EDM Webinar

Google Cloud:
Automating Cloud Data Governance

A conversation with

Mose Tronci
Industry Solution Architect
Google Cloud

Mark Tomlinson
Principal Architect
Google Cloud

Google Cloud

EDMCouncil

Today's panel

Moderator

Jim Halcomb
Head of Product Management
EDM Council

Mose Tronci
Industry Solution Architect
Google Cloud

Mark Tomlinson
Principal Architect
Google Cloud

Poll #1

EDMCouncil

What type of organisation do you work for?

- Enterprise
- Public sector
- Technology provider / Hyperscaler
- Consulting
- Other

Cloud Data Management Capabilities (CDMC)

CDMC Working Group

CDMC™ EDMCouncil Cloud Data Management Capabilities

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

EDMCouncil

* Source: Flexera, 2022 State of the Cloud

Cloud Data Management Capabilities (CDMC) Industry Engagement

CDMC EDMCouncil Cloud Data Management Capabilities

100+ Organizations – 300+ SME participants – Cross Industry Leading Organizations

CDMC Working Group

LSEG
Morgan Stanley
J.P.Morgan
citi
Goldman Sachs
Standard Chartered
UBS
HSBC
SOCIETE GENERALE
WELLS FARGO
Deutsche Bank
BARCLAYS
TP ICAP
LLOYDS BANK
NORTHERN TRUST
DTCC
Tradeweb
PayPal
TD Bank
BNP PARIBAS
Nasdaq
FreddieMac
We make home possible®
KPMG
CAPCO
EY

Cloud & Technology Providers

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

CDMC Working Group Delivered release v1.1
28 September 2021

Regulatory Engagement

EDMCouncil

- **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

EDMCouncil

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

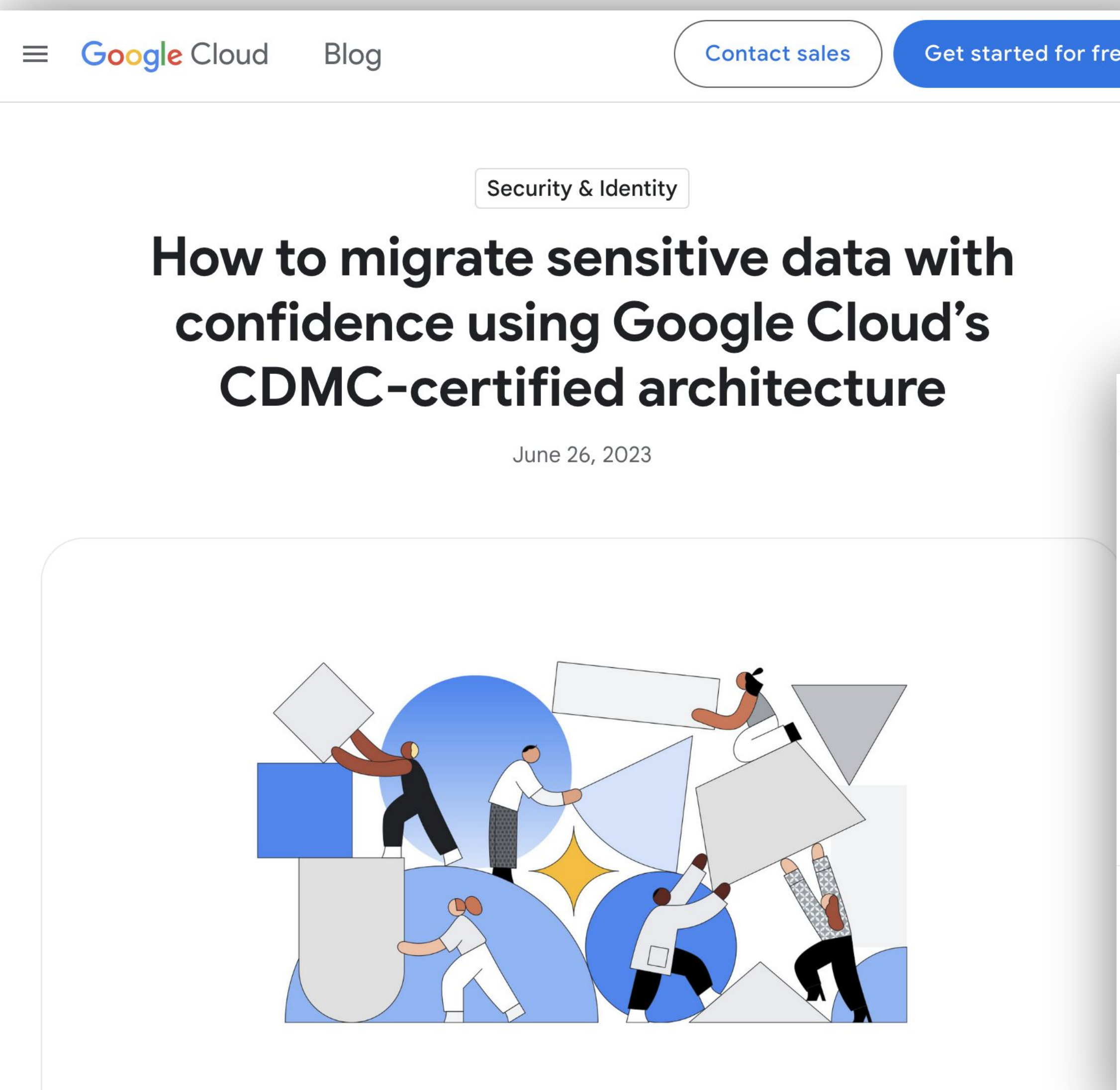
2022
CDMC Framework Validated Cross Industry

CDMC EDMCouncil

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

Google Cloud becomes the next CDMC Cloud Certified Solution

EDM Council CDMC™ CERTIFIED CLOUD SOLUTION



Security & Identity

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

June 26, 2023

Google Cloud

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 Dataplex, has been independently

Recent Announcements See all

Aldefi becomes the newest member to join the EDM Council

Google Cloud

Automating Cloud Data Governance

Mose Tronci, Industry Solution Architect

Mark Tomlinson, Principal Architect

EDM Council
CDMC™
CERTIFIED
CLOUD
SOLUTION

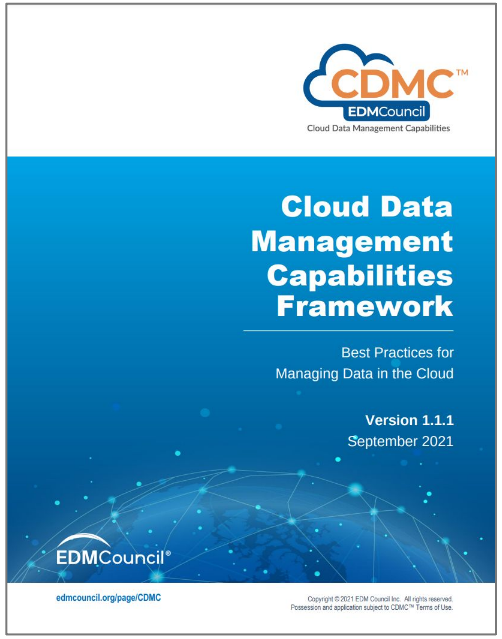
Poll #2

EDMCouncil

What is your role?

- Data Management professional
- CDO
- Data Engineer
- Cloud Architect
- Project Manager
- Consultant
- Sales
- Other

CDMC Main Elements Recap



CDMC™
EDMCouncil
Cloud Data Management Capabilities

Cloud Data Management Capabilities Framework

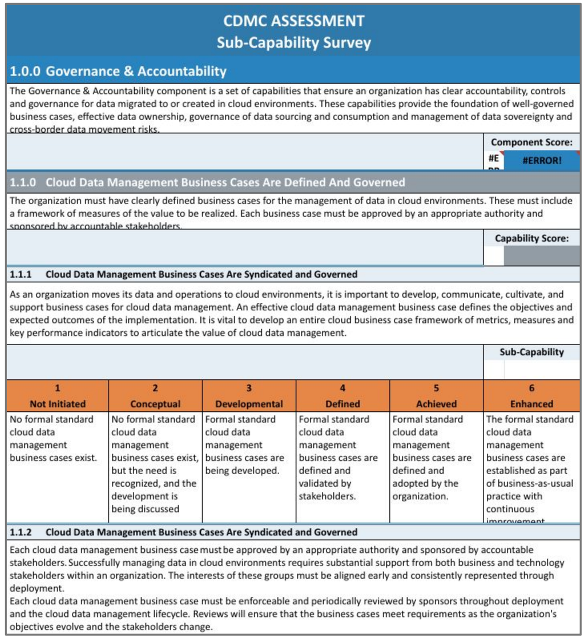
Best Practices for
Managing Data in the Cloud

Version 1.1.1
September 2021

EDMCouncil®

edmcouncil.org/page/CDMC

[CDMC 1.1.1 Framework](#)
Defines the 14 capabilities and 37 sub-capabilities across 6 domains.



CDMC ASSESSMENT
Sub-Capability Survey

1.0.0 Governance & Accountability

The Governance & Accountability component is a set of capabilities that ensure an organization has clear accountability, controls and governance for data migrated to or created in cloud environments. These capabilities provide the foundation of well-governed business cases, effective data ownership, governance of data sourcing and consumption and management of data sovereignty and cross-border data movement risks.

Component Score:
#E #ERROR!

1.1.0 Cloud Data Management Business Cases Are Defined And Governed

The organization must have clearly defined business cases for the management of data in cloud environments. These must include a framework of measures of the value to be realized. Each business case must be approved by an appropriate authority and sponsored by accountable stakeholders.

Capability Score:

1.1.1 Cloud Data Management Business Cases Are Syndicated and Governed

As an organization moves its data and operations to cloud environments, it is important to develop, communicate, cultivate, and support business cases for cloud data management. An effective cloud data management business case defines the objectives and expected outcomes of the implementation. It is vital to develop an entire cloud business case framework of metrics, measures and key performance indicators to articulate the value of cloud data management.

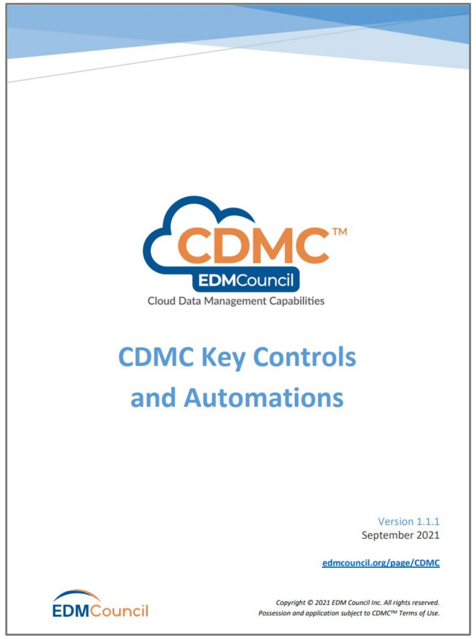
Sub-Capability

1 Not Initiated	2 Conceptual	3 Developmental	4 Defined	5 Achieved	6 Enhanced
No formal standard cloud data management business cases exist.	No formal standard cloud data management business cases exist, but the need is recognized, and the development is being discussed	Formal standard cloud data management business cases are being developed.	Formal standard cloud data management business cases are defined and validated by stakeholders.	Formal standard cloud data management business cases are defined and adopted by the organization.	The formal standard cloud data management business cases are established as part of business-as-usual practice with continuous improvement

1.1.2 Cloud Data Management Business Cases Are Syndicated and Governed

Each cloud data management business case must be approved by an appropriate authority and sponsored by accountable stakeholders. Successfully managing data in cloud environments requires substantial support from both business and technology stakeholders within an organization. The interests of these groups must be aligned early and consistently represented through deployment.
Each cloud data management business case must be enforceable and periodically reviewed by sponsors throughout deployment and the cloud data management lifecycle. Reviews will ensure that the business cases meet requirements as the organization's objectives evolve and the stakeholders change.

