Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

Irish Revenue - Unified Data Platform

"Our governance journey to the Cloud"

Donnchadh Irish (Head of Analytics and Information Branch)
Kenneth Sythes (Statistician)

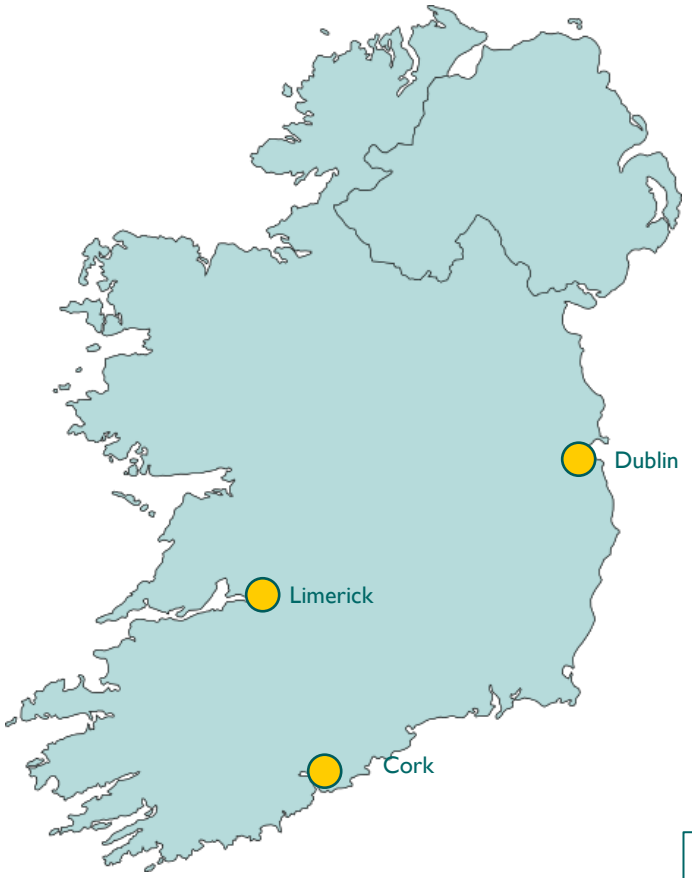
19th October 2026

Revenue – Irish Tax & Customs

Analytics & Information Management Branch

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

AIM Branch
Brainse Anailísíocht & Bainistíocht Faisnéise
Analytics & Information Management



Analytics & Information Management Branch

Donnchadh Irish

iXBRL Unit

Overseeing Revenue's iXBRL regime including:

- Development and maintenance of taxonomies
- Parsing and risk-assessment of returns

Analytics Units

Developing insight from Revenue's data holdings

- Deliver bespoke projects with high operational value
- Leveraging advanced ML and Statistical analysis
- Support divisions to improve their knowledge of their case bases.

Third Party Data Unit

Managing Revenue's third-party data regime including:

