ARVEST Collibra Google

## Trusted Data, Smarter Decisions:

# Arvest Bank's Blueprint for Cloud Migration Success

October 28th

# Today’s speakers

**Joshua Cohen**
Data Engineering Manager
Arvest Bank

**George Keller**
Partner Engineering
Google Cloud

**Marc Campabadal**
Customer Advocate
Collibra


**Carlos Tifa**
Partner Account Manager
Collibra

# Agenda

01 Introduction to Arvest Bank and its challenge

02 The strategy for building trust

03 The solution in action

04 What's next?

05 Collibra+Google: Better together

06 Q&A

# About Arvest Bank

With more than $26 billion in assets, Arvest is a full-service bank delivering financial solutions to individuals and businesses of all sizes.

Since 1961, Arvest has been committed to providing financial solutions that help communities thrive and grow. Today, the bank has more than 200 locations throughout Arkansas, Oklahoma, Kansas and Missouri.

Arvest is known for its commitment to the communities it serves and to attracting, hiring and retaining a diverse group of talented people.



ARVEST

# Our footprint


## Number of associates

Kansas = 192
Missouri = 592
Oklahoma = 1,733
Arkansas = 3,541
Other states = 176

**Total = 6,234**

*As of December 31, 2024*

**1,172** associates are 100% remote across 32 states

**4,618** associates are remote-capable in some capacity

Communities: **112** | Branches: **222** | Asset size: **$26.9 billion**

**# of data engineer associates: 36**
**# of data steward associates: 40**



ARVEST


# Enhanced banking services

Credit Cards – Arvest offers a full suite of business and consumer products and services.

Mortgage – The mortgage division has done~ $2 billion in loan production for five consecutive years and has a $35 billion mortgage servicing portfolio.

Small Business – Arvest is an SBA Preferred Lender

Equipment Finance – Arvest Equipment Finance has placed on the Monitor 100 for eight consecutive years.



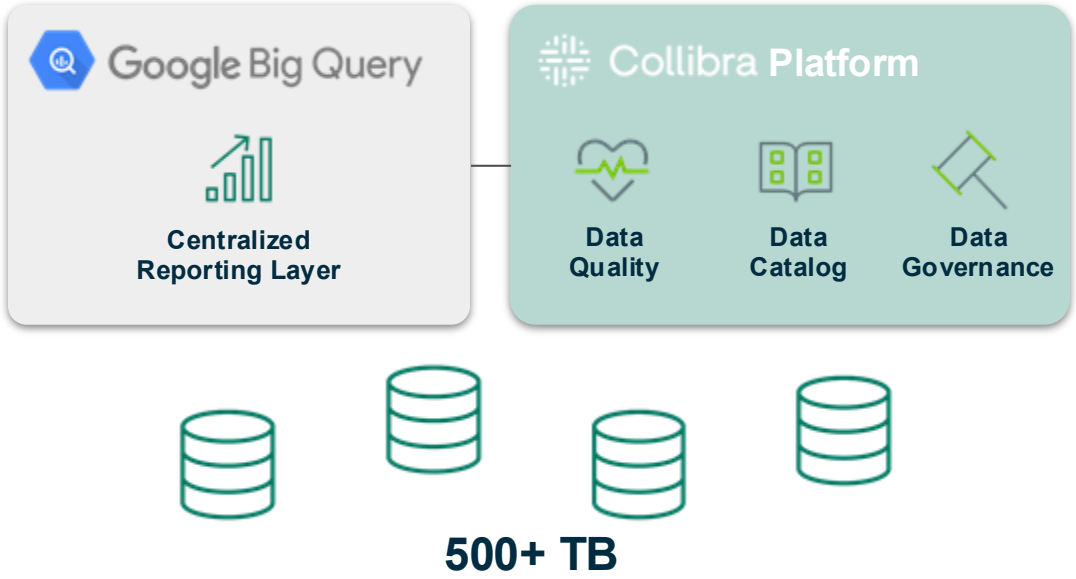
ARVEST

# The Vision

## A unified data platform

### Why?

To break down data silos and enable advanced analytics to support the bank's business objectives and initiatives in a timely manner.