[CDMC 1.1.1 Assessment Template](#)
Spreadsheet to facilitate sub-capability assessments using scoring guide of 1 to 6.



CDMC™
EDMCouncil
Cloud Data Management Capabilities

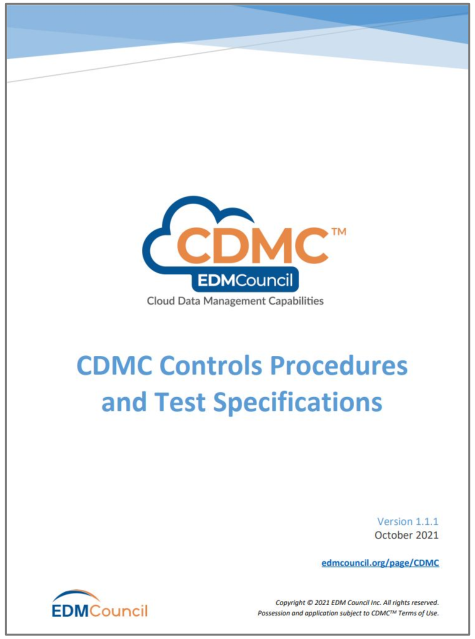
CDMC Key Controls and Automations

Version 1.1.1
September 2021

edmcouncil.org/page/CDMC

EDMCouncil

[CDMC 1.1.1 Key Controls](#)
Specifies 14 key controls across the 6 framework domains.



CDMC™
EDMCouncil
Cloud Data Management Capabilities

CDMC Controls Procedures and Test Specifications

Version 1.1.1
October 2021

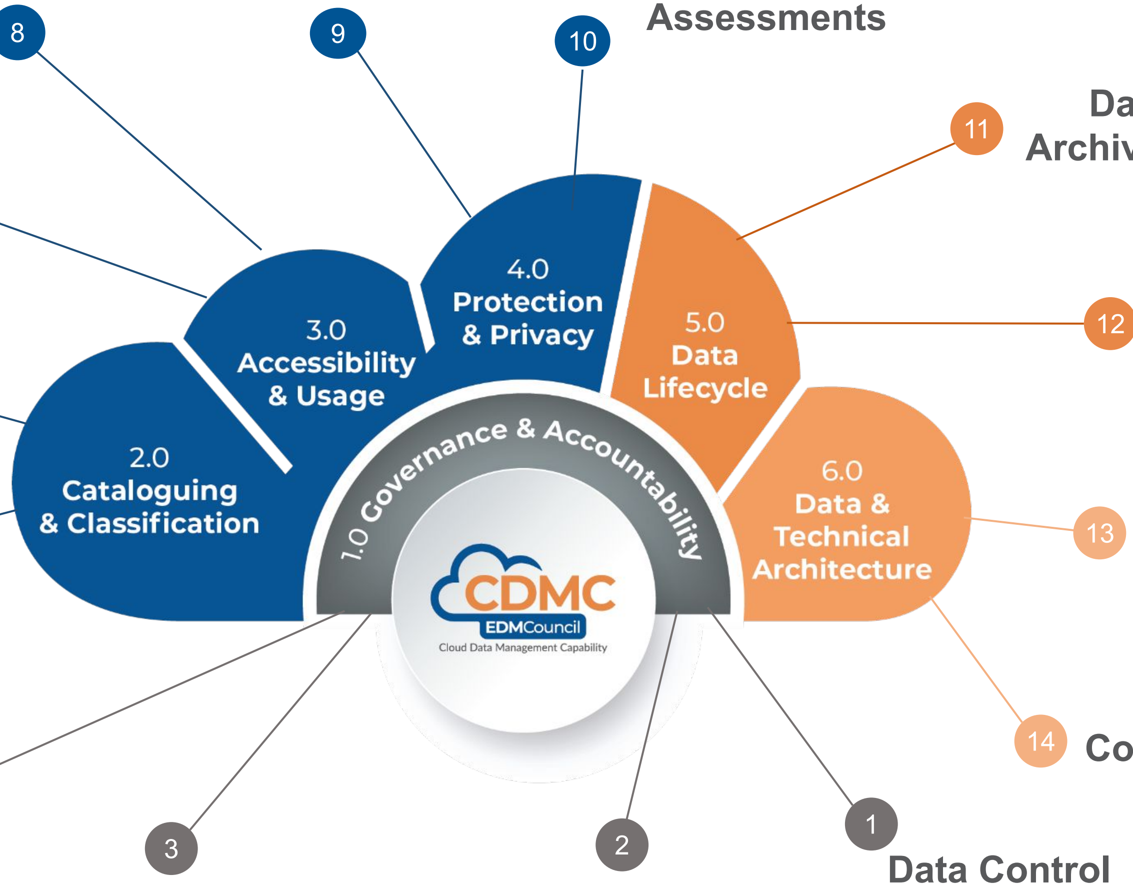
edmcouncil.org/page/CDMC

EDMCouncil

[CDMC 1.1.1 Controls Procedures and Test Specifications](#)
Outlines tests and evidence required for third-party CDMC platform/solution certification

Cloud Data Management Capabilities

14 Key Controls for Managing Data Risk



1.0 Governance & Accountability
2.0 Cataloguing & Classification
3.0 Accessibility & Usage
4.0 Protection & Privacy
5.0 Data Lifecycle
6.0 Data & Technical Architecture
CDMC
EDMCouncil
Cloud Data Management Capability
8
9
10
7
11
6
12
5
13
4
14
3
2
1

1. Data Control Compliance
2. Data Ownership
3. Authoritative Data Sources and Provisioning Points
4. *Data Sovereignty and Cross-Border Movement of sensitive data*
5. Data Cataloguing
6. Data Classification
7. Entitlements and Access for Sensitive Data
8. Data Consumption Purpose
9. Security Controls
10. Data Privacy Impact Assessments
11. Data Retention, Archiving and Purging
12. Data Quality Measurement
13. Data Lineage
14. Cost Metrics

CDMC Assessment types

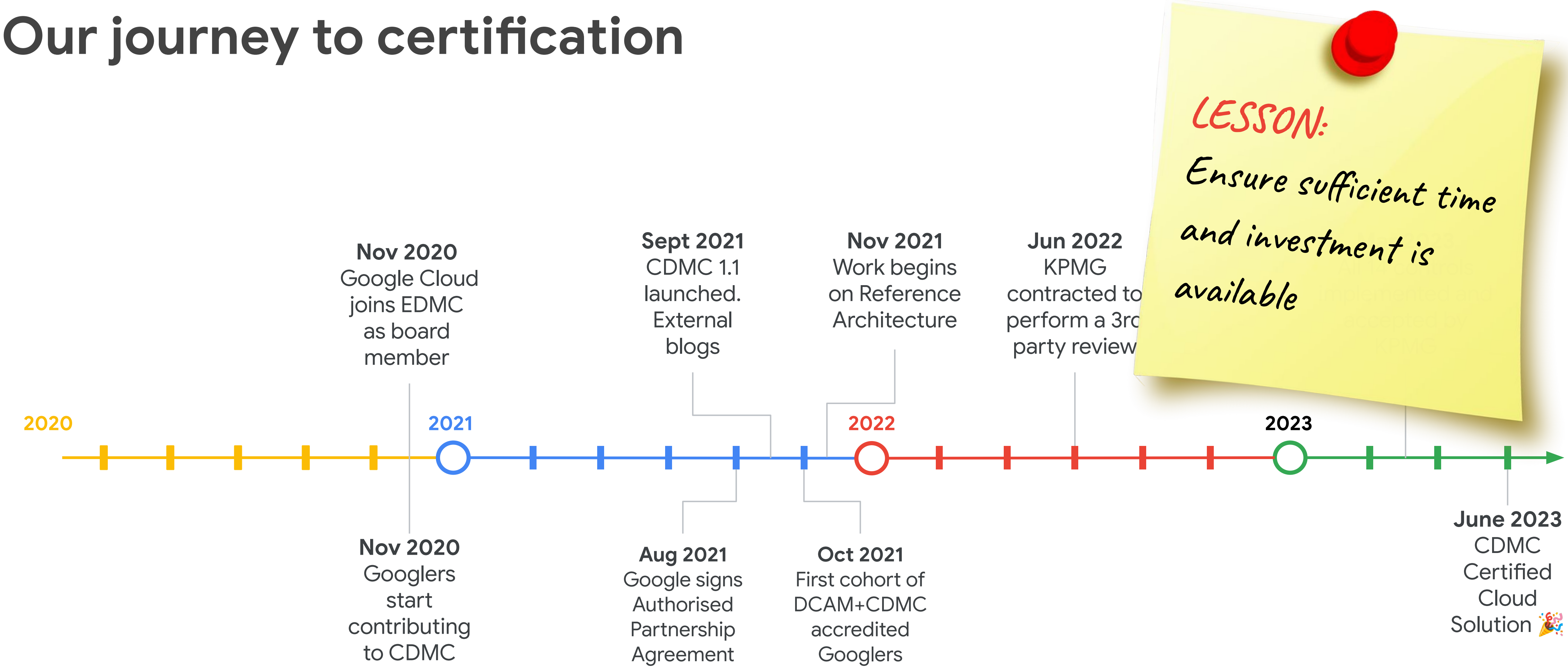
Assessment Type	Target	Use For	Basis	Uses CDMC Capability Framework?	Uses CDMC Key Controls Framework?	Uses CDMC Key Controls Test Procedures?	Assessed By
CDMC Certification	Individual	Personal skills, Planning	Training Attendance	Yes	No	No	Online Exam
CDMC Readiness Assessment	Consumer	Planning and Progress Tracking	Sentiment or Evidence-based	Yes	Optional	Optional	Self- or CDMC Authorized partner
CDMC Certification Assessment	Consumer	Formal Audit, Regulatory Preparation	Evidence-based	Yes	Yes	Yes	CDMC Authorized partner
CDMC Certification Assessment	Cloud Providers	Planning and Independent Evidence	Evidence-based	No	Yes	Yes	Independent CDMC Authorized partner

Focus today

Relevant Lessons

Key Differences

Our journey to certification



Nov 2020
Google Cloud joins EDMC as board member

Nov 2020
Googlers start contributing to CDMC

Aug 2021
Google signs Authorised Partnership Agreement

Sept 2021
CDMC 1.1 launched. External blogs

Oct 2021
First cohort of DCAM+CDMC accredited Googlers

Nov 2021
Work begins on Reference Architecture

Jun 2022
KPMG contracted to perform a 3rd party review

June 2023
CDMC Certified Cloud Solution 🎉

LESSON:
Ensure sufficient time and investment is available

Introducing the Google CDMC Team

LESSON:

Check your team has the right composition of skills

Executive Sponsors

Prajakta Damle
Director, Prod Mgmt
(Data Gov)
EDMC Board

Andy Chang
Group Prod Mgr
(Security)
Blueprint Sponsor

CDMC GCP Reference Architecture Team

Mose Tronci
Industry Solns Architect
EDMC Lead

Mark Tomlinson
Principal Architect
Project Lead

Eduardo Marreto
Cloud Data Engineer

Erlander Lo
Senior Product Manager, Security
Blueprint Lead

John Swain
Customer Engineer, Smart Analytics

Shirley Cohen
Solutions Architect
Tag Engine Lead

Svitlana Gavrylova
Data Management Practice Lead UK

CDMC 1.1 Specification Working Group Contributors

Mose Tronci
Industry Solns Architect
EDMC Lead

Andrew Ikonnikov
Senior Product Manager, Data Governance

Kishore Gopalan
Principal Architect

Prajakta Pandharkar
Customer Engineer, ML

Rochak Lamba
Innovation Lead, Financial Services, PSO

Svitlana Gavrylova
Data Management Practice Lead UK

Thinh Ha
Cloud Data Engineer, PSO

Google Cloud CDMC Reference Architecture core principles

#	Principle
1	Build on prior work and maximize reuse
2	Implement all 14 CDMC controls
3	Align to product roadmap, but minimize dependencies on future items
4	Prefer native Google Cloud features over third party services
5	Strong preference for full automation, but pragmatic, iterative approach
6	Author with intent to publish externally for customers and partners
7	Provide value-add for both new and established Google Cloud customers

LESSON:
Consider bespoke vs.
reusable
implementation

LESSON:
Consider starting
with limited
automation and
iterate

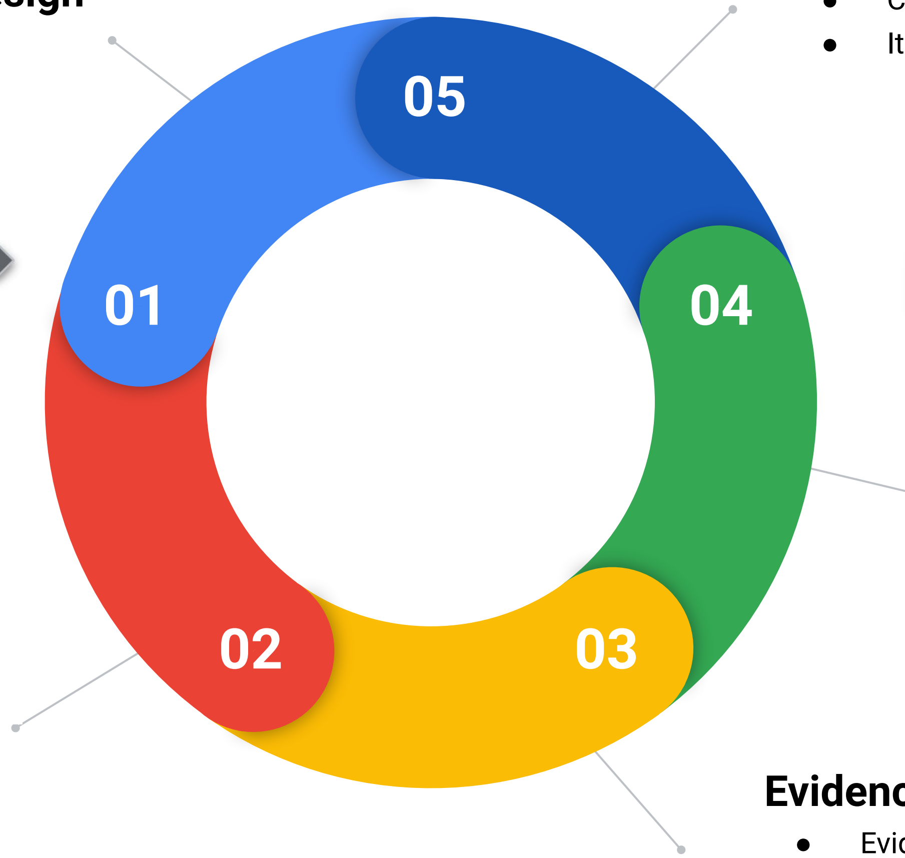
CDMC
EDMCouncil

Delivery process overview

Analysis & Design

- Draft
- Review
- Rework

C01: Data Control Compliance
C02: Ownership
C03: Data Sources
C04: Data Sovereignty
C05: Cataloguing
C06: Classification
C07: Entitlements
C08: Data Consumption
C09: Security
C10: Data Protection
C11: Retention/Archiving
C12: Data Quality
C13: Cost Metrics
C14: Data Lineage



Feedbac

- Capt
- Itera

LESSON:
Identify sponsoring clients early and engage throughout

Documentation & Publication

- Github docs
- Architecture whitepaper

Implementation

- Control detector
- Tag Engine

Evidence & Assessment

- Evidence doc
- Review
- Rework

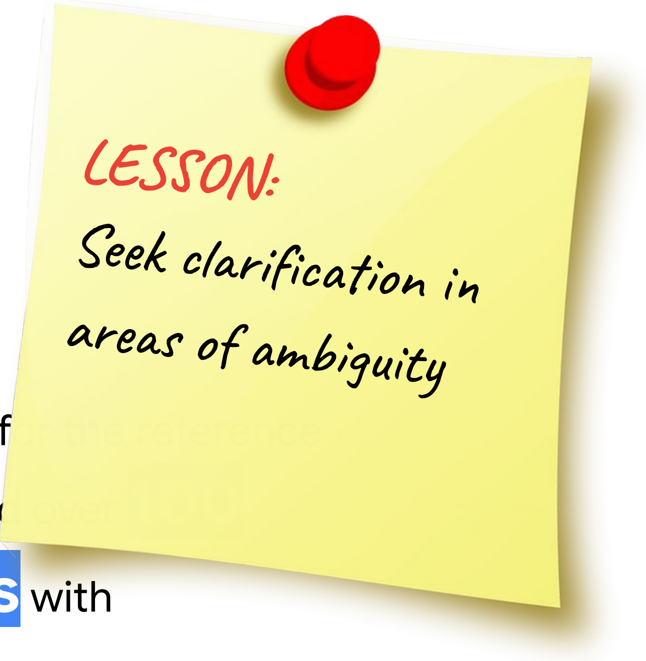
Poll #3

EDMCouncil

Where are you on your CDMC journey?

- Not started
- Currently investigating
- Project in progress
- Already assessed
- Already certified

Analysis & Design - key stats



LESSON:

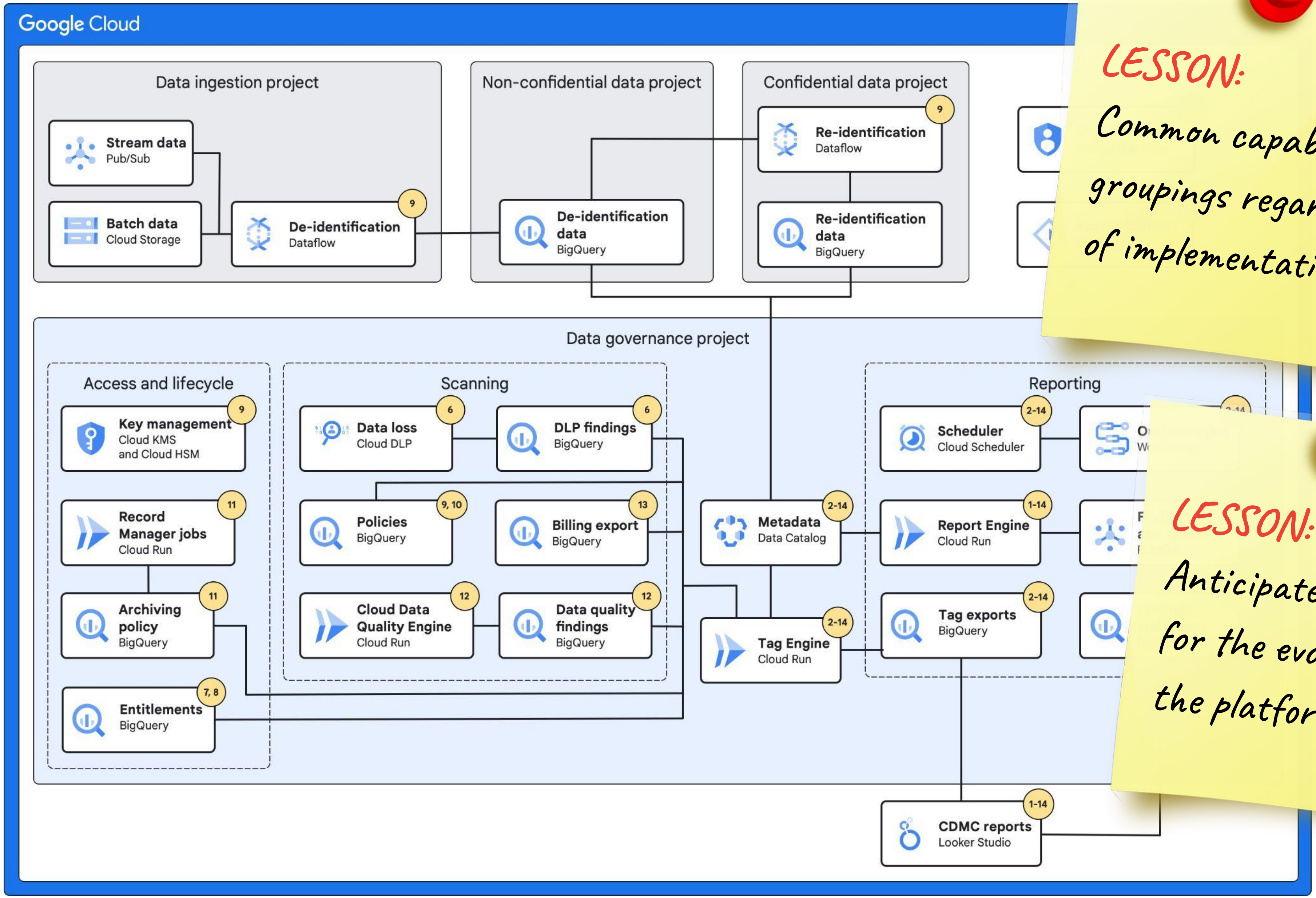
Seek clarification in areas of ambiguity

The CDMC framework specification extends to **165 pages** and details **37 sub-capabilities** of data management and governance

The internal design for the reference architecture had over [illegible] **pages** with **60 figures**,and **136 design decisions**

