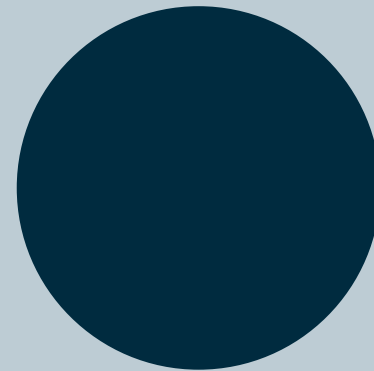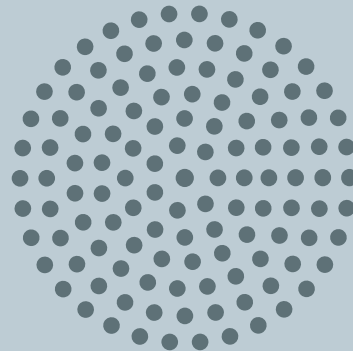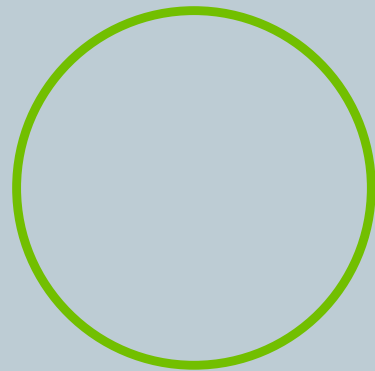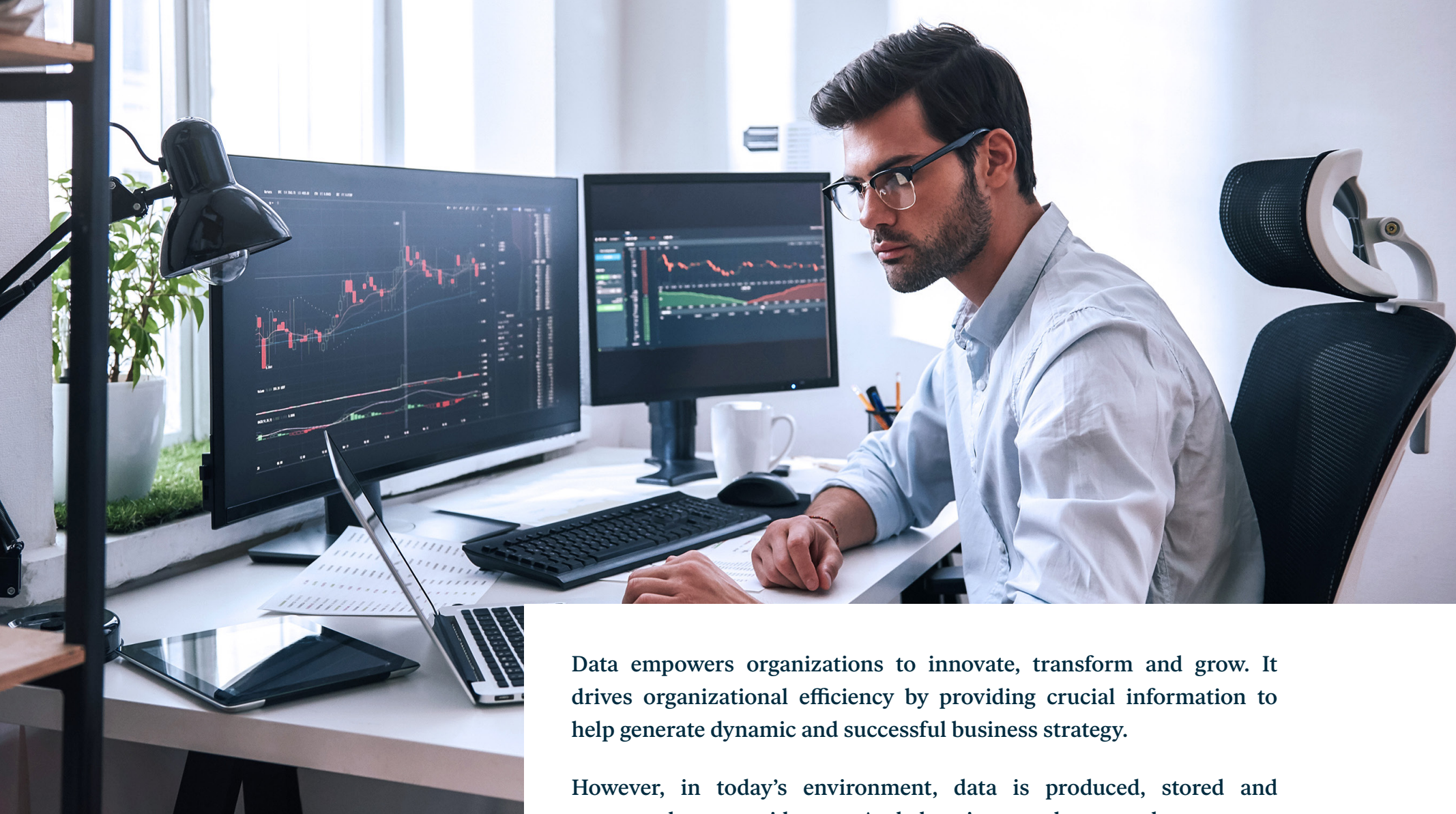Collibra