# The Challenge

"We don't trust this data"

## Core problems

1- User adoption was slightly low.

2- Lack of trust in the data:

- *"Is this data accurate? It doesn't match the old system."*
- *"Is this data complete? I think records are missing."*
- *"Who can I even ask about this data?"*

# The strategy for building trust

# The Shift

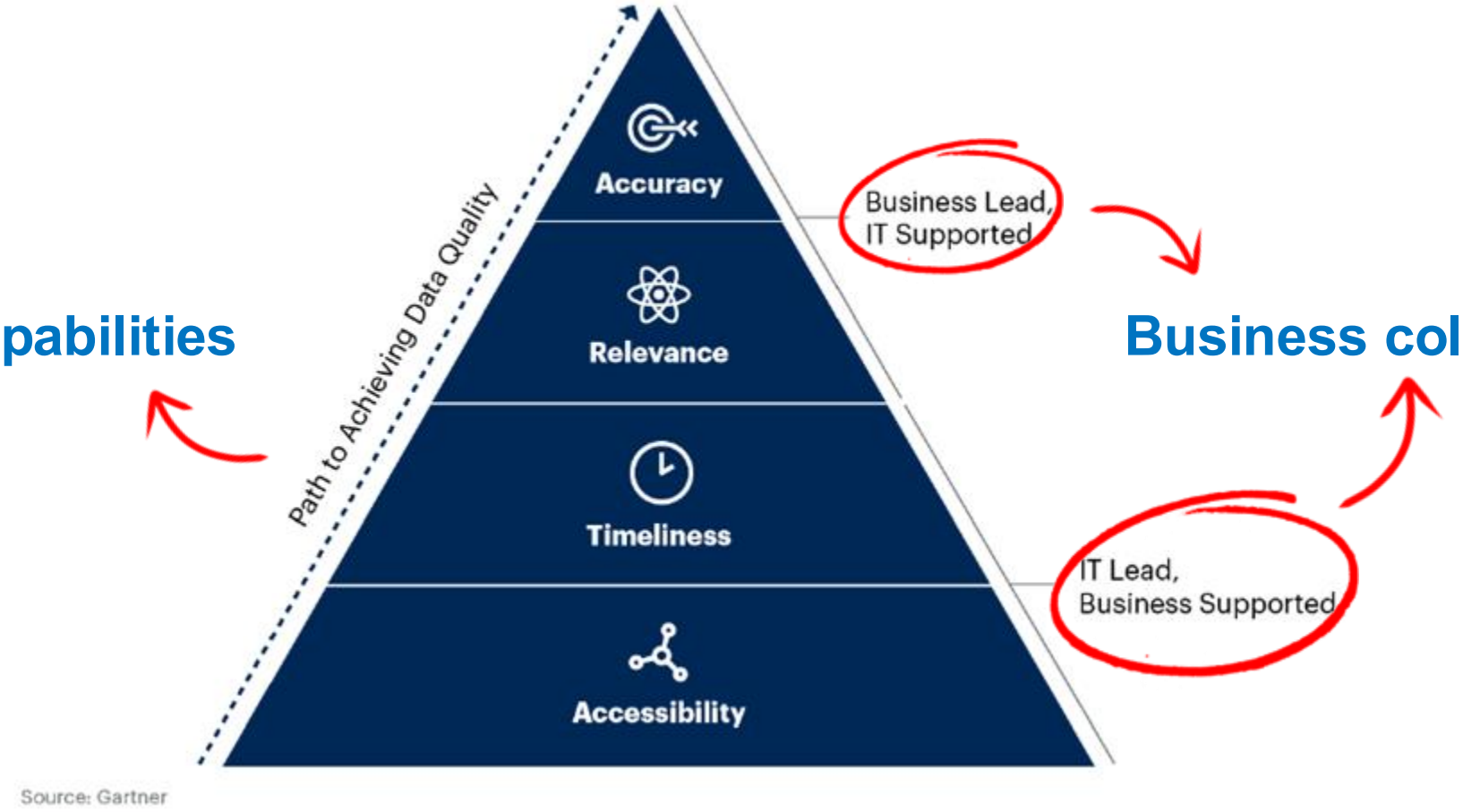
## The Data Engineering team took the lead implementing the solution

### Why the Change?

- Support the adoption of the GCP platform directly.
- Traditional governance felt separate from the platform build.

