metadata</td>
<td></td>
</tr>
<tr>
<td>Data Stewardship</td>
<td></td>
</tr>
<tr>
<td>Business Intelligence &amp; Analytics</td>
<td></td>
</tr>
<tr>
<td>Data Science</td>
<td></td>
</tr>
</tbody>
</table>

To register: EDMCouncil.org > Training > E-Learning
Cloud Data Management Capabilities (CDMC)
Cloud Data Management Capability (CDMC)

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

* Source: Flexera, 2020 State of the Cloud
CDMC: Industry Engagement

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

CDMC Workgroup

Cloud & Technology Provider Certification

Regulatory Engagement

Go-to-market Support

Cloud & Technology Provider Certification

EDM Council

- 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

CDMC Workgroup Publication on track for Q3 2021

Go-to-market Support

EDM Council

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

2H 2021 – 1H 2022 Other Industries

EDM Council

- Life Sciences
- Telecommunications
- Manufacturing
- Retail / Services
- Consumer Tech
- Government/Defense
- Others
Cloud Data Management Capabilities (CDMC) Key Controls

**Security Controls** must be enabled for sensitive data and evidence must be recorded.

- **Data Consumption Purpose** is required for all Data Sharing Agreements involving sensitive data.
- **Entitlements and Access for Sensitive Data** must be defaulted to creator and owner and all access must be tracked.
- **Classification** must be automated for all data.
- **Cataloguing** must be automated for all data.

**Data Cataloguing & Classification** must be populated for all data assets containing sensitive data.

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

**Data Quality Metrics** are required for all sensitive data.

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

**Data Lineage** information is required for all sensitive data.

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

**Classification** 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

---

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

**The Ownership field** in a data catalog must be populated for all sensitive data.
<table>
<thead>
<tr>
<th>Component</th>
<th>Capability</th>
<th>Sub-Capability</th>
<th>CDMC Automation Checklist (DRAFT)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Data Accountability &amp; Governance</strong></td>
<td>1.1 The Cloud Data Management business case is defined and measurable</td>
<td>1.1.1 TBA</td>
<td>Ownership field required for specified classifications (eg PI, MNPI, ISC=HR</td>
</tr>
<tr>
<td></td>
<td>1.2 Data ownership established for both migrated &amp; cloud-generated data</td>
<td>1.2.1 Data Owner roles and responsibilities are defined and agreed</td>
<td>Authorization field required on cataloged data assets for specified classifications (eg PI, MNPI, ISC=HR</td>
</tr>
<tr>
<td></td>
<td>1.3 Data sourcing and consumption are governed and supported by automation</td>
<td>1.3.1 Data sourcing is governed for accuracy, data duplication control and data protection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.4 Data Sovereignty, Residency and Localisation are actively managed</td>
<td>1.3.2 Data consumption is governed for accuracy, trustworthiness and usage</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.3.3 Data sources are authorized and certified</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.3.4 Data provisioning and consumption are standardised and automated</td>
<td></td>
</tr>
<tr>
<td><strong>2. Data Cataloguing &amp; Classification</strong></td>
<td>2.1 Data catalogues are implemented, used and interoperable</td>
<td>2.1.1 Data cataloguing is defined, scoped and actively used</td>
<td>Auto Cataloging Supported</td>
</tr>
<tr>
<td></td>
<td>2.2 Data classifications are defined and used</td>
<td>2.1.2 Metadata is discoverable, enriched, managed and exposed in Data Catalogues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.1.3 Data catalogues are interoperable across multi and hybrid cloud environments</td>
<td></td>
</tr>
<tr>
<td><strong>3. Data Accessibility &amp; Usage</strong></td>
<td>3.1 Data entitlements are managed, enforced and tracked</td>
<td>3.1.1 Data entitlement rights and obligations are captured as metadata</td>
<td>Entitlements automatically restricted to creator &amp; owner for sensitive data (ISC=HR</td>
</tr>
<tr>
<td></td>
<td>3.2 Ethical use and purpose of data use are tracked</td>
<td>3.1.2 Data entitlement rights are enforced</td>
<td>Access tracked for all sensitive data (ISC=HR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.1.3 Enforcement of data entitlement rights is evidenced</td>
<td>Purpose required for all DSAs for all PI &amp; sensitive data (ISC=HR</td>
</tr>
<tr>
<td><strong>4. Securing Data &amp; Privacy</strong></td>
<td>4.1 Data is secured and controls are evidenced</td>
<td>4.1.1 Data is encrypted at rest, in motion and while in use</td>
<td>Security controls and DLP are automatically enabled for all sensitive data (ISC=HR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1.2 Implementation of data security controls is evidenced</td>
<td>Security control evidence is sent to Catalog for all sensitive data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1.3 Data obfuscation techniques are defined, scoped and applied based on risk* and utility</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1.4 A Data Loss Prevention regime is in place</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.2 A data privacy framework is defined and operational</td>
<td>4.2.1 A data privacy framework is defined and agreed</td>
<td>Recommended data privacy framework workflows (eg jurisdictionally aware PIAs) are automatically triggered for all personal data</td>
</tr>
<tr>
<td><strong>5. Data Lifecycle Management</strong></td>
<td>5.1 Data quality is managed</td>
<td>5.1.1 Data Quality Rules Management</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.2 The data lifecycle is planned and managed</td>
<td>5.1.2 Data Quality Profiling and Validation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.1.3 Data Quality outputs publication</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.1.4 Data Quality issue management</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.2.1 Archiving and purging of data are managed</td>
<td></td>
</tr>
<tr>
<td><strong>6. Data &amp; Technical Architecture</strong></td>
<td>6.1 Data provenance and lineage are understood</td>
<td>6.1.1 TBA</td>
<td>Provenance information required for specified classifications (eg PI, MNPI, ISC=HR</td>
</tr>
<tr>
<td></td>
<td>6.2 Technical design principles are established and applied</td>
<td>6.1.2 Data back-up, resilience &amp; recovery - (Rewording TBA)</td>
<td>Well architected framework evaluated metrics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.1.3 Data back-up, resilience &amp; recovery - (Rewording TBA)</td>
<td>Costs directly associated with managing data assets are available to view by Data Owners in the catalog</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.2.1 Portability &amp; exit plan - (Rewording TBA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.2.2 Usage &amp; cost transparency - (Rewording TBA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.2.3 Usage &amp; cost transparency - (Rewording TBA)</td>
<td></td>
</tr>
</tbody>
</table>
CDMC: Accelerating Cloud Adoption

