IDC MarketScape

IDC MarketScape: Worldwide Data Intelligence Platform Software 2024 Vendor Assessment

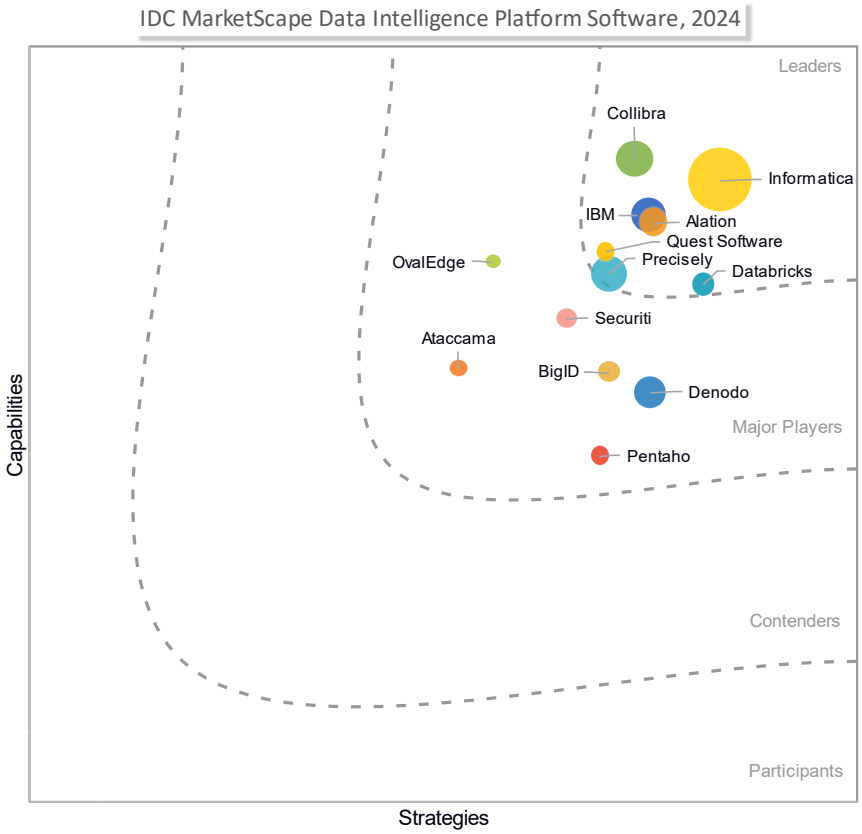
Stewart Bond Marlanna Bozicevich

THIS EXCERPT FEATURES COLLIBRA AS A LEADER

IDC MARKETSCAPE FIGURE

FIGURE 1

IDC MarketScape Worldwide Data Intelligence Platform Software Vendor Assessment



Source: IDC, 2024

November 2024, IDC #US51467224e

Please see the Appendix for detailed methodology, market definition, and scoring criteria.

ABOUT THIS EXCERPT

The content for this excerpt was taken directly from IDC MarketScape: Worldwide Data Intelligence Platform Software 2024 Vendor Assessment (Doc # US51467224).

IDC OPINION

Data intelligence, that is, intelligence about data, is a term IDC has been using since 2016 and really started to leverage it in 2018 when organizations were looking for technology to help with new regional regulations that were emerging (i.e., GDPR). As software capabilities matured and technology companies began to latch onto the data intelligence term, a new market category emerged. IDC has been tracking and forecasting the size of the data intelligence software market category since 2018, by aggregating four submarkets: data quality (DQ), master data intelligence, database life-cycle management, and metadata management inclusive of data cataloging and lineage, business glossaries, and data dictionaries.

IDC MarketScape evaluations of data catalog software published in 2020 and 2022 found that data catalogs were often bundled with complementary capabilities such as stewardship, lineage, quality, and product hubs or internal data marketplaces, clearly indicating consolidation of such capabilities into broader platforms. The result is an evolution of the IDC data catalog software evaluation into a data intelligence software platform evaluation in 2024, with a focus on data stewardship, cataloging, lineage, quality, and data product hub/marketplace functionality.

While IDC includes master data intelligence (MDI) in the data intelligence market definition, MDI is related to the broader master data management (MDM) competitive market that IDC tracks, and it would be better evaluated as part of an IDC MarketScape focused on MDM.