# Gartner’s path to achieving data quality

Technical capabilities enable technical and business teams for a solid data quality foundation




**Technical capabilities**

**Business collaboration**

**Data Quality Foundation**

# The three pillars of data trust at Arvest

## Clear ownership & stewardship

## Foundational data quality

## Radical transparency

# The solution in action

# Step 1: Validating migrated data

## Ensuring the integrity of the data from source to target

### How we did it

- For every table moved to BigQuery, we established a baseline of trust.
- We implemented foundational DQ checks that were simple but powerful:
  a. Row counts match: Does the number of records in BigQuery match the source system?
  b. Primary Keys are valid: Are all primary keys unique and not null?
  c. Column counts match: Is the table structure identical to the source?

### Result

- This immediately catches migration errors and provided a first layer of confidence.

# Step 2: Custom Rules

## Ensuring Critical Data Elements (CDE) meet the right expectations

### How we did it

- Collaboration with the product team to identify CDEs, as defined by Data Governance.
- With the Data Stewards we define the conditions that each CDE must meet to establish a business logic.
- Then we translate the business logic into custom rules and apply them to the CDEs

### Result

- Every CDE has a data quality rule and a score that we can report on for the Data Governance team driving shared ownership

# Step 3: Making data readiness visible

## Ensuring data consumers have access to trusted data

### Our two-flag system for data readiness in Collibra:

- **Status:** `Accepted`
  a. **What it means:** A business steward has reviewed the data asset in Collibra.They've enriched it with a clear business definition, identified critical data elements (CDEs), and confirmed its business purpose.
- **Status:** `Implemented`
  a. **What it means:** The Data Engineering team has attached Collibra DQ rules. Foundational *and* critical business rules are now running against the table in BigQuery.

### Result

- Our users can now filter for "Implemented" tables, knowing they are fully documented, owned, and monitored. This approach enables our **curated data model.**

# Step 4: Streamlining issue management

Ensuring remediation actions are taken and documented

## Our proactive issue management process

1. **Review**: Our team reviews DQ job results weekly.
2. **Create**: When a rule fails, we create a formal issue directly within Collibra.
3. **Assign**: The issue is automatically linked to the specific table asset and assigned to the designated owner.
4. **Track**: Everyone can see the issue, who is working on it, and its status. No more chasing emails or spreadsheets.

## Result

- This creates **visibility, transparency, and accountability.**

# What’s next?

# What’s next?

## A glimpse into the future

### Lineage & diagrams

Provide full, end-to-end visibility of data flow from source to report.

### Data marketplace

Empower users to easily "shop" for certified, trusted data assets for their projects.

### Custom reporting

Report on accuracy, completeness, timeliness and validity and be able to report across the critical data elements so DG team can focus on remediation.

### Improved issue response

Reduce the response time by reviewing the issue management process and implementing new processes and enablement.