- Returns compliance,
- Data validation, and
- Matching

Data Sharing Agreements Unit

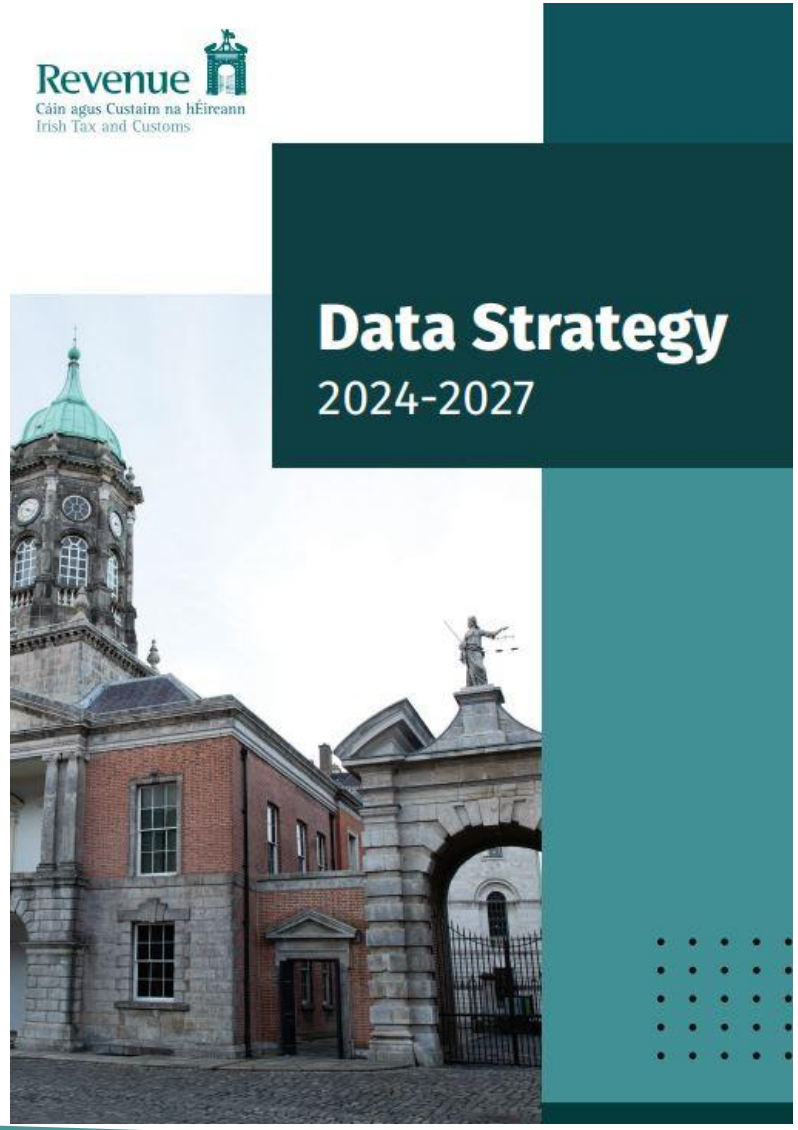
Managing Revenue's Data Sharing Agreements

- Catalogue of DSA's
- Identify designated roles under the DSA

AIM Branch focus is on **deriving value from data** and progressing **strategic data initiatives…**

Vision – Data Strategy…

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs



Data Principles
Seven clear and over-arching standards for data usage and management across the organisation

Strategic Pillars
Five key areas of focus and associated high-level outcomes required to achieve the Data Principles

Capabilities
Three functions that will be needed in Revenue to implement and realise our strategic vision



Collecting the right data by design
We will collect data that is necessary and relevant – ensuring that it is appropriately underpinned by legislation.

Ensure data security throughout
We will make data available and accessible in a secure and responsible way to those individuals and teams in Revenue that need it.

Data as 'a single source of truth'
We will collect data once and actively maintain it as a 'single source of truth' to avoid duplication and facilitate its re-use to the greatest extent allowable.

