

Collibra Data Governance with Cloudera Enterprise Data Hub.

A solution brief

Collibra NV VAT BE 0899.079.439

Oorlogskruisenlaan 116 (t) +32 2 894 79 60
1120 Brussels - BELGIUM (f) +32 2 706 56 76

Collibra Polska NIP PL 8971791391

Ul. Św. Antoniego 2/4, brama D (t) +48 71 780 47 24
50-073 Wrocław – POLAND

Collibra Inc TIN 80-0924168

25 Broadway, 9th Floor (t) +1 646 963 6513
New York, NY 10004 – USA (f) +1 646 786 3696

Collibra UK

1 Fore Street (t) +44 203 695 6965
London, EC24 5EJ – UK (f) +44 203 006 8844

Information-driven organizations need big data governance.

Becoming a data driven organizations is a critical to meeting today's challenges. Using an enterprise data hub with comprehensive big data governance will maximize the value of your data assets.

Big data gives us unprecedented power to produce new insight. But that insight is limited if we don't have confidence that it is the right data, of high quality, from the right places, combined in the right way and used properly. Data Governance gives you the visibility and the tools to find and make maximum use of your data. Business users need to be able to locate the data that they need to answer their questions. They must also know what the data means, where it is coming from, and who owns it in order to trust and use it. Data Scientists need to locate and manipulate the data. Organizations need to have clear agreements controlling their data sharing. Business users need to know the quality and applicability of the data. And everyone needs to be able to constantly improve the data, its quality and completeness. None of this is possible without a comprehensive data governance solution that is designed to work in the new big-data world of your enterprise information hub.

Collibra Data Governance Center: Data Governance for Business Users

In the Big Data world, where data variety is increasing exponentially, comprehensive automated data governance is more important than ever before.

Collibra Data Governance Center covers all key data governance and stewardship activities. One integrated product suite that empowers data stewards to catalog data, become the trusted data authority and get more business value from data. Collibra Data Governance Center's incredible flexibility, artificial intelligence assisted automation, and configurable operating model are tailored for the new big data world. Some data governance tools require an up-front approach with lots of labor before you get anything valuable, but Collibra enables you to build out your governance as you build your enterprise data hub, and apply it as you find new uses for your data.

Catalog Your Data

An enterprise data hub is a powerful tool, and as an engine for processing data and queries it is excellent. However, it needs a mechanism to enable business users to discover the data that they need. Without means of "shopping" for data in a way that is comprehensible to a business user and linked to the business terms he or she understands, the data stored in the data hub will not be fully utilized. Collibra provides an easy to use catalog for all your data, that uses machine learning and AI techniques to help categorize and describe all of your data. This data is linked to business terms, so that business users can identify the data that they need. It is linked to the other artifacts that use it

Collibra Data Governance Center automates all of the information collection, integration and decision making needed to govern your data. Data governance becomes a seamless part of your daily activity, and is constantly adapting and improving as your enterprise data hub

(reports, analytical models, processing jobs, etc.) giving complete lineage and change impact analysis. Data quality is easily understood through the role specific dashboards. Collibra Data Governance Center delivers the capabilities that business users need to unlock the value in your enterprise data hub.

Implement Your Data Governance Operating Model

Collibra Data Governance Center goes well beyond the collection of metadata, and enables your organization to implement a complete governance operating model. It can accommodate different parts of the organization with different ideas about using and governing data. It can also govern the entire set of data from its creation or intake into the organization, through to its use or dissemination.

ORGANIZATION - Set up your council, stewardship committee, working groups, and other structures. Collibra supports both centralized and federated organizations.

PEOPLE & PROCESS - Formalizing roles and responsibilities creates transparency and let's you automate your data stewardship processes.

SHARING AND USAGE – Manage the uses of data through data sharing agreements.

QUALITY- Get a comprehensive view of the quality of your data through data quality dashboards. Automatically or manually raise issues when quality or other data attributes are not correct.

TEMPLATES - Configure what you will govern. This could be just a subset of the data in the enterprise hub, or the entire assets of your organization. The possibilities are endless.

INTEGRATION – Collibra Data Governance Center is fully integrated with Cloudera Navigator, and comes with connectors and integration templates for many other toolsets in your landscape.

COMPLETENESS – Govern all of your data, and other assets such as analytical models, map/reduce jobs, business terms and visualizations across the enterprise hub and beyond.