Empower your business with trusted data:

Become more data driven with an enterprise data catalog

Data empowers organizations to innovate, transform and grow. It drives organizational efficiency by providing crucial information to help generate dynamic and successful business strategy.

However, in today's environment, data is produced, stored and consumed at a rapid pace. And data is spread across departments, stored in various systems, and in many cases, no one is in charge of it. As a result, organizations struggle to organize, maintain and utilize their data effectively and efficiently.

80%

of analysts' time is spent simply discovery and preparing data

Harvard Business Review

In this modern, complex data environment, most organizations fail to unlock the value of their data. Business analysts, data scientists and other business users who use data, sometimes daily, cannot do their jobs effectively because they struggle to find and access accurate, complete and trustworthy data. As a result, they spend more time searching for data than actually using data to generate analyses and impactful insights.

And when business users finally find the data they need, they face another challenge — they aren't sure they can trust the data. They lack visibility into where the data comes from, what it means, and whether it fits their needs. And of course, bad data leads to inaccurate and incomplete analysis and reporting.

Although the inability to easily access trusted data directly impacts business users' day to day, ultimately the entire organization suffers.

It leads to

- Poor business decisions

- Slow pace of innovation

- An inability to grow and be competitive

So how do you solve this problem?

How do you combat the inability to access trusted data which is plaguing innovation within your organization?

The answer?

A data catalog!

48%

of companies achieved a better understanding of their data to drive insights and actions after implementing data catalogs.

"Machine Learning Data Catalogs Put the Entire Business in Full View", Forrester, 2019

Data catalogs help you find data that matters.

Data catalogs allow business users to quickly and easily shop for trusted data. And with trusted data, organizations can make better business decisions that drive business value and growth.

A data catalog helps you:

- Gain a unified view of all your data

- Improve trust and confidence in your data

- Increase productivity and operational efficiency

- Accelerate time to insight

Companies using a data catalog are 93% more likely to be satisfied with data quality.

The Intelligent Data Catalog: A Foundation for Analytic Excellence, Aberdeen Catalog, 2019

Gain a unified view of all your data

A data catalog gives you visibility into all of your data, regardless of where it is stored — whether it's in a data warehouse, a SaaS application, on-premises or in the cloud. It enables you to seamlessly search for and filter through all the data across your data ecosystem to find the right data for your needs.

Improve trust and confidence in your data

A data catalog helps the business have trust and confidence in your organization's data. How, you may ask? By providing full transparency into the data within your data ecosystem. A data catalog enables users to see what the data is, where it's coming from, how it's used, and better understand its quality. Being able to trust data enables you to truly unlock the value of your data and generate meaningful and trusted business insights.