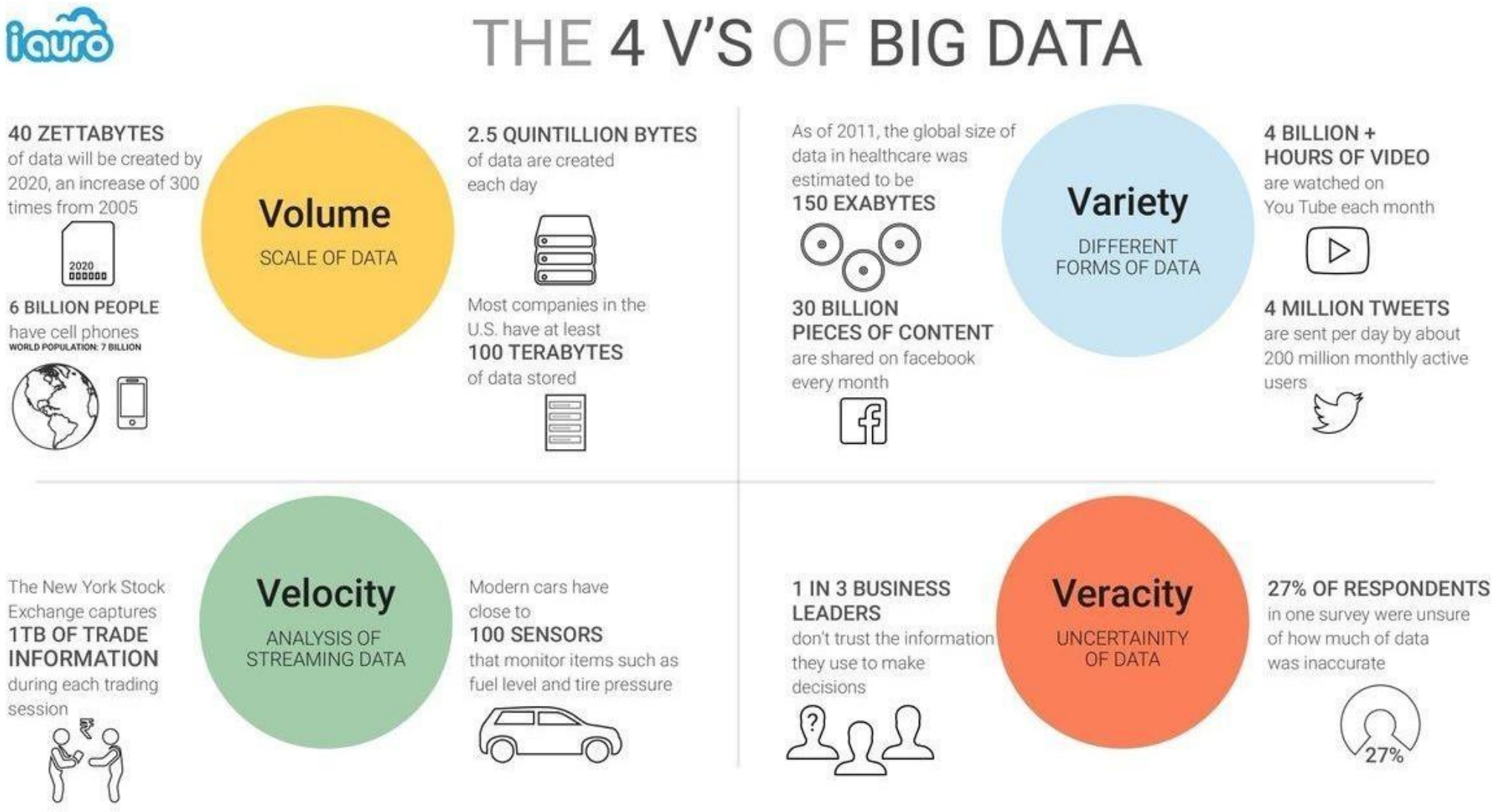
Single taxpayer-centric view
We will ensure that our data strategy supports the further development of our single taxpayer-centric model.

Consistent usage and management practices
We will use standardised and consistent practices to ensure the quality of data being submitted externally and for data usage and management across the organisation.

Business ownership for data
Every data asset will have a business owner in-place to ensure it is managed effectively throughout its life cycle.

Data for public service
We will collaborate with the wider public service through responsible sharing of data with other public service bodies.

Imperatives to evolve…



Reference : http://www.ibmbigdatahub.com/infographic/four-vs-big-data

Growing Data, Shrinking Control…

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

The Governance Gap

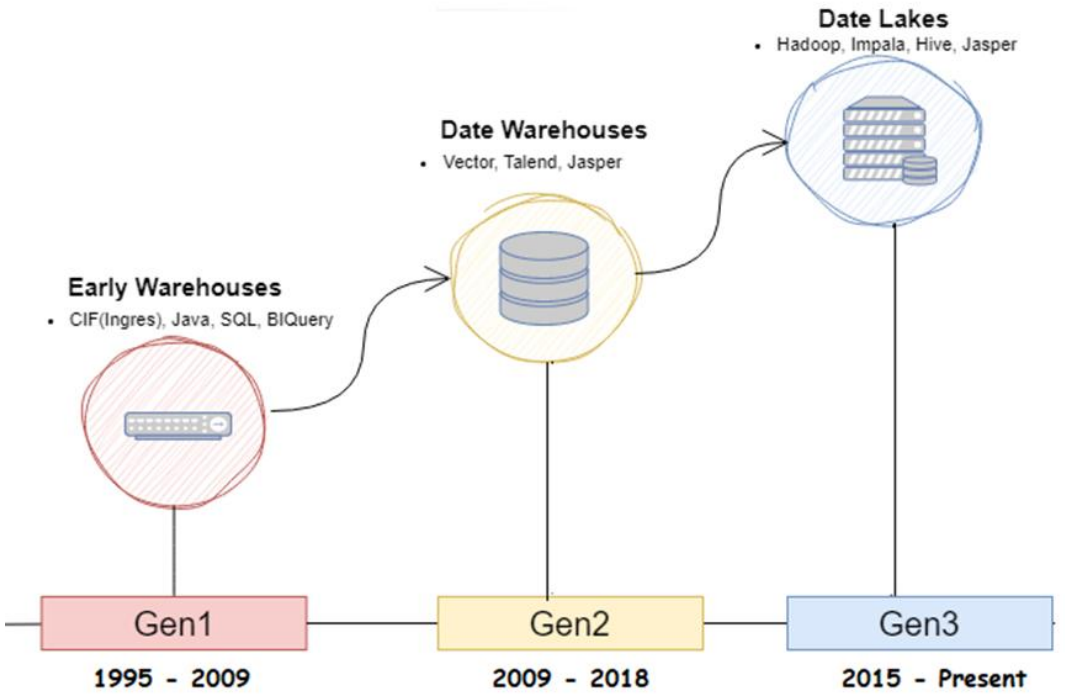
- Data volume and variety are growing far faster than our current governance frameworks
- No single, organisation-wide metadata catalogue - so even skilled staff can't always find, trust, or understand data that already exists
- Data classification is cumbersome – can take time to implement
- Weak governance also blocks opportunity - responsible AI/ML-driven compliance detection requires well-governed, representative data
- A 'copy-based' data culture still dominates: extract, duplicate, reconcile - creating multiple conflicting versions of the truth
- No overarching applied business ownership model - meaning accountability for a dataset's quality and lifecycle often falls back to the technicians