# Collibra & Google Cloud Better Together

# A strategic partnership for your data journey

## Multi-year partnership agreement

Funded resources

Engineering

## Capital G Investments

Advisory

Portfolio companies

Customer roundtables

## Executive sponsorship

CEO

Product management

Sales

Alliances

Marketing

## GTM Strategy

Data & Analytics Specialists

Events

Better Together

Marketplace

Awards

## Technology collaboration

Dataplex

BigQuery

Vertex-Gemini

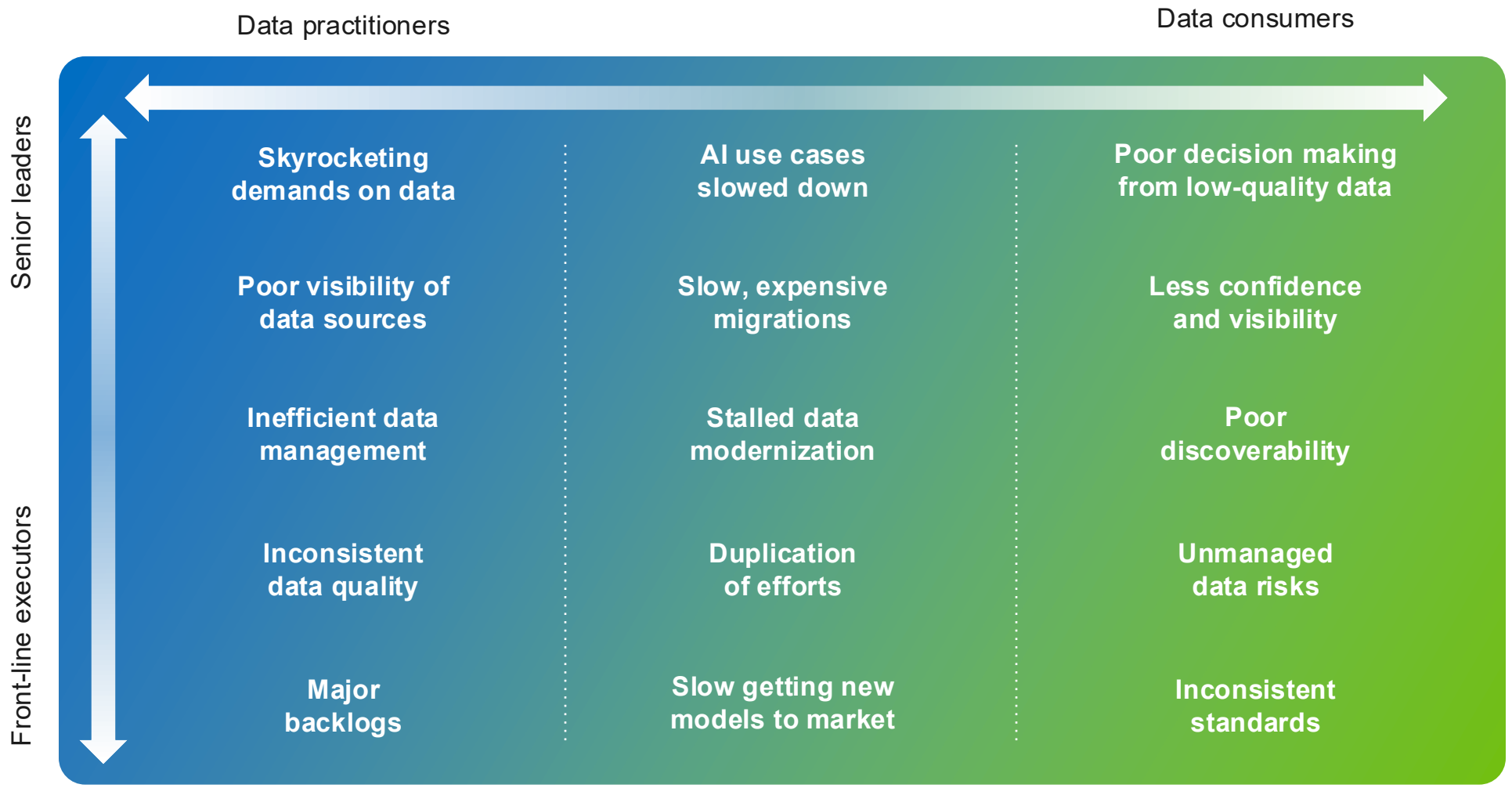
Collibra

# The challenge: fragmented data and stalled AI innovation



Collibra

# Governance fragmentation can cause pain




Collibra

# Fragmentation leads to poor data dissemination to consumers

**44%**

Data professionals say they **don't have access to data that they need** to do their jobs [1]

**41%**

Executives point to "**substandard data** as the facet of their data operations most in need of improvement" [2]

**33%**

Leaders cite addressing **untimely data delivery** as a top priority [2]