Increase productivity and operational efficiency

A data catalog enables business analysts and data scientists to spend less time searching for data and reports and more time performing analyses. It also reduces duplicative and repetitive work by quickly identifying certified data and reports, so that users can focus their efforts on deriving new insights from their data.

47%

of respondents indicate that analytic consumers are having difficulty locating and accessing data content."

Dresner Data Catalog Study, Dresner, 2019

Accelerate time to insight

With a data catalog, organizations can quickly and successfully find and access data, thus facilitating faster business insights. This allows organizations to adapt to the trends of the market as they occur and spend more time innovating.

By 2021, organizations that offer a data catalog of internal and external data will realize twice the business value from their data and analytics than those organizations that do not use a data catalog.

Augmented Data Catalogs: Now an Enterprise Must-Have for Data Analytics Leaders, Gartner, 2019

Now you know WHY you should invest in a data catalog, but how do you know WHAT data catalog to invest in?

According to a survey conducted by Forrester[1], 39% of companies will implement data catalogs within the next 12 months. However, due to the abundance of data catalog choices available, organizations struggle to find the right data catalog solution.

Here are the seven must-have capabilities for data catalogs:

Ability to deploy across an enterprise

Broad metadata connectivity

Machine learning-powered automation
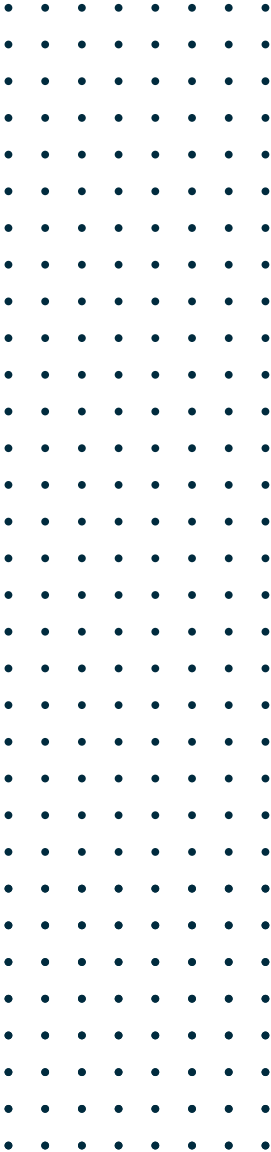
Automated data lineage

Collaboration

Embedded data governance and data privacy

Data shopping experience

1. *Machine Learning Data Catalogs Put the Entire Business in Full View*, Forrester, 2019

Ability to deploy across an enterprise

Organizations must look for a data catalog that can be strategically deployed across the enterprise, not just to support one data source or one department. Tactical data catalogs can be useful for specific situations and use cases, but they create metadata siloes. Data catalogs must fit into the larger metadata initiatives of an organization and catalog sources across the enterprise to provide the most value. They must also be able to support detailed policies, complex operating models, various organizational structures and diverse data sets — and have the ability to define roles and access permissions to ensure security and data privacy.

74%

of business analysts say that data management tools make it easier to search for existing analysis and reports that fit their needs.

Forrester Data Intelligence Report, Forrester, 2020

Broad metadata connectivity

Organizations have data spread across a huge number of data sources, from operational databases to data warehouses, data lakes and other enterprise applications. Data catalogs must be able to connect to and ingest metadata from all these data sources, creating an inventory of business-critical data sets that are easily accessible, regardless of where they are stored. Additionally, data catalogs should integrate with the BI tools of your choice, so that you can catalog reports, share insights and create a seamless workflow between your data catalog and BI tools.

Machine learning-powered automation

Due to the rapid growth in and diversity of data sources, manual organization of data is just not feasible. A data catalog with ML-powered automation capabilities reduces the impractical and onerous task of manually sorting through and organizing all data assets. Automation assists the data steward in adding context to data by automatically classifying data and suggesting links between physical data assets, logical data models and conceptual business terms. This greatly speeds up the time it takes to populate a data catalog and get it ready for use.

Automated data lineage

Organizations should look for a data catalog that provides native, automated, and integrated business and technical lineage. End-to-end data lineage provides users the necessary context about data. It helps them better understand and trust their data by allowing them to see where data comes from, how data is used, and to validate reports. Data lineage also helps organizations comply with regulations such as GDPR, CCPA and BCBS-239 and perform detailed impact analyses.