**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](http://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
EXPLOSION OF DATA VOLUME & COMPLEXITY

Source: Deloitte, 2017

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.

Total size of the digital universe
Data useful if analyzed
Data from mobile devices and people
Data from embedded systems

2013
4.4 ZB
22%
17%
2%

2020
44 ZB
37%
27%
10%

Up to 90% of this data is unstructured
Business people are asking increasingly complex questions across structured and unstructured data.

Complex analysis often requires distributing and blending data from multiple sources, multiple business units and increasingly expanding new data sources.

Analyzing this at scale is not practical using traditional 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
GRAPH Solutions = Holistic Data Views & Enriched Data Insights

Harmonizing Data From Data Silos

- Sales Data
- Customer channels
- Customer data
- Finance Data
- Product Data
- CRM Data

High Data Quality
Integrated Data Management
High Return on Analytics

High Return on Analytics
Is Your Modern Data Platform Using **Graph**?

Graph Data Solutions Used By Leading Organisations Today

The Knowledge Graph

A visualization of the Knowledge Graph
KNOWLEDGE GRAPHS

- Don’t just watch....get involved!
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

Regulators

Suppliers

Corporate Reporting Entities

Data Aggregators, Ratings, Research

Consumers

Investment, Product Creation, & Operations

Asset Owners & Holders

Industry Standard Organizations

EDM Council
EDMConnect – Connect & Collaborate

https://edmcouncil.org/

Welcome to the EDMConnect Community!
The largest world-wide community dedicated to the advancement of data management.

CONNECT
Connect with thousands of experts and peers in data management.
More Info ▶

DISCOVER
Access a treasure trove of documents, webinars, videos and virtual classes!
More Info ▶

COLLABORATE
Join discussions, interest communities and work groups to collaborate and contribute!
More Info ▶

REGISTER NOW: edmcouncil.org/NewMemReg
Get involved with EDM Council!

LEARN MORE

• Visit edmCouncil.org – Go to “Groups”
• Contact us at info@edmCouncil.org