Collibra

www.collibra.com

61 Broadway, 31st Floor, New York, NY 10006, USA
Tel: +1 646 893 3042   Email: contact@collibra.com

MUTABLE AWARD 2024

GOLD

Collibra broadens its data governance capability to data intelligence

The company

Collibra was founded in 2008 in Brussels. The company has grown to over a thousand employees with over 700 enterprise customers across the globe, and now has offices from Canada to Australia. The company has attracted a lot of capital investment, with the last financing round in November 2021 valuing the company at $5.25 billion. Customers include Mattel, McDonald's, Heineken, Shell, Mars and Johnson Matthey. The company is privately held.

What is it?

Collibra offers a comprehensive data governance solution that has expanded in scope over the years, and it now prefers the term "data intelligence platform". Enterprises have a broad range of data assets, located on-premise and in public and private clouds, in a wide variety of applications and data sources. They need to control that range of data assets and ensure that their data is reliable, unambiguous and of high quality. Collibra offers a full-function data governance solution that can be tailored to various categories of business and technical users such as data stewards, data scientists and data owners. The product is geared towards the management of business data rather than technical metadata. A major release of the software occurred in Q2 2024,

This **Mutable Quadrant** is derived from 13 high level metrics, the more the image covers a section the better. **Execution** metrics relate to the company, **Technology** to the product, **Creativity** to both technical and business innovation and **Scale** covers the potential business and market impact.

> *This is probably the easiest integration I've ever seen happen between two tools.*
>
> Marla Dans, Head of Data Management and Governance, Chicago Trading Company
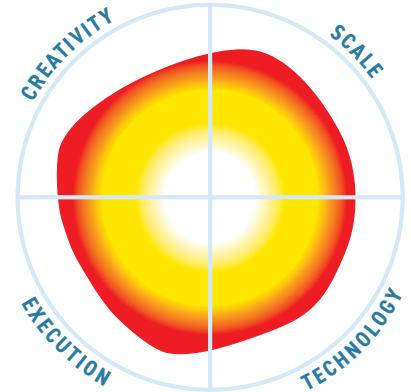
including better data observability, a revamped user interface and the ability to generate SQL from natural language statements.

What does it do?

At its heart is a data catalog and business glossary that documents the data assets of an organisation. Data can be classified and profiled, and business ownership can be assigned to the various data assets. Policies can be defined and enforced, so for example, workflows can be defined to check the authorisation of changes to a particular business hierarchy. Collibra has visual tools to assist users in navigating their data landscape, including a comprehensive data lineage solution that tracks the flow of data and can be used for impact analysis. This is the company's own intellectual property, so they no longer rely, as they once did, on a partnership with a third-party data lineage
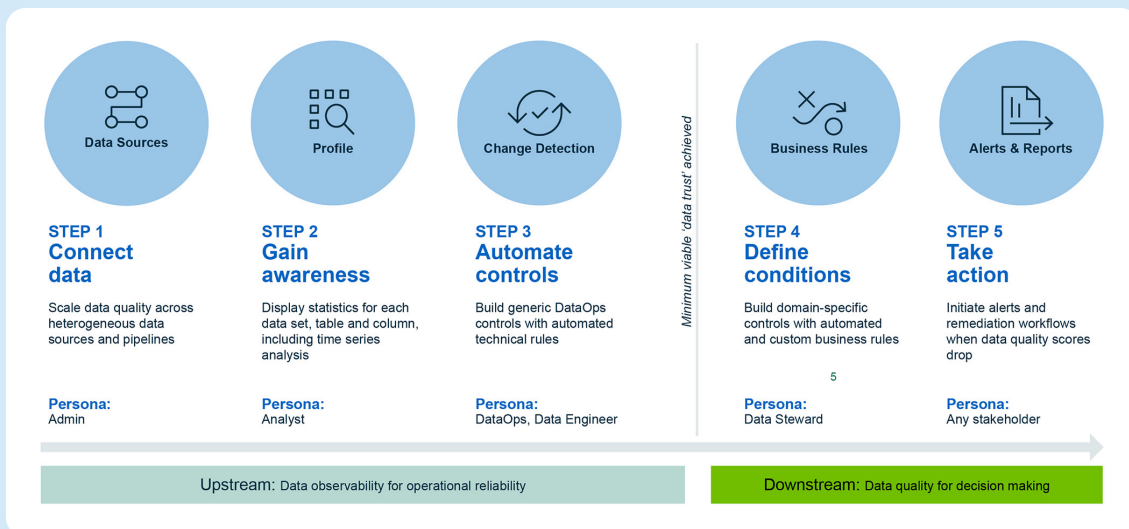
Data Sources