No perfect answer - need to put a stake in the ground

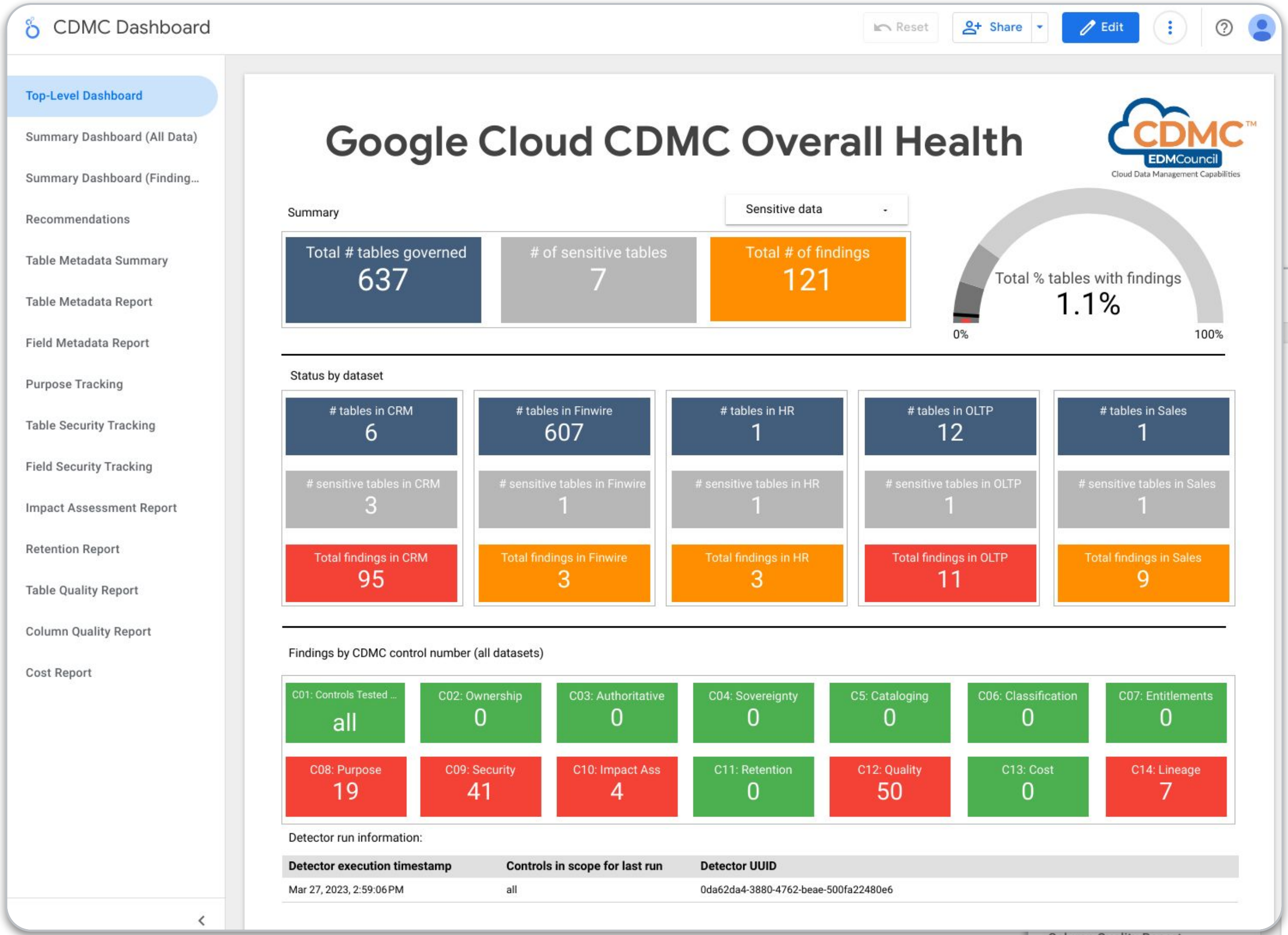
Implementation: Overview and Control Map



LESSON:
Common capability groupings regardless of implementation

LESSON:
Anticipate and budget for the evolution of the platform

All findings in one place



Google Cloud CDMC Overall Health

Summary

Total # tables governed	# of sensitive tables	Total # of findings
637	7	121

Total % tables with findings 1.1%

Status by dataset

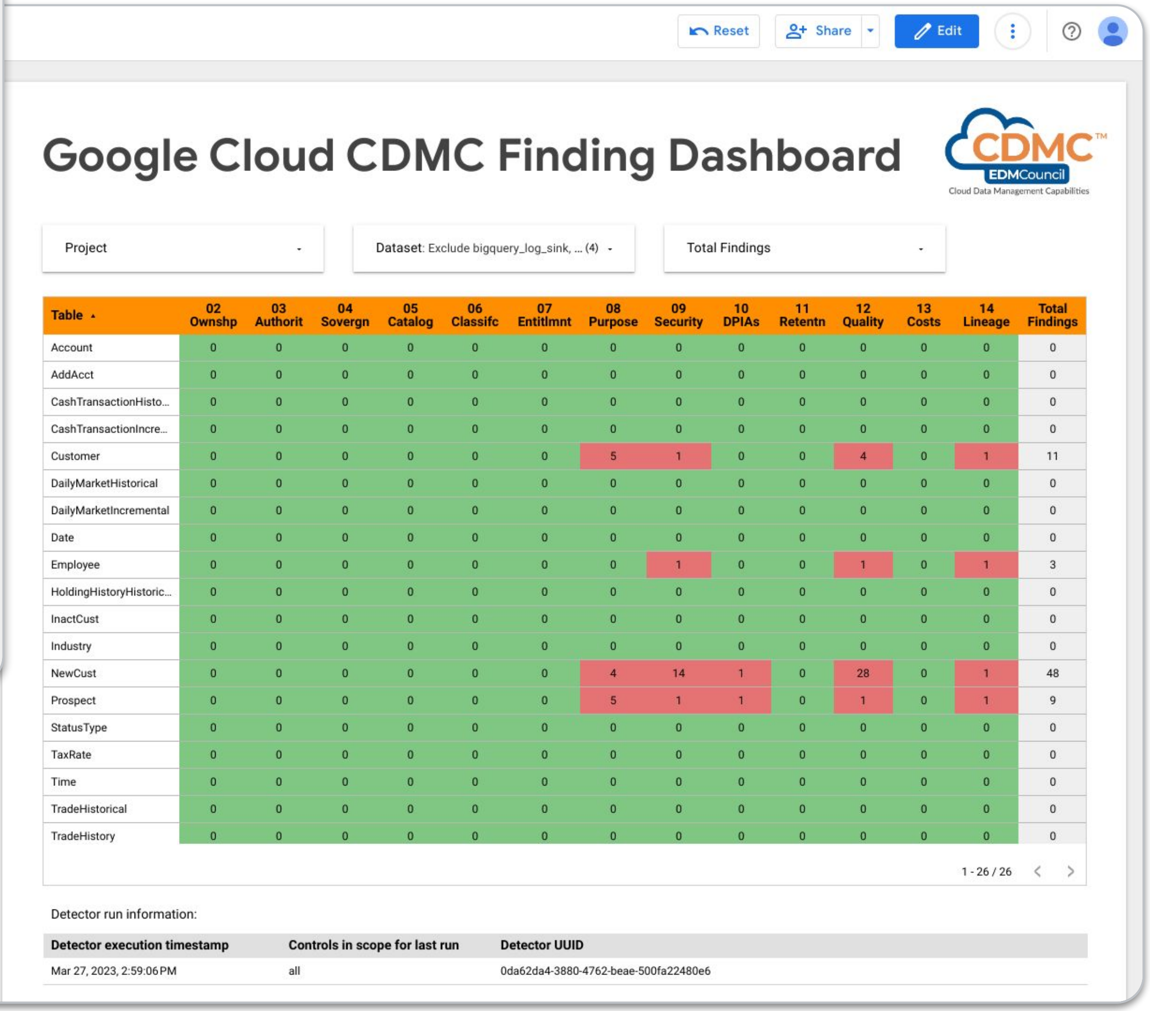
# tables in CRM	# tables in Finwire	# tables in HR	# tables in OLTP	# tables in Sales
6	607	1	12	1
# sensitive tables in CRM	# sensitive tables in Finwire	# sensitive tables in HR	# sensitive tables in OLTP	# sensitive tables in Sales
3	1	1	1	1
Total findings in CRM	Total findings in Finwire	Total findings in HR	Total findings in OLTP	Total findings in Sales
95	3	3	11	9

Findings by CDMC control number (all datasets)

C01: Controls Tested ...	C02: Ownership	C03: Authoritative	C04: Sovereignty	C5: Cataloging	C06: Classification	C07: Entitlements
all	0	0	0	0	0	0
C08: Purpose	C09: Security	C10: Impact Ass	C11: Retention	C12: Quality	C13: Cost	C14: Lineage
19	41	4	0	50	0	7

Detector run information:

Detector execution timestamp	Controls in scope for last run	Detector UUID
Mar 27, 2023, 2:59:06 PM	all	0da62da4-3880-4762-beae-500fa22480e6