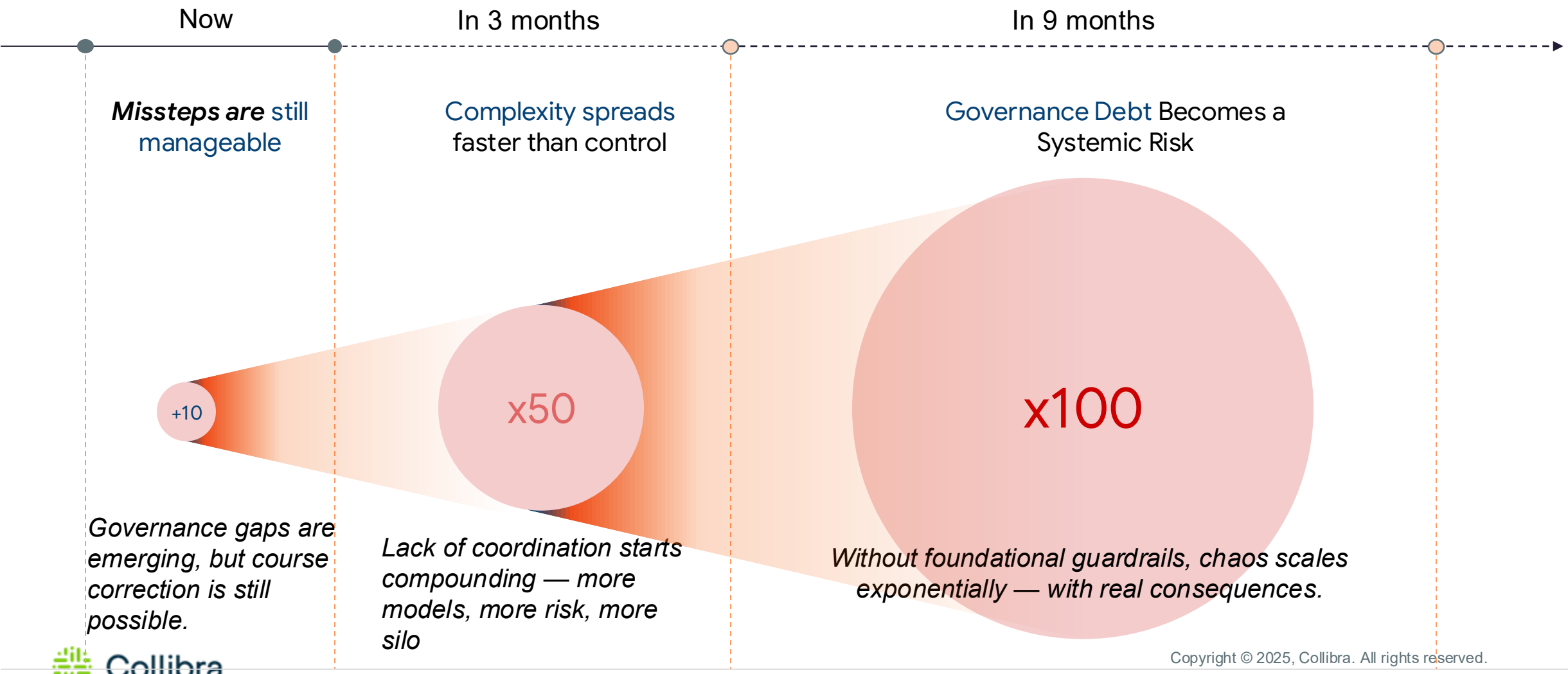


These factors result in **underutilized data**, **reduced productivity**, **eroded trust** and **compromised decision-making.**

**Sources:**
1. IDC PlanScape: Data as a Product, May 2024
2. Modernizing data with strategic purpose - MIT Technology Review, April 2024
3. Navigating the Planes of Enterprise Intelligence Architecture, June 2023