Profile

Change Detection

Business Rules

Alerts & Reports

Minimum viable 'data trust' achieved

STEP 1
Connect data

Scale data quality across heterogeneous data sources and pipelines

Persona:
Admin

STEP 2
Gain awareness

Display statistics for each data set, table and column, including time series analysis

Persona:
Analyst

STEP 3
Automate controls

Build generic DataOps controls with automated technical rules

Persona:
DataOps, Data Engineer

STEP 4
Define conditions

Build domain-specific controls with automated and custom business rules

Persona:
Data Steward

STEP 5
Take action

Initiate alerts and remediation workflows when data quality scores drop

Persona:
Any stakeholder

Upstream: Data observability for operational reliability

Downstream: Data quality for decision making

Figure 1 – The 5-step process to data quality and observability

specialist. There is also a "knowledge graph" to help visualise the connections between data assets as well as between data and other governance assets such as policies.

There is a search capability to allow data assets to be discovered, and this also has a natural language interface. Collibra has a reputation for having an interface that is digestible by business users, and this will go further in the next version with a comprehensive set of changes, including the ability to tailor the interface to specific data assets as well as different business roles. Collibra extends its capability to cover data access and security and overlaps to a degree with special security governance products.

Collibra's platform leverages artificial intelligence in several ways, some of which appeared in their Q2 2024 release. This included the generation of content for data asset description as well as suggesting possible description data duplication. The vendor (correctly) believes that although generative AI can be very productive, it still needs human review for generated content, as AI content can sometimes be overly imaginative (the so-called "hallucination" problem inherent in large language models). Collibra can be used to manage AI models just like other data assets and has a comprehensive AI Governance application so customers can benefit from a unified data and AI governance solution. Earlier this year, Collibra introduced AI governance, another key pillar of their platform, that enables organizations to deliver trusted AI safely and effectively.

Collibra has extensive partnerships, so for example can exchange metadata with tools such as Microsoft Purview and SAP. Interestingly, there is what seems to be a quite productive partnership with SAP and Collibra provides their joint customers end-to-end data governance across the entire data landscape including both SAP and non-SAP data. Additional bridges exist for products like Databricks and Snowflake, with over a hundred connectors provided. Collibra aims to meet users where they already are and will extend this approach with integrations to productivity tools like Slack and Teams also, enabling for example alerts in Collibra to be visible in Slack. This is the company's own intellectual property, so they no longer rely, as they once did, on a partnership with a third-party data lineage specialist.

Why should you care?

Collibra is the current market leader in data governance and is a pioneer in the space. It competes with Alation, Ataccama, Informatica and other data catalog and data governance products. It offers a fully functional data catalog and business catalog with powerful search and discovery tools, from visuals and knowledge graphs to a natural language interface. The vendor continues to attract investment, and so is well positioned to continue to plough money into research and development to continue to improve its product. The scope of the platform has increased in recent years as it has developed its own data lineage and data quality capabilities, which at one time it used to rely on a partner to provide. This is a natural development in a market that is beginning to mature, with customers often seeking broader solutions to their data management issues. The vendor needs to ensure that the acquisitions that it has made are smoothly integrated within the overall user interface and customer experience, and it is evidently investing in this area.

The bottom line

Collibra is one of the market leaders in data governance and did a great deal to popularise this field. The vendor continues to expand its customer base and invest in new and improved product areas. As such it is a natural product to be considered for major enterprise data governance needs.

Roll-up dimensions

- Provide clear insights within the Quality tab on asset pages
- Use standard DQ dimensions that your business recognizes

High-level scores

- Calculations for DQ scores summarized for end users

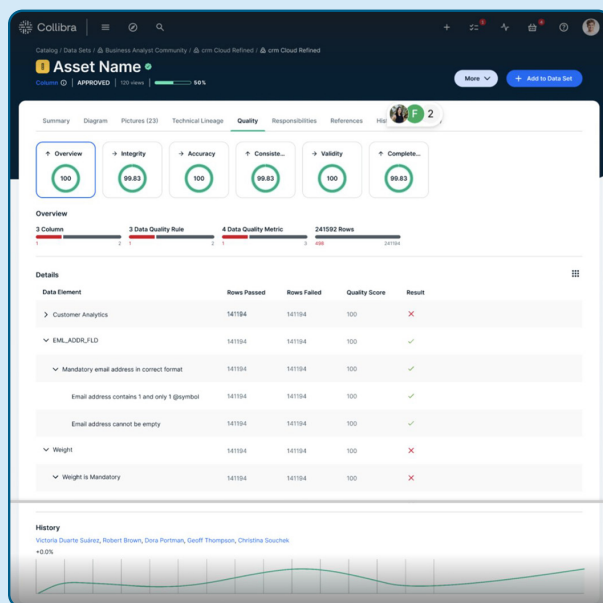### Drill-in detail

- Users have access to more detailed information when needed

Figure 2 – Simplify DQ rules for business users