This Photo by Unknown Author is licensed under CC BY-SA-NC

Where we are…

- A burning platform
  - Three generations of legacy architecture becoming more difficult to maintain
- OECD Analytics Maturity Model 2022
  - an opportunity to progress
- Total volume c.200TB



Unified Data Platform- Google Cloud

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

- Bring all information **together into a single, secure place**.
- State-of-the-art, **unified AI-Lakehouse architecture** on Google Cloud.
- Leveraging BigQuery, Google Cloud Storage, and BigLake
- **Break down data silos** and establish a **single source of truth** for all data practitioners.
- **Analyse data and find insights** without having to constantly move or copy information.
- **Accelerate time-to-insight** and lower the technical skills barrier
- Enhanced by **powerful AI capabilities** - Gemini as embedded assistant within BigQuery Studio
- Intelligent **code completion**, translate **natural language prompts** into complex SQL queries and **automatically generate visualisations** for business analysts.



A Digital Future for Ireland: How Revenue is Partnering with Google Cloud

10/06/2026 Ireland's Office of the Revenue Commissioners is embarking on a seven-year partnership with Google Cloud to modernise its data infrastructure.

Key implementation principles

- Ability to disconnect Cloud at any time due to
  - Sovereignty issues
  - Disaster recovery of front-line systems
  - Cost
- No time-critical / front-line processing done on Cloud

Data Management and Governance is integral to efficient and effective use of the platform. The Revenue Unified Data Platform integrates as much of this as possible by design making data Governance seamless.