Collibra

# Lack of data dissemination will eventually lead to AI projects stalling or inevitably failing



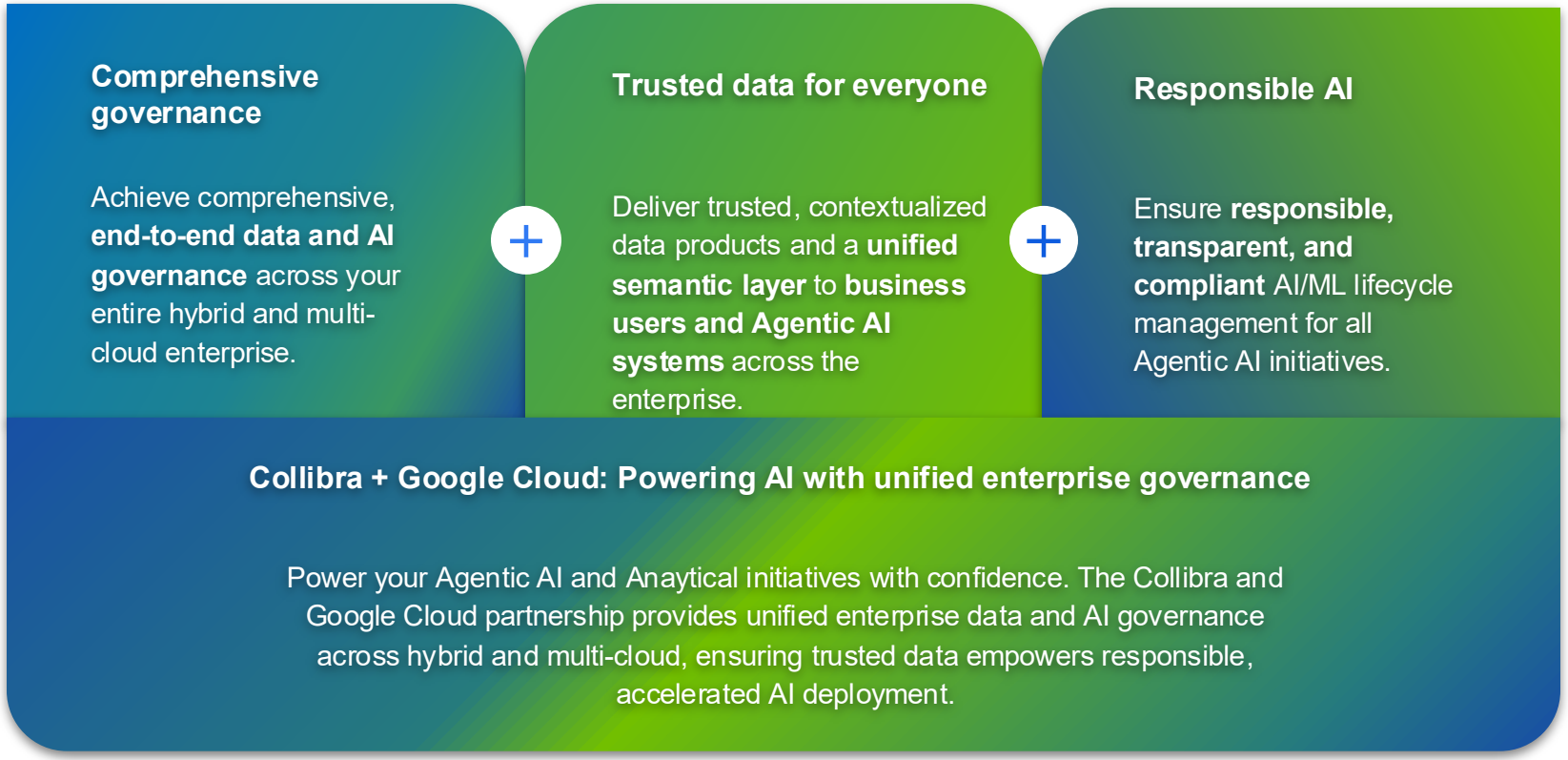

Collibra

# The solution: Collibra and Google Cloud



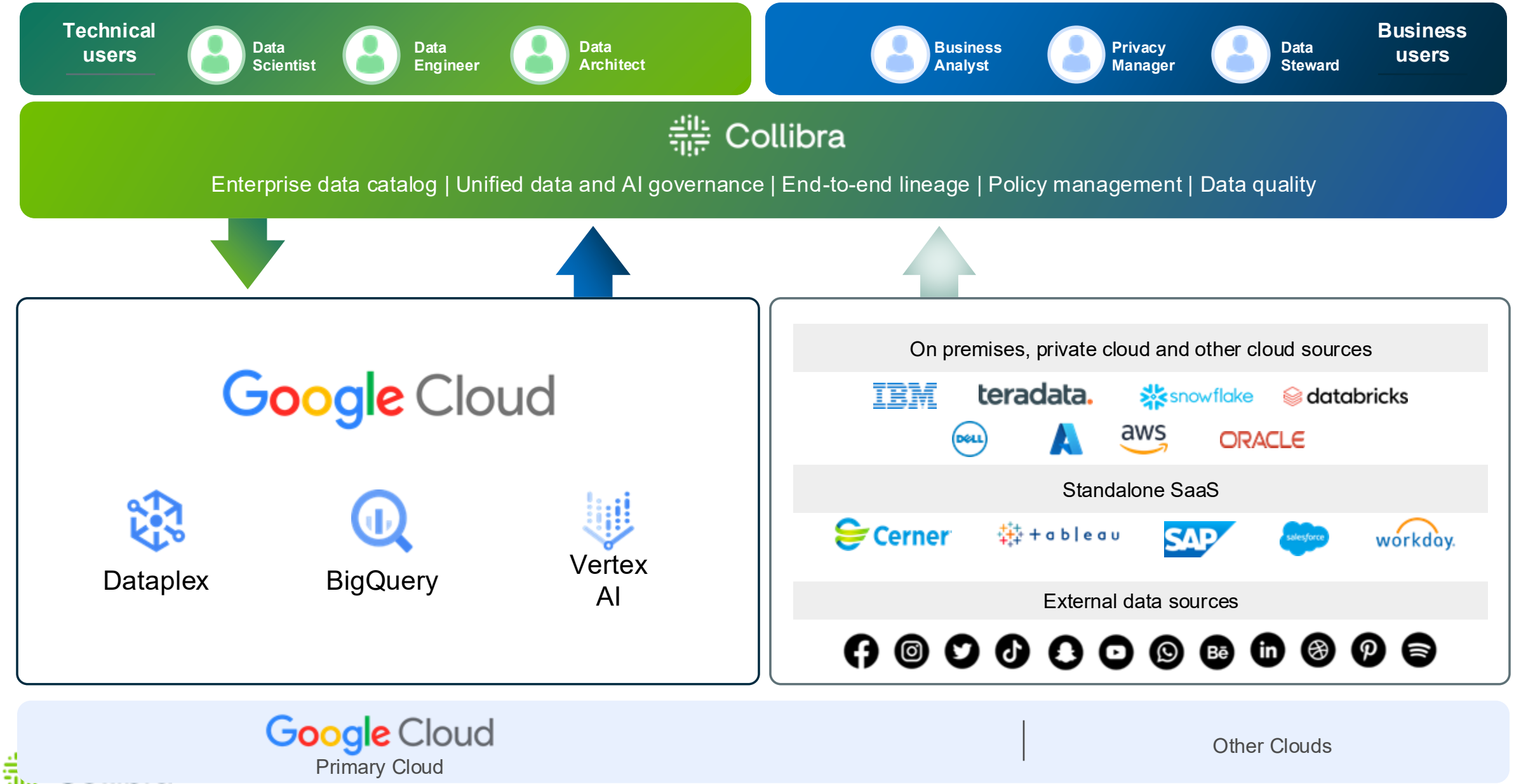
Collibra

# A unified approach to data and AI governance




Collibra

# Collibra & Google Cloud: better together



# Partnership in action at Arvest


## The Scalable Platform

Provided the powerful, scalable infrastructure with **BigQuery** to handle over 500 TB of data from dozens of source systems.


## The Confidence Layer

Provided the tools for **Governance, data quality, and issue management** to make the data on GCP reliable.

Collibra was embedded directly into the **Data Engineering team** to **support GCP platform adoption** and own data confidence.