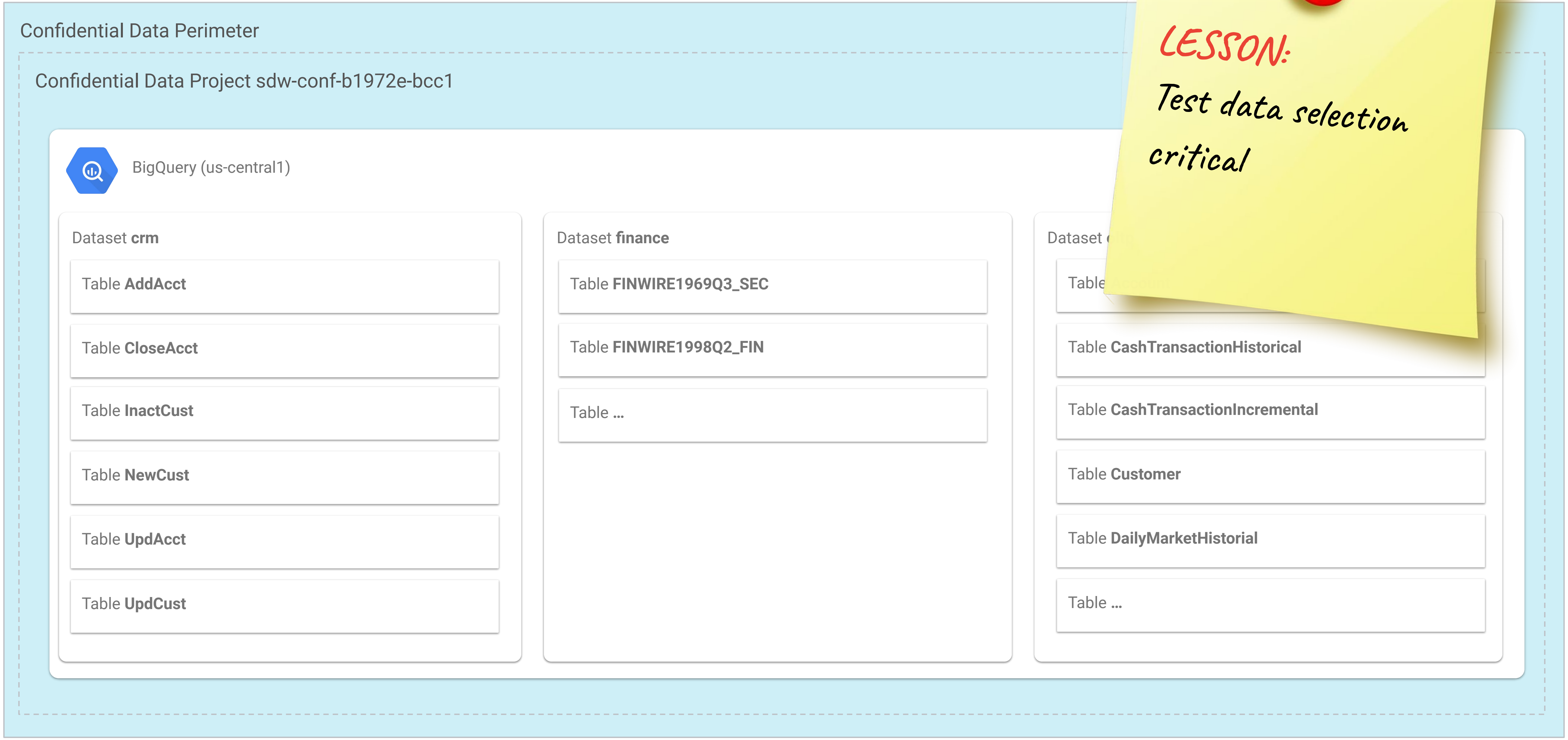
Google Cloud CDMC Finding Dashboard

Table ▲	02 Ownshp	03 Authorit	04 Sovergn	05 Catalog	06 Classifc	07 Entitlmnt	08 Purpose	09 Security	10 DPIAs	11 Retentn	12 Quality	13 Costs	14 Lineage	Total Findings
Account	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AddAcct	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CashTransactionHisto...	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CashTransactionIncre...	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Customer	0	0	0	0	0	0	5	1	0	0	4	0	1	11
DailyMarketHistorical	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DailyMarketIncremental	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Date	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Employee	0	0	0	0	0	0	0	1	0	0	1	0	1	3
HoldingHistoryHistoric...	0	0	0	0	0	0	0	0	0	0	0	0	0	0
InactCust	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Industry	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NewCust	0	0	0	0	0	0	4	14	1	0	28	0	1	48
Prospect	0	0	0	0	0	0	5	1	1	0	1	0	1	9
StatusType	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TaxRate	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Time	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TradeHistorical	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TradeHistory	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Detector run information:

Detector execution timestamp	Controls in scope for last run	Detector UUID
Mar 27, 2023, 2:59:06 PM	all	0da62da4-3880-4762-beae-500fa22480e6

Implementation: Test Datasets, subset of TPC-DI



Confidential Data Perimeter

Confidential Data Project sdw-conf-b1972e-bcc1

BigQuery (us-central1)

Dataset **crm**

- Table **AddAcct**
- Table **CloseAcct**
- Table **InactCust**
- Table **NewCust**
- Table **UpdAcct**
- Table **UpdCust**

Dataset **finance**

- Table **FINWIRE1969Q3_SEC**
- Table **FINWIRE1998Q2_FIN**
- Table **...**

Dataset

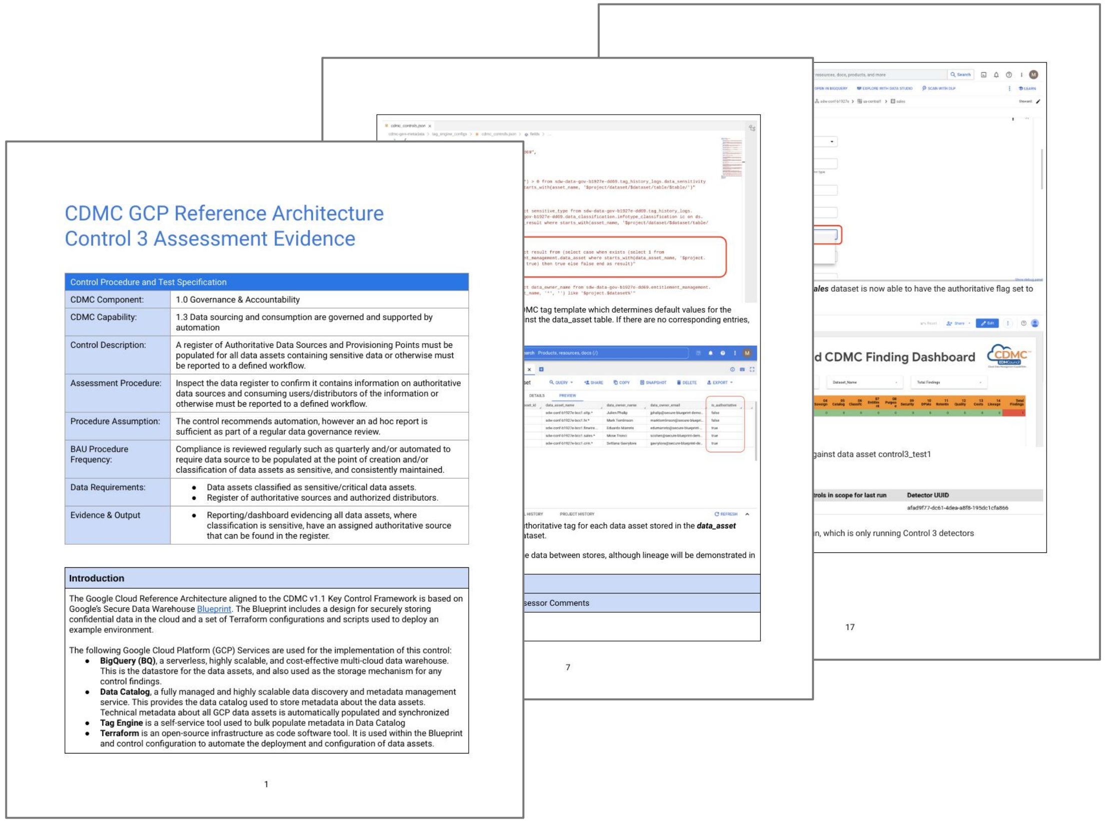
- Table
- Table **CashTransactionHistorical**
- Table **CashTransactionIncremental**
- Table **Customer**
- Table **DailyMarketHistorial**
- Table **...**

LESSON:

Test data selection critical

Our TPC-DI-derived test data provides a large selection of data classifications and meaningful volumes to demonstrate our capabilities.

KPMG Assessment process



CDMC GCP Reference Architecture Control 3 Assessment Evidence

Control Procedure and Test Specification	
CDMC Component:	1.0 Governance & Accountability
CDMC Capability:	1.3 Data sourcing and consumption are governed and supported by automation
Control Description:	A register of Authoritative Data Sources and Provisioning Points must be populated for all data assets containing sensitive data or otherwise must be reported to a defined workflow.
Assessment Procedure:	Inspect the data register to confirm it contains information on authoritative data sources and consuming users/distributors of the information or otherwise must be reported to a defined workflow.
Procedure Assumption:	The control recommends automation, however an ad hoc report is sufficient as part of a regular data governance review.
BAU Procedure Frequency:	Compliance is reviewed regularly such as quarterly and/or automated to require data source to be populated at the point of creation and/or classification of data assets as sensitive, and consistently maintained.
Data Requirements:	• Data assets classified as sensitive/critical data assets. • Register of authoritative sources and authorized distributors.
Evidence & Output	• Reporting/dashboard evidencing all data assets, where classification is sensitive, have an assigned authoritative source that can be found in the register.

Introduction

The Google Cloud Reference Architecture aligned to the CDMC v1.1 Key Control Framework is based on Google's Secure Data Warehouse Blueprint. The Blueprint includes a design for securely storing confidential data in the cloud and a set of Terraform configurations and scripts used to deploy an example environment.

The following Google Cloud Platform (GCP) Services are used for the implementation of this control:

- **BigQuery (BQ)**, a serverless, highly scalable, and cost-effective multi-cloud data warehouse. This is the datastore for the data assets, and also used as the storage mechanism for any control findings.
- **Data Catalog**, a fully managed and highly scalable data discovery and metadata management service. This provides the data catalog used to store metadata about the data assets. Technical metadata about all GCP data assets is automatically populated and synchronized
- **Tag Engine** is a self-service tool used to bulk populate metadata in Data Catalog
- **Terraform** is an open-source infrastructure as code software tool. It is used within the Blueprint and control configuration to automate the deployment and configuration of data assets.

1

MC tag template which determines default values for the
nst the data_asset table. If there are no corresponding entries,
thoritative tag for each data asset stored in the data_asset
taset.
e data between stores, although lineage will be demonstrated in
ssessor Comments
7
ales dataset is now able to have the authoritative flag set to
d CDMC Finding Dashboard
CDMC
gainst data asset control3_test1
trols in scope for last run
Detector UUID
n, which is only running Control 3 detectors
17

GCP CDMC RA Evidence Docs

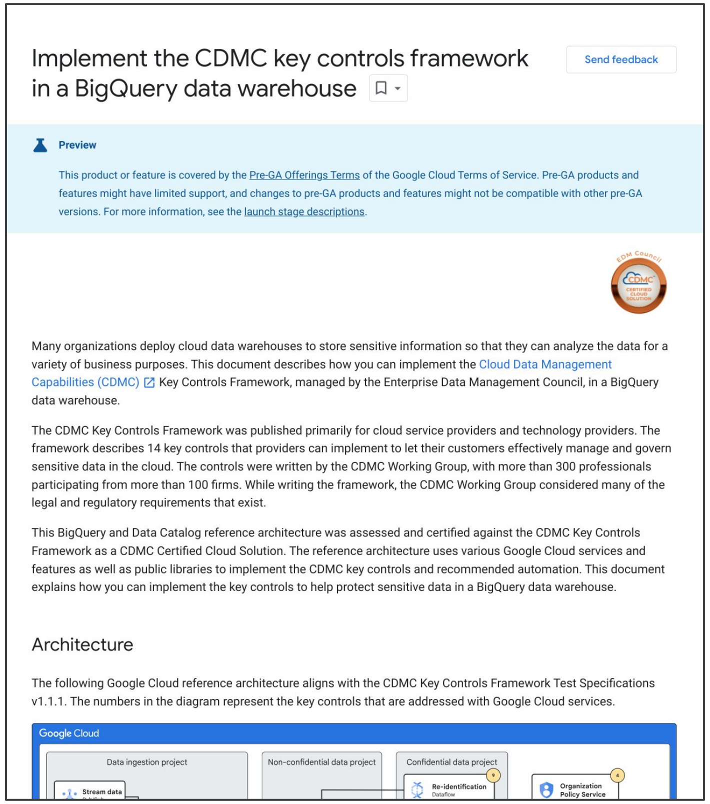
Evidence documents for each Control validating alignment to criteria in test specification (300+ pages)

#	Description		s
1	Data Control Compliance		
2	Ownership Field		
3	Authoritative Data Sources & Provisioning		
4	Data Sovereignty and Cross-Border Mov		
5	Cataloging		
6	Classification	Implemented	Accepted
7	Entitlements and Access for Sensitive Data	Implemented	Accepted
8	Data Consumption Purpose	Implemented	Accepted
9	Security Controls	Implemented	Accepted
10	Data Protection Impact Assessments	Implemented	Accepted
11	Data Retention, Archiving and Purging	Implemented	Accepted
12	Data Quality Measurement	Implemented	Accepted
13	Cost Metrics	Implemented	Accepted
14	Data Lineage	Implemented	Accepted

LESSON:

Validate the minimum evidence required to satisfy the assessor

Documentation & Publication:
Our primary technical deliverables



Implement the CDMC key controls framework in a BigQuery data warehouse

Send feedback

Preview

This product or feature is covered by the Pre-GA Offerings Terms of the Google Cloud Terms of Service. Pre-GA products and features might have limited support, and changes to pre-GA products and features might not be compatible with other pre-GA versions. For more information, see the launch stage descriptions.