Interacting with the UDP…

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

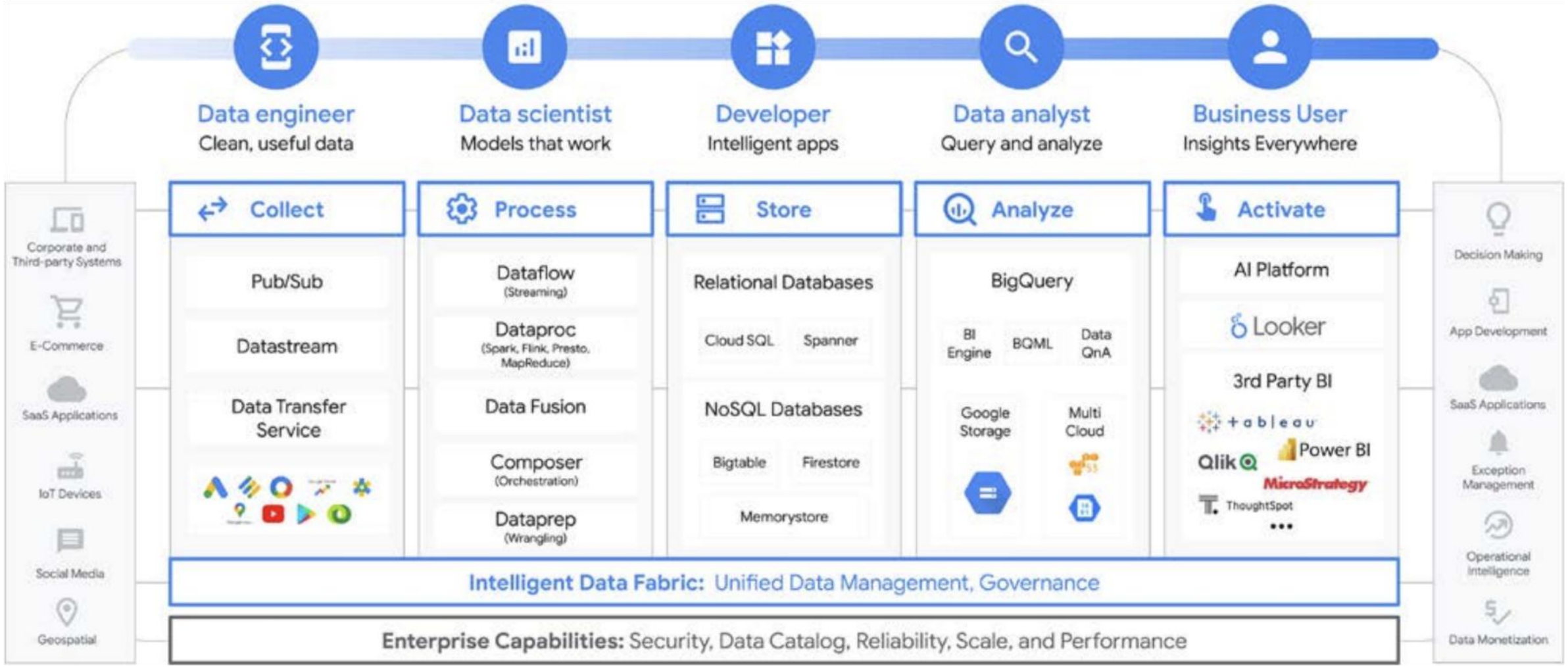


Image source: *https://cloud.google.com/resources/googlecloud-unified-analytics-data-platform-paper*

Implementation roadmap

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

May 2026

Kicked off engagement with Google Professional Services

June 2026

Setup of GCP landing zone and Data Platform

July – Aug 2026

Migration of first pilot Dataset to Cloud Data Platform

Stood up "Metadata Hit Squads"

Explore Agentic AI solutions to aid data and metadata assimilation

Oct 2026 – on

Continue Migration

Incorporate Open Knowledge Format for metadata

Explore data analytics capabilities

Seek to leverage Agentic AI

Determine education requirements and pathways

May-Aug 2026

Google-led twice weekly workshops to understand requirements and architect GCP land zone and data platform