The data intelligence platform software market is rapidly evolving, driven by the need for AI-ready data. According to IDC's *Office of the CDO Survey* fielded in the summer of 2024, 83% of organizations have changed their data strategy since the emergence of generative AI (GenAI). That change is an increase in focus, with over half of the organizations stating that data management focus has significantly increased or is now the top focus for the organization. Objectives include supporting AI initiatives by improving the quality of data and analytics products and improving data security and privacy. Data intelligence software inclusive of data stewardship, data cataloging, data

quality, data lineage, and data product hubs (internal data marketplaces) provide key capabilities in support of AI initiatives and the need to deliver AI-ready data.

Data governance ensures compliance and security, while data cataloging and lineage provide transparency and traceability. High data quality is essential for accurate AI outcomes, and data product hubs or internal marketplaces facilitate data sharing and utilization. These capabilities are underpinning data management transformation within an enterprise intelligence architecture, where intelligence about highly distributed, diverse, dynamic, and dark data informs policy and control of data access, movement, and utilization of integrated, curated, and annotated domain-oriented data products. Data observability monitors data drift and shift in real time as data is synthesized with analytical and AI models for use in data-driven business activities and AI-fueled business.

Many of the vendors evaluated in this IDC MarketScape helped define the data intelligence software market category. There is still some interpretation of the definition as we see software vendors focused on data privacy and security leveraging data intelligence in a slightly different manner than those focused on data intelligence for analytics. There are also vendors that extend the definition to include leveraging intelligence about the data to deliver intelligence from the data, leveraging intelligent automation. These multiple interpretations are proof of the utility of data intelligence, applicable across multiple domains, applications, and use cases, from data engineering to governance and control and from analytics to machine learning and AI. Intelligence about data is critical in enabling the intelligent enterprise in modern times.

IDC MARKETSCAPE VENDOR INCLUSION CRITERIA

A critical point in this research effort is to meet the following inclusion criteria:

- The vendor offers a data intelligence software platform as one product with data catalog, data governance/stewardship, data lineage, data quality management, and optionally an internal data marketplace. At a minimum, data catalog and basic lineage capabilities need to be native intellectual property of the software vendor. Partner- or OEM-provided IP is acceptable for data governance/stewardship, advanced data lineage, data quality management, and internal data marketplace capabilities.
- The data intelligence software platform being evaluated by IDC is the strategic data intelligence solution for the vendor.
- The vendor has at least $25 million in revenue for 2023, specifically from the data intelligence platform software offering regardless of where it is deployed.

- The vendor operates in North America and recognizes a minimum of 5% of revenue from the Americas and at least one additional global region (EMEA and AP).

ADVICE FOR TECHNOLOGY BUYERS

- As with any technology procurement initiative, understanding the requirements of your organization is the most critical first step. Procuring software that is not able to meet your requirements or expectations will result in less-than-optimal adoption and possibly even a failed initiative. Challenge vendors to demonstrate effectiveness of their technology with your data, within your technology environment, applied to your data intelligence use cases.
- Vendors should be used that provide strong AI and machine learning integration for automated data discovery, classification, and quality management, as these features enhance productivity and data accuracy.
- There is a growing demand for intelligence about all data, structured and unstructured, primarily driven by the need to support generative AI use cases. The amount of unstructured data in an organization is greater than the amount of structured, and much of it is dark — meaning it hasn't been discovered, tagged, or classified. Organizations should include unstructured data types in their list of data intelligence requirements and look at vendors in the market able to support structured and unstructured data for meeting future needs of AI use cases, increasing insights, and reducing liability.
- Decision-makers should evaluate data intelligence software platforms based on support of multicloud and hybrid environments, ensuring flexibility and scalability in data management and governance across diverse data ecosystems.
- The use of data intelligence platforms with comprehensive capabilities should be prioritized, including data cataloging, governance, lineage, quality, and data product hubs in support of delivering AI-ready data to AI initiatives.
- Software vendors with robust partner ecosystems and customer success programs are critical for effective implementation, ongoing support, and maximizing the value of data intelligence investments.

VENDOR SUMMARY PROFILES

This section briefly explains IDC's key observations resulting in a vendor's position in the IDC MarketScape. While every vendor is evaluated against each of the criteria outlined in the Appendix, the description here provides a summary of each vendor's strengths and challenges.