Many organizations deploy cloud data warehouses to store sensitive information so that they can analyze the data for a variety of business purposes. This document describes how you can implement the Cloud Data Management Capabilities (CDMC) Key Controls Framework, managed by the Enterprise Data Management Council, in a BigQuery data warehouse.

The CDMC Key Controls Framework was published primarily for cloud service providers and technology providers. The framework describes 14 key controls that providers can implement to let their customers effectively manage and govern sensitive data in the cloud. The controls were written by the CDMC Working Group, with more than 300 professionals participating from more than 100 firms. While writing the framework, the CDMC Working Group considered many of the legal and regulatory requirements that exist.

This BigQuery and Data Catalog reference architecture was assessed and certified against the CDMC Key Controls Framework as a CDMC Certified Cloud Solution. The reference architecture uses various Google Cloud services and features as well as public libraries to implement the CDMC key controls and recommended automation. This document explains how you can implement the key controls to help protect sensitive data in a BigQuery data warehouse.

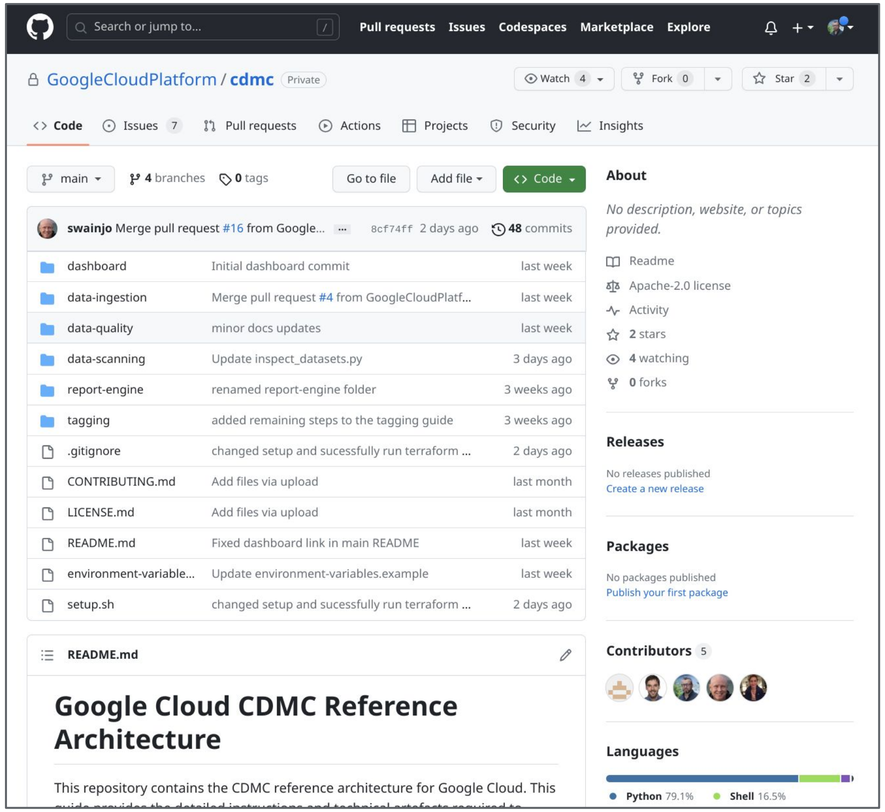
Architecture

The following Google Cloud reference architecture aligns with the CDMC Key Controls Framework Test Specifications v1.1.1. The numbers in the diagram represent the key controls that are addressed with Google Cloud services.

Google Cloud
Data ingestion project
Stream data
Non-confidential data project
Confidential data project
9
Re-identification
Dataflow
4
Organization Policy Service

GCP CDMC Whitepaper (40+ pgs)

Overview of the reference architecture and controls



GoogleCloudPlatform / **cdmc** Private

Watch 4 · Fork 0 · Star 2

4 branches · 0 tags

swainjo Merge pull request #16 from Google... 8cf74ff 2 days ago **48** commits

dashboard	Initial dashboard commit	last week
data-ingestion	Merge pull request #4 from GoogleCloudPlatf...	last week
data-quality	minor docs updates	last week
data-scanning	Update inspect_datasets.py	3 days ago
report-engine	renamed report-engine folder	3 weeks ago
tagging	added remaining steps to the tagging guide	3 weeks ago
.gitignore	changed setup and sucessfully run terraform ...	2 days ago
CONTRIBUTING.md	Add files via upload	last month
LICENSE.md	Add files via upload	last month
README.md	Fixed dashboard link in main README	last week
environment-variable...	Update environment-variables.example	last week
setup.sh	changed setup and sucessfully run terraform ...	2 days ago

README.md

Google Cloud CDMC Reference Architecture

This repository contains the CDMC reference architecture for Google Cloud. This

About

No description, website, or topics provided.

Readme
Apache-2.0 license
Activity
2 stars
4 watching
0 forks

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

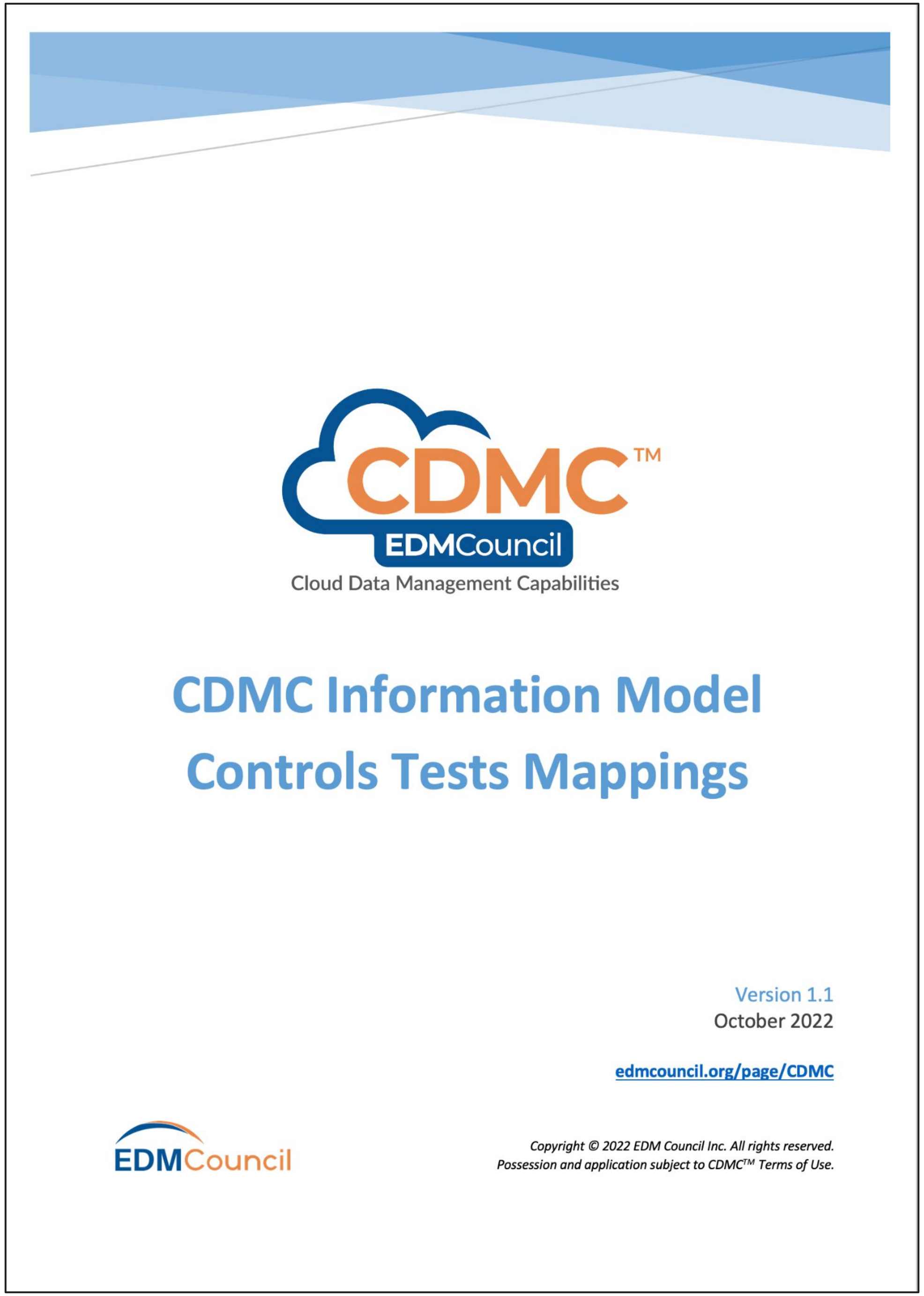
Contributors 5

Languages

Python 79.1% · Shell 16.5%

GitHub Repo

Publication of example assets (scripts, detectors, setup guide) into new standalone repo.