June-July 2026

Google and BIRA test and refine data migration

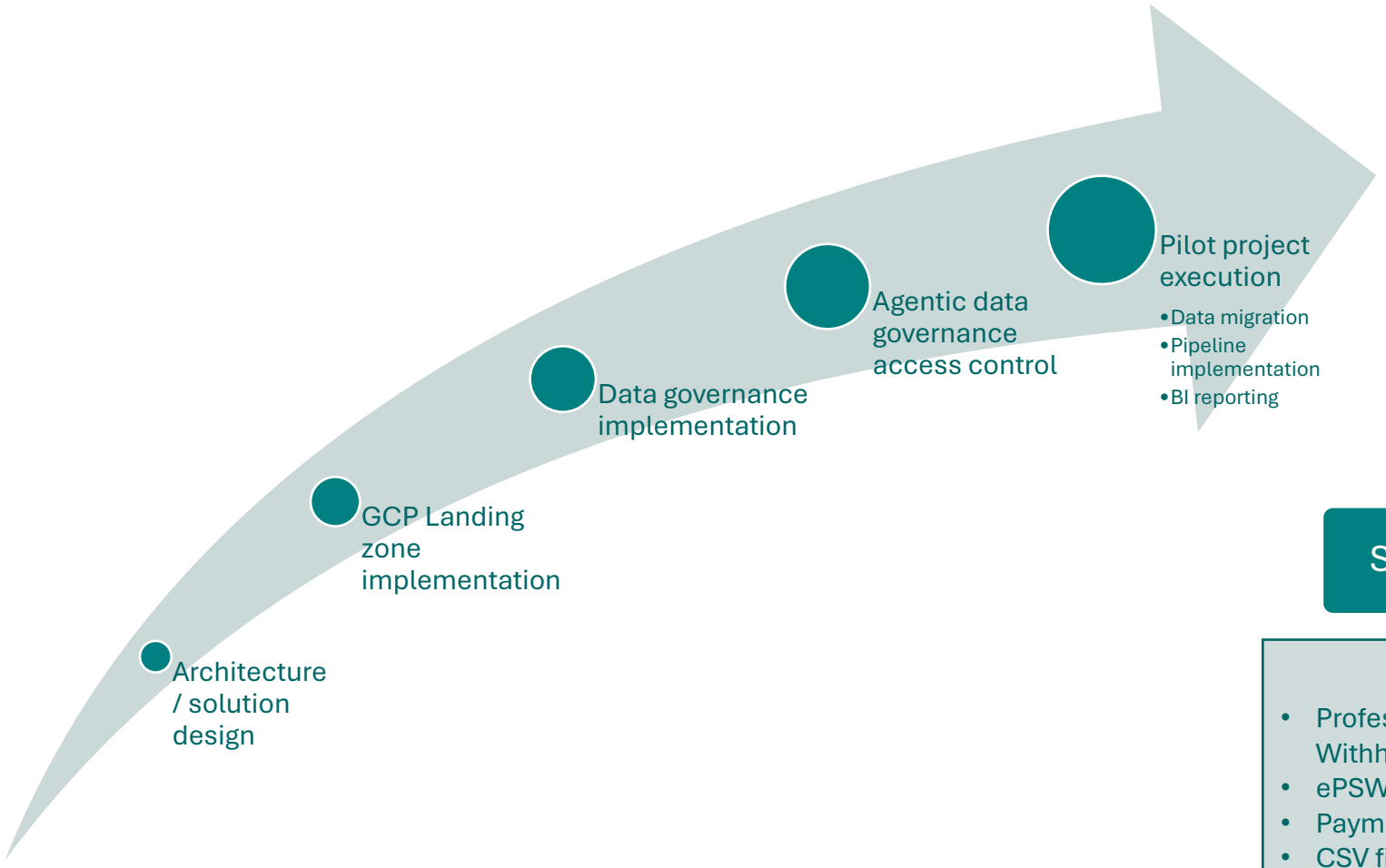
Workshops exploring Data Domains

Sept 2026

Migration of richer datasets to explore data analytics capabilities of platform

Implemented the metadata markdown – Open Knowledge Format

Staged migration approach



Pilot project execution

- Two pilot projects
  - (i) ePSWT, and
  - (ii) Form 11
- BI reporting
  - Will point existing Jaspersoft tools at new data platform
  - Replicate existing reporting
  - Ensure accuracy of migrated data
- 100GB data