Collibra

After a thorough evaluation of Collibra's strategies and capabilities, IDC has positioned the company in the Leaders category in this 2024 IDC MarketScape for worldwide data intelligence platform software.

Collibra's Data Intelligence Platform offers a comprehensive, integrated solution designed for enterprisewide scale and adoption. With over 100 native integrations and a vast range of innovative capabilities, the company supports data discovery, classification, and governance across various ecosystems. The platform supports seamless collaboration among business and technical users through multi-persona experiences, social collaboration features, and customizable workflows. Collibra's adaptive governance model combines data cataloging, quality, observability, and policy enforcement, ensuring compliance and traceability. The platform's AI-driven curation and data marketplace enhance data usage and value while minimizing risk and inefficiencies. The Collibra Data Intelligence Platform continues to address the evolving needs of data and AI governance.

Quick facts about Collibra are:

- **Product name:** Collibra Data Intelligence Platform
- **Company headquarters:** New York City, New York, United States, and Brussels, Belgium
- **Year founded:** 2008
- **Employees:** 1,100
- **Top industry areas:** Financial services, retail, public sector, and education
- **Cloud:** Collibra Data Intelligence Platform is available as a service on Amazon Web Services, Microsoft Azure, and Google Cloud. It can also be deployed as a single-tenant solution in a private cloud or on customer-managed infrastructure.
- **Pricing model:** Collibra is available on subscription or perpetual license terms, priced by user, compute capacity, volume of metadata, and by the number of sources scanned.
- **Interesting fact:** In addition to gathering feedback from innovation labs and formal beta testing programs — as well as alpha testing with Collibra employees, including those in its Data Office — Collibra provides an ideation portal where Collibra employees, customers, and partners can submit and vote on product ideas. In 1Q24 alone, customers submitted more than 450 ideas and provided over 3,000 endorsements, resulting in decisions on over 300 of those ideas.

Strengths

- The Collibra Data Intelligence Platform offers seamless collaboration and consumption through multi-persona experiences, enabling both business and

technical users. Social collaboration features including ratings and reviews, notifications, mentions, alerts, and a data helpdesk for managing issues surfaced by data consumers facilitate collaboration among stakeholders. Collibra's Data Notebook is a recent addition to the platform that provides users with the ability to query data as well as share SQL code, visualizations, approaches, best practices, and insights directly connected to the data.

- Collibra currently leverages AI in several places, including allowing users the ability to create data quality rules written in natural language that are then translated to SQL, parsing Power BI data models to display transformation details and complete lineage, including measures and calculated columns. Collibra is also leveraging generative AI to automatically generate asset description recommendations, improving the productivity of data stewards. Machine learning is used in data consumption to suggest similar data sets and make content recommendations based on user profiles. Collibra is leveraging its experience in data intelligence and applying it to AI governance disciplines, providing intelligence about models being used in the enterprise.
- Collibra's customer onboarding and ongoing engagement is effective, with a broad range of training opportunities, including Collibra University, a Collibra user community of 10,000+ members, and dedicated customer success representatives. Collibra also leverages in-product telemetry to understand platform and capability usage and adoption to help drive focus and innovation for features among customers.

Challenges

- A complete refresh of the Collibra Data Intelligence Cloud user interface in 2023 was necessary to improve the user experience, but it still has some work to do improving the developer experience for those organizations that need to customize the solution. A graphical user interface is available for creating tailored assert pages and building workflows, but a command-line interface is not available, and Collibra only supports the Groovy programming language for workflow customization. Collibra does offer a rich set of APIs that can be used with a broader set of programming languages when embedding Collibra capabilities into other applications.
- The ratio of indirect to direct sales is low for Collibra, which may be why it is implementing improvements to its global systems integrator onboarding program to support alignment with the broader technology partner ecosystem. Collibra expects to enable a better interconnected go-to-market strategy with global systems integrators and technology partners to grow its percentage of indirect sales and growth in emerging markets.