SOLUTION HIGHLIGHTS

- > Integrated with all the governance information in Cloudera and Cloudera Navigator
- > Govern all types of assets including analytical models, visualizations, queries, jobs, and reports.
- > Implement governance as and where needed
- > Collaboration and workflow enable rapid decisions
- > Data Sharing Agreements insure information is used as intended
- > Integration with IT work management systems links

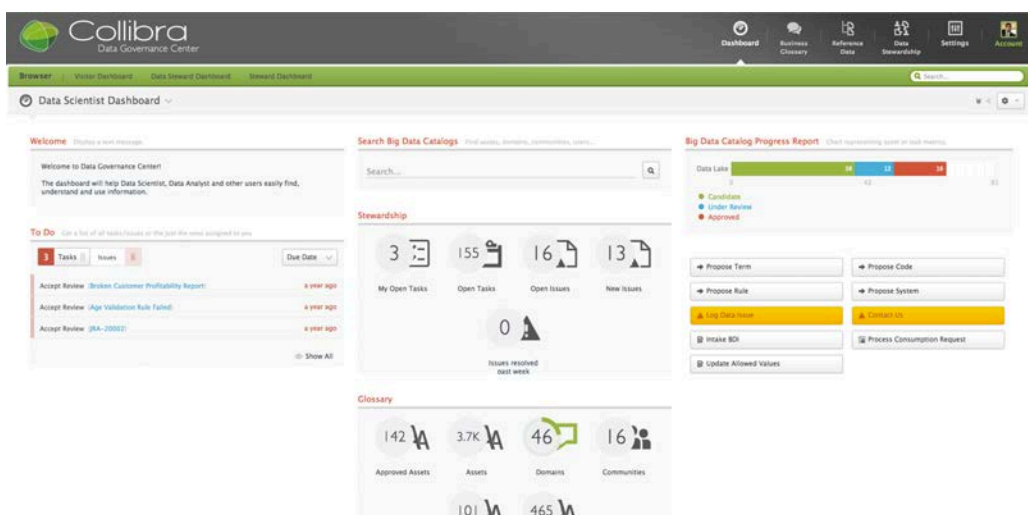
Increase & Sustain End User Adoption

The difference between success and failure of a Data Governance program is end-user adoption.

Collibra ensures adoption by delivering a consumer-like, business-focused web application. It has extensive data publication and consumption options, and integrates with productivity tools like email and Microsoft Office™. Any user can easily understand a business term and which data supports it, find data that they need for an analysis, or report issues when data does not appear to be correct. Data stewards and subject matter experts have data governance tasks delivered to them in the form that they require, so they can be handled easily as part of day-to-day activities, and eliminating manual processes and creating continuous improvement opportunities.

Track Compliance, Maturity & ROI

Measurement is the first step that leads to control and eventually to improvement. Providing key metrics to the organization on the enforcement of the policies and processes will demonstrate the value and ROI of your data governance program. Collibra Data Governance center comes with an array of fully customizable dashboards for CDO's, data scientists, data stewards, and IT professionals These dashboards are a window into the data and its quality, the completeness the maturity of their data governance



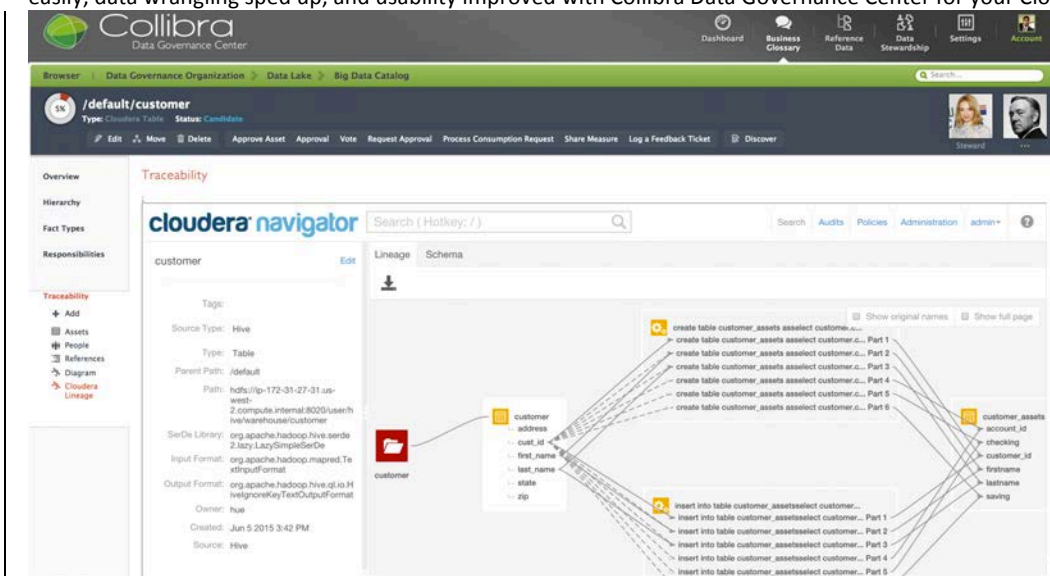
processes, and the individuals and organizations that are contributing to their success.