Straightforward → Increasing complexity

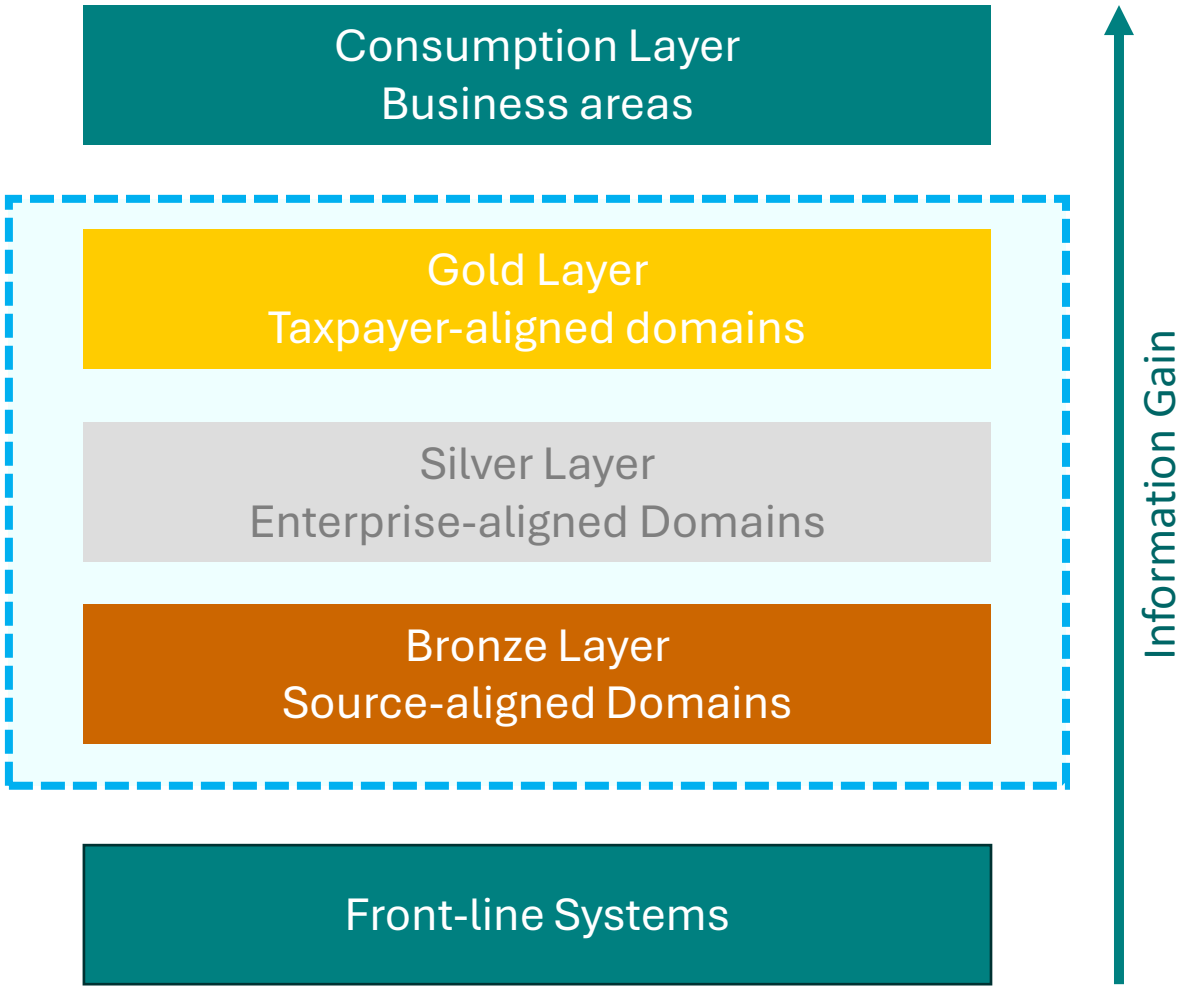
ePSWT

- Professional Services Withholding Tax (PSWT)
- ePSWT implemented July 2021
- Payment Notifications via ROS
- CSV file specification
- Compact, well-defined dataset (12 structured fields)

Form 11

- Income Tax Return and Self-Assessment
- Return of income, charges, capital gains and claim for tax credits, allowances and reliefs
- 42 pages of complex fields including free-text fields

Medallion architecture



Data Domains
- Provide clear view of data lineage and stakeholders

Business Data Ownership
- **Top-down business ownership**
- Defined from the **nine** data domains

Meta-data
- Have set-up meta-data 'hit squads' on priority datasets
- Leveraging previous inventories of datasets and data products across the organisation
- Implemented the metadata markdown – Open Knowledge Format

Cross-cutting relationships
- Reference and Master Data

Data Quality
- Implement stage-gates between medallion layers
- Reflect data quality metrics and data processes

Data Products
- Identify data products and users
- Statistical and aggregation requirements

Personal Division
Domain D1 Individuals

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

Consumption Layer (Business areas)

- Personal Division
- Business Division
- Medium Enterprises Division
- High Wealth & Financial Services Division
- Large Corporates Division
- Customs Division
- Investigations, Prosecutions & Frontier Mgt Division
- Collector General's Division

Gold Layer (Taxpayer-aligned domains)

- Domain D1 Individuals
- Domain D2 Exempt Bodies
- Domain D3 Sole Traders
- Domain D4 Companies
- Domain D5 Estate Trusts & Wealth Management
- Domain D6 Customs
- Domain D7 Non-Resident / Cross-border
- Domain D8 Intermediary / Agents
- Domain D9 Cross-Cutting Relationships

Silver Layer (Enterprise-aligned Domains)

- PMOD
- F11
- F12
- CT1
- F1
- CG1
- IT38 (CAT)
- IT39
- VAT
- VIES
- OSS VAT
- IOSS
- RCT
- LPT
- SAD
- EMCS
- GloBE P2
- CRS/F ATCA
- DAC 1..8
- ROM1
- 3PR

Bronze Layer (Source-aligned Domains)

- ITP_DB
- CRS_DB
- eCustoms_DB
- VIES_DB
- F11_XML
- F12_XML
- CT1_XML

Front-line Systems

- On Premises
- Cloud

Business Division
Domain D3 Sole Traders

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

Consumption Layer (Business areas)