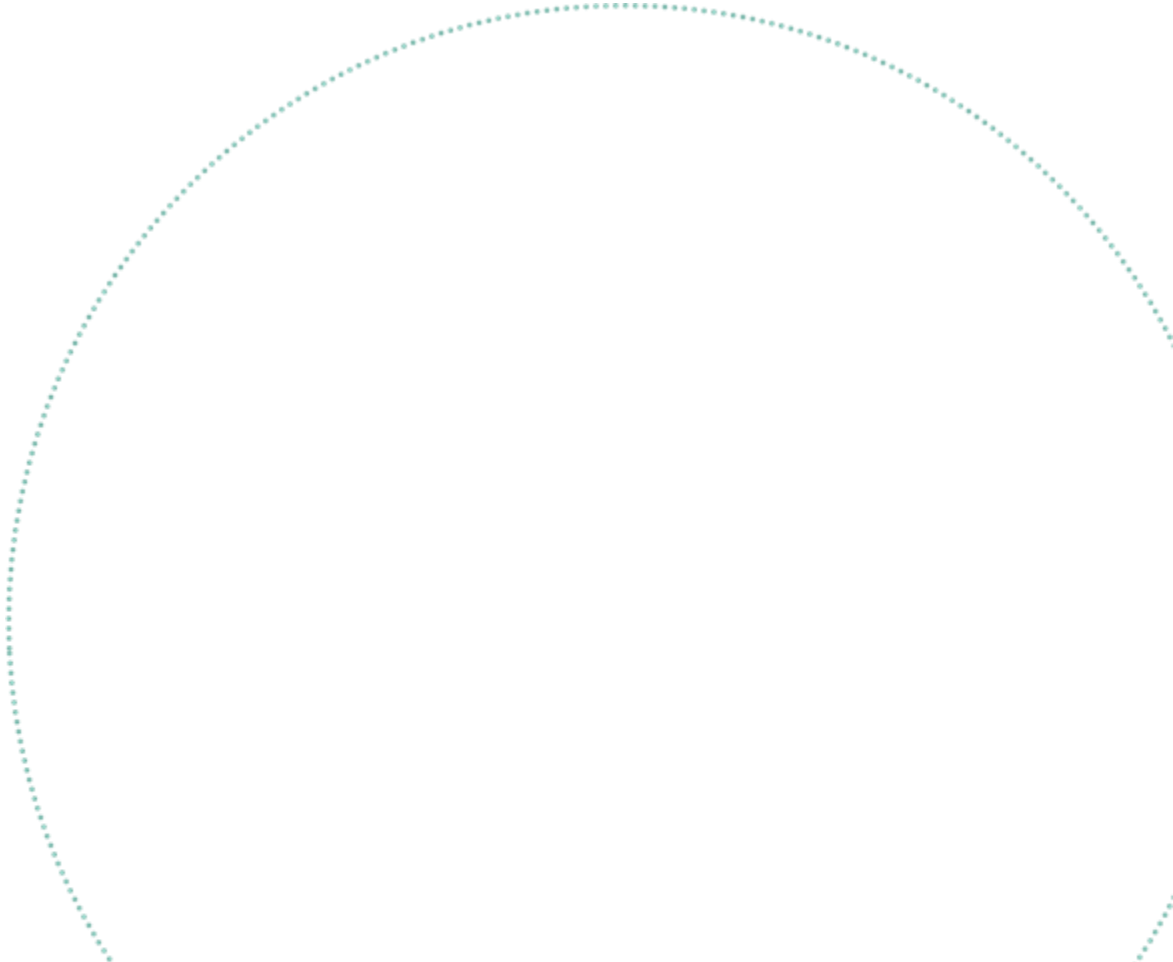

## The Outcome

Achieved the vision of a **centralized reporting layer** in BigQuery, overcoming low user adoption and lack of trust by embedding **Foundational Data Quality** checks to validate and curate 500 TB of migrated data.

Established trust by design with a clear **two-flag system (Accepted/Implemented)** in Collibra, and implemented **streamlined issue management** for immediate data quality failure resolution.

Collibra

# Thank you



Collibra

# Q&A

# APPENDIX