Consider Collibra When

The Collibra Data Intelligence Platform is a single, integrated solution architected for enterprisewide scale and adoption, not limited to one cloud, application, or integration environment. Collibra is well established in highly regulated industries where data governance is a priority as compliance is critical for business operations and reporting to authorities. Top business use cases Collibra is serving include business reporting, unified data and AI governance, and self-service data analytics through its ability to catalog and govern data assets, enable data quality, and deliver data products in an internal marketplace. Collibra is being leveraged by governance managers, privacy officers, and data consumers. Data personas using the technology include data engineers, data stewards, and data scientists. Chief data officers who need transparency into the data they are accountable for will appreciate Collibra's measurement and reporting of governance, risk, and compliance efforts via customizable dashboards, articulation scores, ratings, data maturity reports, and usage analytics.

APPENDIX

Reading an IDC MarketScape Graph

For the purposes of this analysis, IDC divided potential key measures for success into two primary categories: capabilities and strategies.

Positioning on the y-axis reflects the vendor's current capabilities and menu of services and how well aligned the vendor is to customer needs. The capabilities category focuses on the capabilities of the company and product today, here and now. Under this category, IDC analysts will look at how well a vendor is building/delivering capabilities that enable it to execute its chosen strategy in the market.

Positioning on the x-axis, or strategies axis, indicates how well the vendor's future strategy aligns with what customers will require in three to five years. The strategies category focuses on high-level decisions and underlying assumptions about offerings, customer segments, and business and go-to-market plans for the next three to five years.

The size of the individual vendor markers in the IDC MarketScape represents the market share of each individual vendor within the specific market segment being assessed. For this IDC MarketScape, vendor size was determined using IDC's 2023 Software Tracker, which represents an estimate of each vendor's 2023 software revenue. Submarkets markets include data quality, master data intelligence, metadata management, composite data framework, and dynamic data management systems.

IDC MarketScape Methodology

IDC MarketScape criteria selection, weightings, and vendor scores represent well-researched IDC judgment about the market and specific vendors. IDC analysts tailor the range of standard characteristics by which vendors are measured through structured discussions, surveys, and interviews with market leaders, participants, and end users. Market weightings are based on user interviews, buyer surveys, and the input of IDC experts in each market. IDC analysts base individual vendor scores, and ultimately vendor positions on the IDC MarketScape, on detailed surveys and interviews with the vendors, publicly available information, and end-user experiences to provide an accurate and consistent assessment of each vendor's characteristics, behavior, and capability.

Market Definition

Data intelligence leverages business, technical, relational, and operational metadata to provide transparency of data profiles, classification, quality, location, lineage, and context, enabling people, processes, and technology with trustworthy and reliable data. Data intelligence software collects, organizes, classifies, and maintains this collection of metadata that makes up data intelligence.

For this IDC MarketScape, data intelligence platforms include software with data definition, profiling, quality, lineage, cataloging, and stewardship (governance) capabilities.

LEARN MORE

Related Research

- *AI-Ready Data: Foundation for AI-Fueled Business* (IDC #US52607724, September 2024)
- *Architecting the AI-Fueled Business, 2024: Effective AI Adoption Requires a Business Operational Plan and a New Technology Operating Model* (IDC #US52576424, September 2024)
- *IDC Survey Spotlight: Has Generative AI Impacted Data Management?* (IDC #US52535524, September 2024)
- *Making the Case: Data Governance for GenAI — Applying New Data Value Chains in the Enterprise* (IDC #US52520924, August 2024)
- *Worldwide Data Integration and Intelligence Software Market Shares, 2023: Hydrating AI* (IDC #US51712324, June 2024)

Synopsis

This IDC study evaluates the evolving data intelligence software market, driven by needs for AI-ready data. It highlights 13 key vendors — Alation, Ataccama, BigID, Collibra, Databricks, Denodo, IBM, Informatica, OvalEdge, Pentaho, Precisely, Quest Software, and Securiti — assessing their strengths, challenges, and suitability for various data management and AI initiatives. The document underscores the importance of data governance, quality, lineage, and cataloging in supporting AI and data-driven business activities.

“Data intelligence platform software is key to helping organizations deliver AI-ready data products for synthesis and utilization in business activities,“ says Stewart Bond, vice president, Data Intelligence and Integration Software research at IDC. “Each vendor in this evaluation is doing something unique in terms of bringing AI to data using AI automation in the software, improving the productivity and outcomes of data workers in making data ready for AI initiatives in the organization to improve business outcomes.”

ABOUT IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology, IT benchmarking and sourcing, and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives. Founded in 1964, IDC is a wholly owned subsidiary of International Data Group (IDG, Inc.).

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com