CDMC™

EDMCouncil

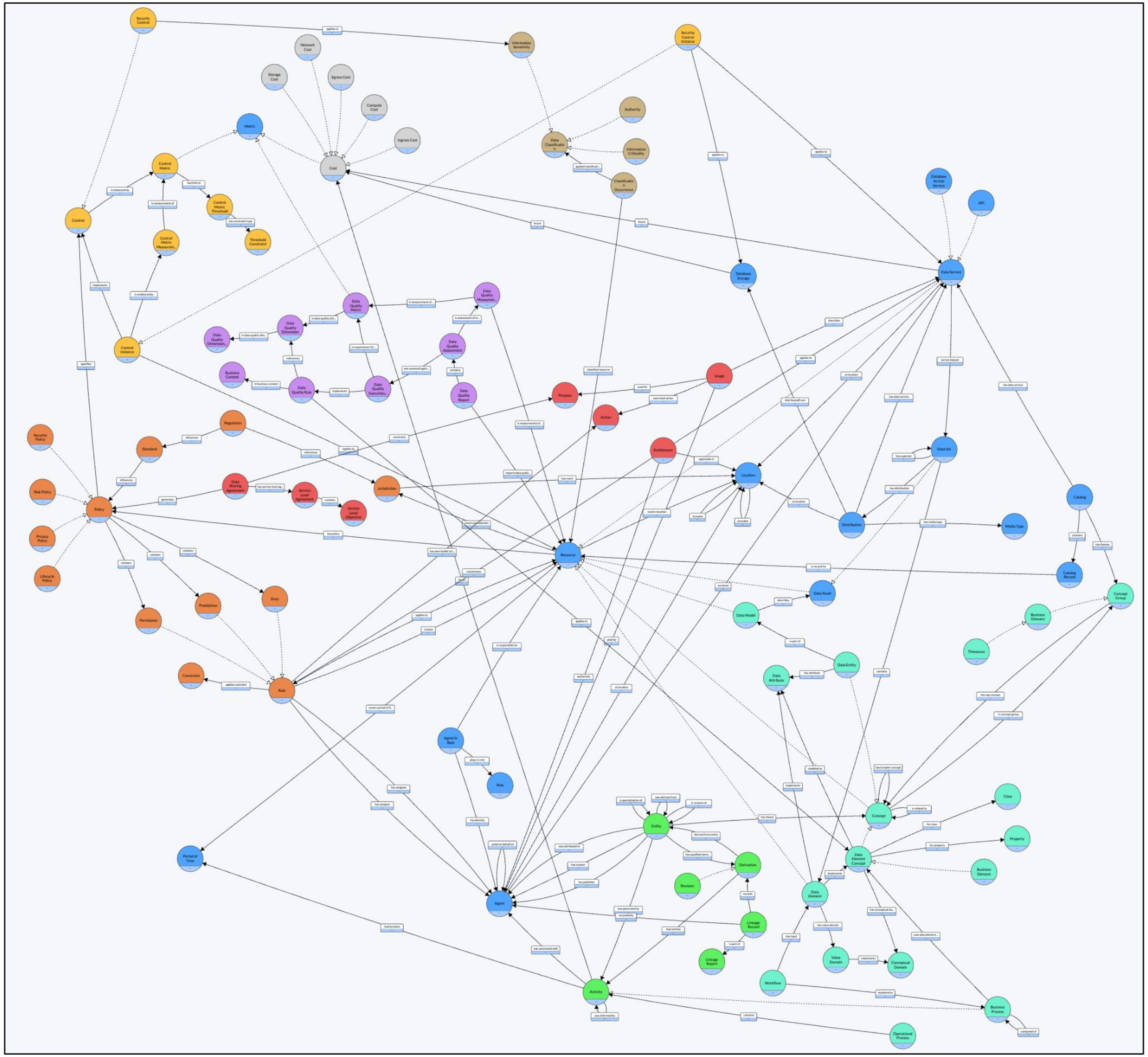
Cloud Data Management Capabilities

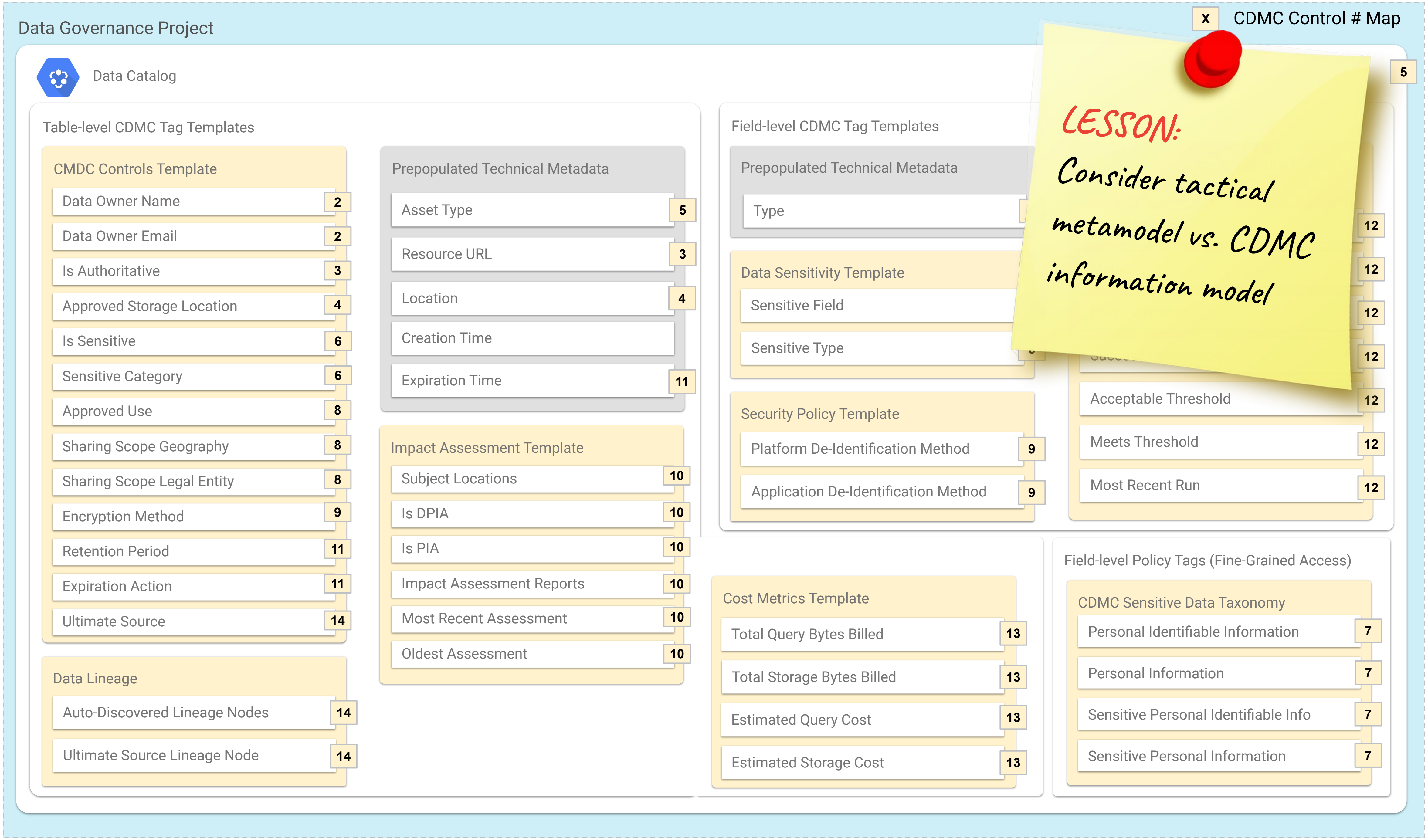
CDMC Information Model
Controls Tests Mappings

Version 1.1
October 2022

edmcouncil.org/page/CDMC

EDMCouncil





Data Governance Project

x CDMC Control # Map

Data Catalog

5

Table-level CDMC Tag Templates

CMDC Controls Template

Data Owner Name	2
Data Owner Email	2
Is Authoritative	3
Approved Storage Location	4
Is Sensitive	6
Sensitive Category	6
Approved Use	8
Sharing Scope Geography	8
Sharing Scope Legal Entity	8
Encryption Method	9
Retention Period	11
Expiration Action	11
Ultimate Source	14

Data Lineage

Auto-Discovered Lineage Nodes	14
Ultimate Source Lineage Node	14

Prepopulated Technical Metadata

Asset Type	5
Resource URL	3
Location	4
Creation Time	
Expiration Time	11

Impact Assessment Template

Subject Locations	10
Is DPIA	10
Is PIA	10
Impact Assessment Reports	10
Most Recent Assessment	10
Oldest Assessment	10

Field-level CDMC Tag Templates

Prepopulated Technical Metadata

Type	

Data Sensitivity Template

Sensitive Field	
Sensitive Type	

Security Policy Template

Platform De-Identification Method	9
Application De-Identification Method	9

	12
	12
	12
	12
Acceptable Threshold	12
Meets Threshold	12
Most Recent Run	12

Cost Metrics Template

Total Query Bytes Billed	13
Total Storage Bytes Billed	13
Estimated Query Cost	13
Estimated Storage Cost	13

Field-level Policy Tags (Fine-Grained Access)

CDMC Sensitive Data Taxonomy

Personal Identifiable Information	7
Personal Information	7
Sensitive Personal Identifiable Info	7
Sensitive Personal Information	7

LESSON:

Consider tactical metamodel vs. CDMC information model

How to use our assets!

Established Google Cloud Clients

1. Read CDMC capability and key controls frameworks
2. Review our whitepaper
3. Perform CDMC capability assessment
4. Perform CDMC controls gap analysis
5. Download and selectively adapt CDMC GitHub repo assets to address control gaps in your environment

New to Google Cloud

1. Read CDMC capability and key controls frameworks
2. Review our whitepaper
3. Evaluate and implement Secure Data Warehouse Blueprint
4. Download and implement CDMC GitHub repo asset in its entirety

Users of Other Clouds

1. Read CDMC capability and key controls frameworks
2. Review our whitepaper to identify any architectural patterns that you could reuse in your own control design
3. Reach out to your own provider and understand about how you can reuse their certified architecture

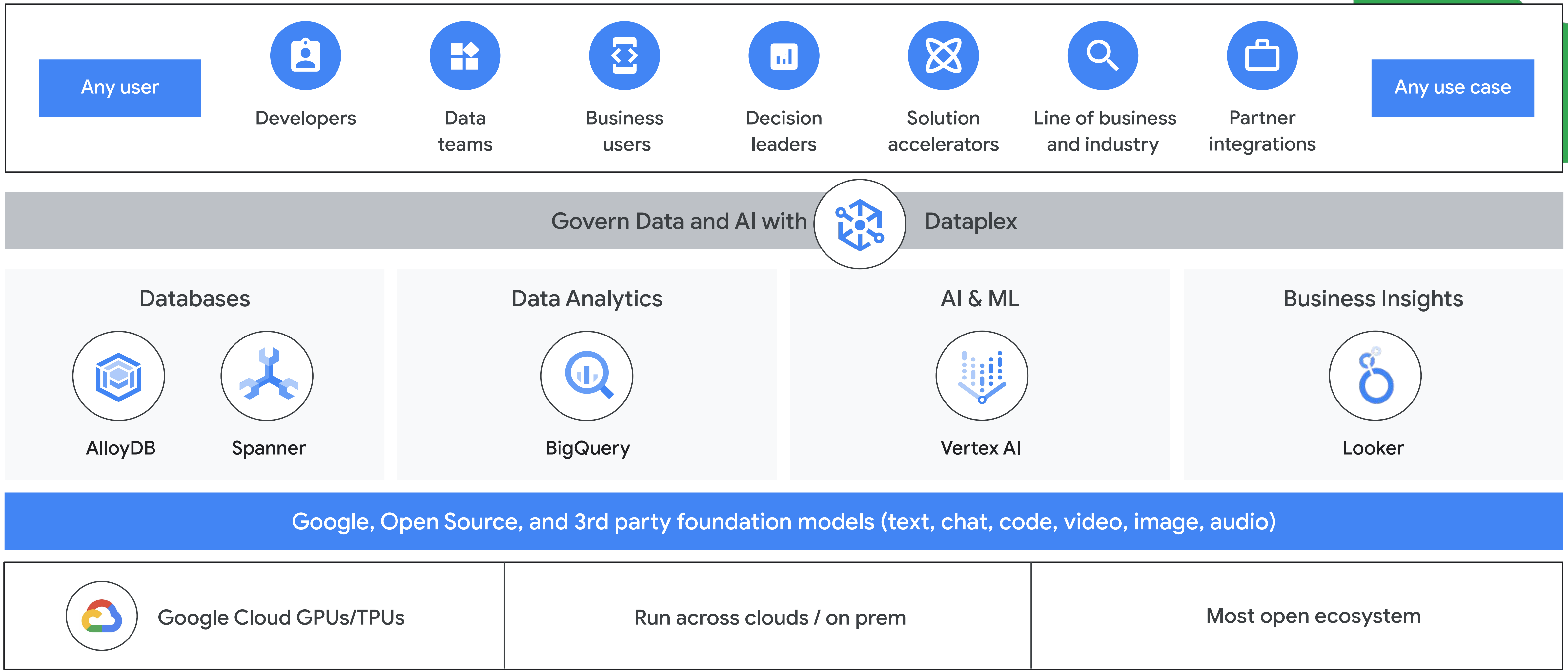
Questions?