- Personal Division
- **Business Division**
- Medium Enterprises Division
- High Wealth & Financial Services Division
- Large Corporates Division
- Customs Division
- Investigations, Prosecutions & Frontier Mgt Division
- Collector General's Division

Gold Layer (Taxpayer-aligned domains)

- Domain D1 Individuals
- Domain D2 Exempt Bodies
- **Domain D3 Sole Traders**
- Domain D4 Companies
- Domain D5 Estate Trusts & Wealth Management
- Domain D6 Customs
- Domain D7 Non-Resident / Cross-border
- Domain D8 Intermediary / Agents
- Domain D9 Cross-Cutting Relationships

Silver Layer (Enterprise-aligned Domains)

- PMOD
- **F11**
- F12
- CT1
- F1
- CG1
- IT38 (CAT)
- IT39
- **VAT**
- VIES
- **OSS VAT**
- **IOSS**
- RCT
- LPT
- SAD
- EMCS
- GloBE P2
- CRS/F ATCA
- **DAC 1..8**
- ROM1
- 3PR

Bronze Layer (Source-aligned Domains)

- **ITP_DB**
- **CRS_DB**
- eCustoms_DB
- **VIES_DB**
- **F11_XML**
- F12_XML
- CT1_XML

Front-line Systems

- On Premises
- Cloud

Third Party Returns
Domain D7 Non-Resident / Cross-border

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

Consumption Layer (Business areas)

- Personal Division
- Business Division
- Medium Enterprises Division
- High Wealth & Financial Services Division
- Large Corporates Division
- Customs Division
- Investigations, Prosecutions & Frontier Mgt Division
- Collector General's Division

Gold Layer (Taxpayer-aligned domains)

- Domain D1 Individuals
- Domain D2 Exempt Bodies
- Domain D3 Sole Traders
- Domain D4 Companies
- Domain D5 Estate Trusts & Wealth Management
- Domain D6 Customs
- **Domain D7 Non-Resident / Cross-border**
- Domain D8 Intermediary / Agents
- Domain D9 Cross-Cutting Relationships

Silver Layer (Enterprise-aligned Domains)

- PMOD
- F11
- F12
- CT1
- F1
- CG1
- IT38 (CAT)
- IT39
- VAT
- VIES
- OSS VAT
- IOSS
- RCT
- LPT
- SAD
- EMCS
- GloBE P2
- CRS/F ATCA
- DAC 1..8
- ROM1
- 3PR

Bronze Layer (Source-aligned Domains)

- ITP_DB
- CRS_DB
- eCustoms_DB
- VIES_DB
- F11_XML
- F12_XML
- CT1_XML

Front-line Systems

- On Premises
- Cloud

Bridging the gap…

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs



#1 – ZERO COPY



#2 – EXTRACT COST MINIMISATION



#3 – SYSTEMS AS RECORDS OF FACT



#4 – DATA MODEL



#5 - IMPACT

Impact

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs



Same but better

From copy-based evidence

(screenshots, ad-hoc extracts)

To connection-based evidence

More focus on the data model and what it delivers.

Better alignment with GDPR and data minimisation – users see only what they are authorised to see.

Improved monitoring – access and queries are logged centrally.

Performance and control – large unfiltered extracts are replaced by controlled queries.

Reproducible data products.

Enabling the Governance foundation in the UDP…

Revenue
Cáin agus Custaim na hÉireann
Irish Tax and Customs

BigQuery Warehouse

Serverless, scalable single source of truth organised into Bronze (raw), Silver (validated), and Gold (curated, BI - ready) layers

Knowledge Catalogue

The organisation's metadata hub - definitions, lineage, and quality information discoverable in one place, tied to operational tags for automated classification

Identity & Access Management (IAM)

Access enforced by role and function, not personell deployment location - implementing data access policy as an executable process.

Dataflow, Cloud Functions

Automated ingestion and transformation pipelines connecting source systems into the governed cloud platform, replacing manual import and ad-hoc ingestion

Automated Data Quality Checks

Quality rules from our framework run continuously within pipelines, with results published back to the Knowledge Catalogue for full disclosure

Thank you for your attention

We hope that you found this useful

We are happy to take any questions or comments



"... if AI is being used for one of these critical functions, companies must now...show extensive transparency around how that system works. They have to produce a performance monitoring plan to show that that system is being monitored over time with thresholds that are acceptable for that use case."

Louise McCormack, Analytics Summit, 2025