71%

of business analysts say that data lineage makes it easier to see where data has come from and how it has been changed.

Forrester Data Intelligence Report, Forrester, 2020

Collaboration

Collaboration creates trust in data by allowing all users and teams to contribute by commenting on, rating, reviewing and tagging data sets. Data catalogs with strong collaboration capabilities break down organizational silos and enable the sharing of data, knowledge and insights across an organization. What's even better is a data catalog that can automate common processes, such as data certification and data access requests, with customizable workflows — eliminating the need for back and forth emails.

Embedded data governance and data privacy

Since data catalogs are all about enabling access to data across an enterprise, it's crucial that data governance and data privacy are at the core of any data catalog. Without this, there is a risk that data can be accessed by people who shouldn't have access to it and used in a non-compliant or even an unethical manner — putting the organization's reputation in jeopardy. Organizations should look for a data catalog that provides the ability to:

- Create and maintain a single, centralized business glossary

- Establish standard governance processes

- Certify data sets

- Automatically identify and tag sensitive data

- Enforce governance and data privacy policies and access controls

Data shopping experience

To make it easy for all data consumers to find and access the data they need, a data catalog should provide a business-friendly, eCommerce-like data shopping experience. Gone are the days when the business was at the mercy of IT when it came to getting access to clean, quality data. A data catalog should allow all data consumers to quickly and easily shop for and check out data sets whenever they need – and enable them to choose the best data for their needs by surfacing insights about data quality.

64%

of business analysts say shopping for data in a central location enables effective outcomes.

Forrester Data Intelligence Report, Forrester, 2020

We have answered the WHY and the WHAT. Now we must answer the HOW.

How would an organization use a data catalog? A data catalog will help you...

- Get more value from your data and analytics investments, such as data lakes and BI tools

- Accelerate your move to the cloud

- Enable self-service analytics for the business
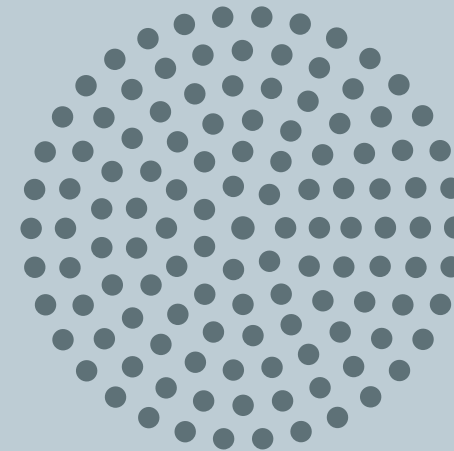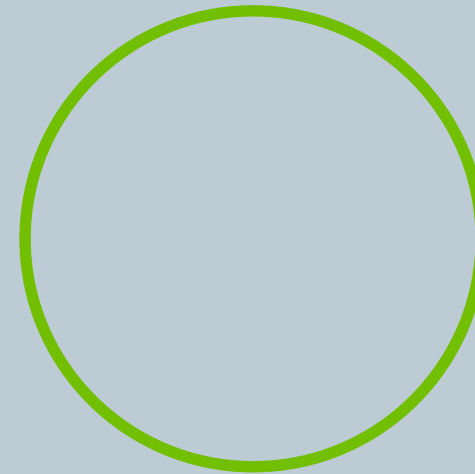
Find data that matters!

A data catalog is a powerful tool. It enables business users to discover, understand, trust and access the data they need, when they need it.

A data catalog will help your business:

- Eliminate the pain of searching through chaotic data swamps to find the right data

- Ensure confidence and trust in data

- Speed up time to insight so that you can innovate and transform faster

Data catalogs are changing the way businesses interact with their data, empowering data citizens to use data to drive business outcomes.

At Collibra, we see data catalogs as a crucial part of an organization's journey to achieving Data Intelligence and an important factor in driving revenue, improving operational efficiency, and generating innovation and growth.

If you are interested in learning more, please visit collibra.com.