Google Cloud

EDM Webinar

Google's Data and AI Cloud

Simplifying data and AI experiences for all users with Gen AI



Any user

Developers

Data teams

Business users

Decision leaders

Solution accelerators

Line of business and industry

Partner integrations

Any use case

Govern Data and AI with Dataplex

Databases

AlloyDB

Spanner

Data Analytics

BigQuery

AI & ML

Vertex AI

Business Insights

Looker

Google, Open Source, and 3rd party foundation models (text, chat, code, video, image, audio)

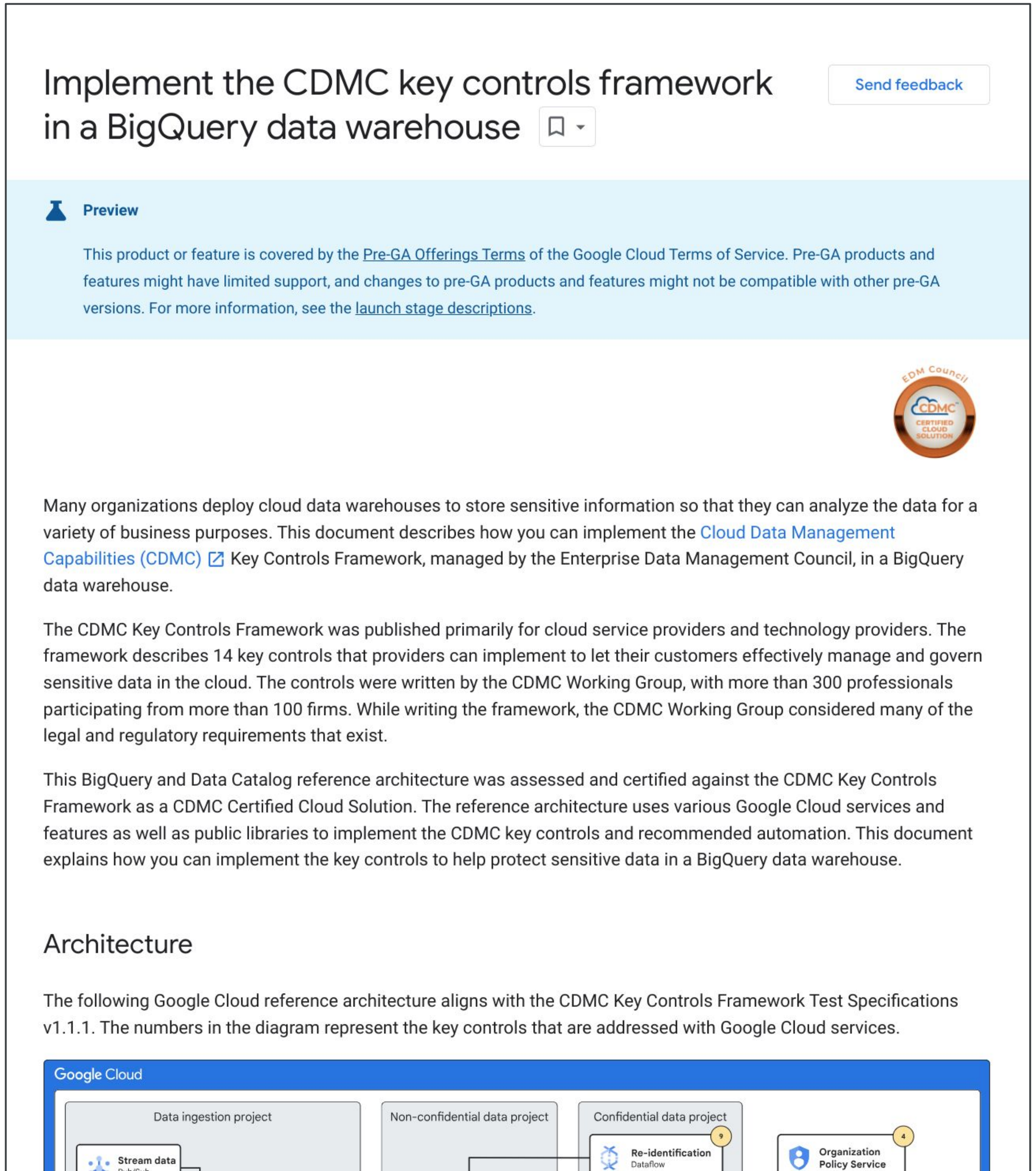
Google Cloud GPUs/TPUs

Run across clouds / on prem

Most open ecosystem

Recap of our lessons

1. Ensure **sufficient time** and investment is available
2. Check your team has the **right composition of skills**
3. **Seek clarification** in areas of ambiguity
4. Anticipate and budget for the **evolution of the platform**
5. Consider **bespoke vs. reusable** implementation
6. Consider starting with **limited automation** and iterate
7. **Common capability groupings** regardless of implementation
8. **Test data selection** critical
9. **Identify sponsoring clients early** and engage throughout
10. Validate the **minimum evidence** required to satisfy the assessor
11. Consider **tactical metamodel** vs. CDMC information model



Implement the CDMC key controls framework in a BigQuery data warehouse

Preview

This product or feature is covered by the Pre-GA Offerings Terms of the Google Cloud Terms of Service. Pre-GA products and features might have limited support, and changes to pre-GA products and features might not be compatible with other pre-GA versions. For more information, see the launch stage descriptions.

Many organizations deploy cloud data warehouses to store sensitive information so that they can analyze the data for a variety of business purposes. This document describes how you can implement the Cloud Data Management Capabilities (CDMC) Key Controls Framework, managed by the Enterprise Data Management Council, in a BigQuery data warehouse.

The CDMC Key Controls Framework was published primarily for cloud service providers and technology providers. The framework describes 14 key controls that providers can implement to let their customers effectively manage and govern sensitive data in the cloud. The controls were written by the CDMC Working Group, with more than 300 professionals participating from more than 100 firms. While writing the framework, the CDMC Working Group considered many of the legal and regulatory requirements that exist.

This BigQuery and Data Catalog reference architecture was assessed and certified against the CDMC Key Controls Framework as a CDMC Certified Cloud Solution. The reference architecture uses various Google Cloud services and features as well as public libraries to implement the CDMC key controls and recommended automation. This document explains how you can implement the key controls to help protect sensitive data in a BigQuery data warehouse.

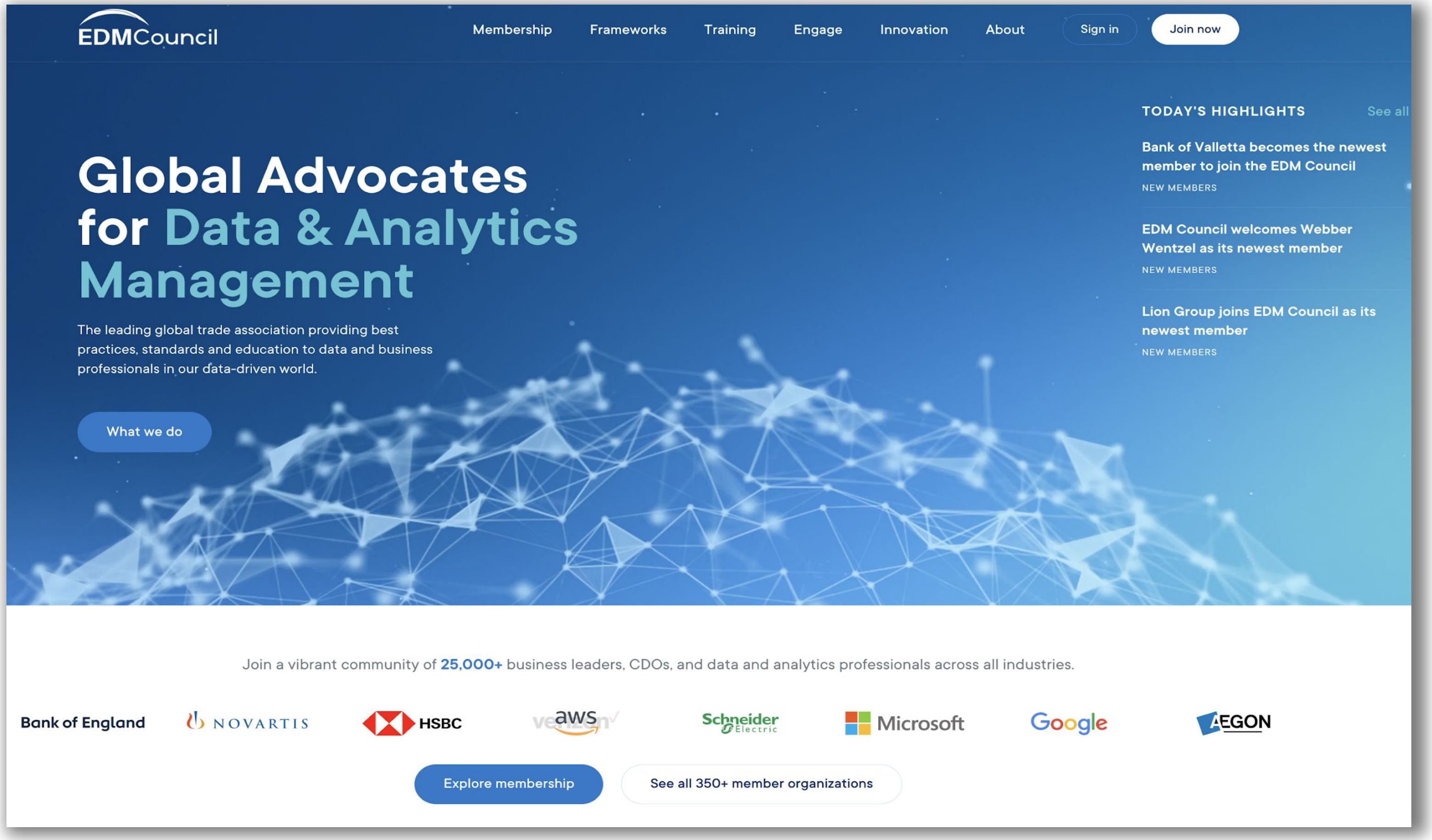
Architecture

The following Google Cloud reference architecture aligns with the CDMC Key Controls Framework Test Specifications v1.1.1. The numbers in the diagram represent the key controls that are addressed with Google Cloud services.

Google Cloud
Data ingestion project
Stream data
Non-confidential data project
Confidential data project
9
Re-identification
Dataflow
4
Organization Policy Service

Join EDM Council and our membership community of companies…

EDMCouncil



TODAY'S HIGHLIGHTS See all

Bank of Valletta becomes the newest member to join the EDM Council
NEW MEMBERS

EDM Council welcomes Webber Wentzel as its newest member
NEW MEMBERS

Lion Group joins EDM Council as its newest member
NEW MEMBERS

Global Advocates for Data & Analytics Management

The leading global trade association providing best practices, standards and education to data and business professionals in our data-driven world.

What we do

Join a vibrant community of **25,000+** business leaders, CDOs, and data and analytics professionals across all industries.

Explore membership

See all 350+ member organizations

350+ Member Firms
Cross-industry, including Regulators

25,000+
Professionals

Worldwide
Americas, Europe, Africa, Asia, Australia

edmcouncil.org

DCAM

CDMC EDMCouncil

Data ROI EDMCouncil

ESG EDMCouncil

Open Knowledge Graph INNOVATION LAB

eLearningCurve

EDM Webinar

Thank you!

FOR MORE INFORMATION:

Mose Tronci
Industry Solution Architect
Google Cloud
mtronci@google.com

Google Cloud

EDMCouncil