Benefits of the joint offering

Collibra enables you to get the most value out of all of your data. It provides the most flexible and automated data governance for your big data, and for all your data. It is fully integrated with the metadata in Cloudera and with Cloudera Navigator.

Leverage Cloudera Navigator and Collibra for Data Governance that gives your enterprise data hub the speed, quality and value it needs.

Collibra Data Governance Center implements your catalog and operating model, tailored to the needs of your business. You can govern your data from the time it is created or brought to the organization, until it is used or delivered to another party. Roles and responsibilities, processes, assets, policies, agreements are all kept up to date and linked to the appropriate people, data and technology. All of this information is connected to the metadata within Cloudera Navigator. Collibra and Cloudera Navigator together provide a seamless user experience. Issues are uncovered and addressed, data quality is improved, the right data is founds easily, data wrangling sped up, and usability improved with Collibra Data Governance Center for your Cloudera Enterprise Data Hub.



Find the data you need and Speed Data Wrangling

Organizations spend up to 75% of the elapsed time for creating analytics engaged in data wrangling activities. This process of finding appropriate data and massaging it can be dramatically sped up using Collibra Data Governance Center

- Data catalog is linked to business terms
- AI technology fills in the gaps in the overall content and metadata
- Metadata can be found and used easily
- Advanced visualizations quickly lead data scientists and BI professionals to the right data
- Data ingestion governance allows organizations to add data quickly to the data hub

Govern your unstructured data and its queries

Insures that your queries return the right data so the analytical metrics based on that data can be trusted. Data scientists, owners and users can insure that the correct data values, references and results are used.

Efficient & structured coordination between producers, consumers and data scientists to ensure all parties are

- aware of data changes that might impact results
- No more inexplicable changes / errors in analytics results
- Reduce time-consuming error analysis & resolution
- Increase quality and trust in analytics
- Increase self-service ability and understanding by all

Manage changes to your data

One of the biggest benefits of an enterprise data hub is the flexibility to use data in new and different ways. Sometimes this means changing or adding new attributes to the information you have.

- Easily analyze the data you have and initiate change requests when needed
- Structured process for changes that includes all of the business and technical stakeholders

Trust and verify your data and its use

Collibra Data governance gives you complete control and visibility into your data, its policies and its attributes.

- Ingestion control to insure that data and its lineage are known
- Complete governance of all types of assets, including analytical models, map/reduce jobs, queries, visualizations, reports and any artifact that uses the data.
- Complete visualization of any type of relationship, including lineage relationships and context.

Benefits of Collibra

- **Active connection** between Cloudera and Collibra Data Governance center.
- **Bidirectional exchange** of information and changes
- **Integrated** user experience
- **Flexible** asset types, roles, policies and processes
- **Automated** rules and tasks eliminate manual work
- **Artificial Intelligence** assistance to make work simpler
- **Complete** stewardship support
- **Dynamic** governance process

About Collibra

Collibra Corporation is the industry's only global data governance provider founded to address data management from the business stakeholder perspective. Delivered through a cloud-based or on-premise solution, Collibra is the trusted data authority that provides data stewardship, data governance and data management for the enterprise business user. Collibra automates data governance and management processes by providing business-focused applications where collaboration and ease-of-use come first. Collibra's data governance platform embraces the new requirements of big data solutions, with automation, machine learning, and the flexibility to govern data assets from source to visualization. Increase the value of your data through strong and flexible data governance. Find out more at <http://www.collibra.com>

About Cloudera

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data: The Enterprise Data Hub. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Founded in 2008, Cloudera was the first and is still today the leading provider and supporter of Hadoop for the enterprise. Cloudera also offers software for business critical data challenges including storage, access, management, analysis, security and search. With over 15,000 individuals trained, Cloudera is a leading educator of data professionals, offering the industry's broadest array of Hadoop training and certification programs. Cloudera works with over 700 hardware, software and services partners to meet customers' big data goals. Leading organizations in every industry run Cloudera in production, including finance, telecommunications, retail, internet, utilities, oil and gas, healthcare, biopharmaceuticals, networking