

## Market Guide for Active Metadata Management

Published 14 November 2022 - ID G00756612 - 14 min read

By Analyst(s): Guido De Simoni, Mark Beyer

Initiatives: [Data Management Solutions](#)

Active metadata management is a set of capabilities across multiple data management markets, led primarily by recent advancements in how metadata can be used for continuous analysis. Data and analytics leaders must consider the market evolution as transformational in all data-enabling technologies.

### Overview

#### Key Findings

- Organizations are seeking new approaches to automate and optimize; organizations are demanding “active metadata” to assure augmented data management capabilities.
- Existing metadata management tools are increasingly incapable of fulfilling comprehensive metadata needs in the enterprise.
- Adjacent data management platforms such as databases, data integration, data quality and data governance tools have consistently increased their capabilities relative to accessing and managing external metadata sources.
- Metadata capabilities embedded in most data management solutions must expand to provide metadata analysis for all types of metadata categories from any platform, or that function will be replaced by embeddable, advanced metadata functionality from mature metadata solutions.
- There is a lack of common metadata standards, which makes metadata sharing and interoperability a major challenge across multiple metadata management solutions in the market.

#### Recommendations

Data and analytics leaders investing in data management solutions must:

- Adopt metadata management solutions that can create prescriptive, parameterized recommendations to alter design inputs by exploiting adjacent data management tools for operations such data observability tools.
- Assess current data management tools and platforms by evaluating their capability to share internal metadata to support broader platform-to-platform orchestration.
- Start capturing “runtime metadata” like data usage, data affinity and user behaviors to exploit metadata value by automation.
- Identify tools and platforms that can export their own metadata and import “foreign” metadata relative to existing metadata, processing instructions and optimization strategies.
- Enable support for automated system changes in adjacent data management systems by leveraging tools that support metadata analytics workflow management capabilities, including collaborative design capabilities.

## Strategic Planning Assumption

- Through 2024, organizations that adopt aggressive metadata analysis across their complete data management environment will decrease time to delivery of new data assets to users by as much as 70%.

## Market Definition

The active metadata market offers a broad spectrum of data management tools (see Note 1) that have specifically enabled significant active metadata use within their platform. Active metadata management is a set of capabilities that enables continuous access and processing of metadata that support ongoing analysis over a different spectrum of maturity, use cases and vendor solutions. Active metadata outputs range from design recommendations based upon execution results and reports of runtime steps through, and indicators of, business outcomes achieved. The resulting recommendations from those analytics are issued as design inputs to humans or system-level instructions that are expected to have a response.

## Market Description

Active metadata management can be identified in conjunction with metadata management technology maturity. Active metadata is the continuous analysis of all available users, data management, systems/infrastructure and data governance experience reports to determine the alignment and exception cases between data as designed versus actual experience.

Active metadata management includes the capability of operationalizing these analytic outputs in the form of operational alerts and generated recommendations. It identifies the nature and extent of patterns in data operations, ultimately resulting in AI-assisted reconfiguration of data itself and operations that use that data in active metadata utilization.

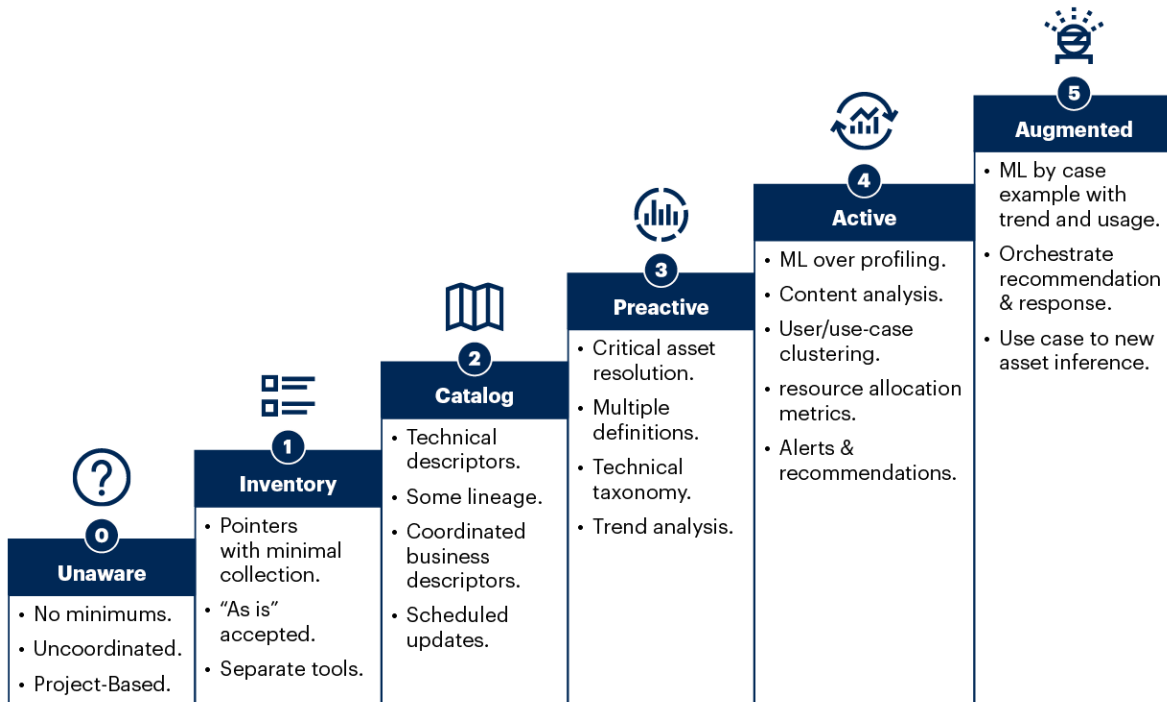
Capabilities that are essential in this market space are:

- **ML over profiling:** A technique used to analyze and gain a better understanding of raw data. It is the first step in determining what insights data can yield when you run it through machine learning algorithms to make predictions. This is an advanced metadata step that is proactive and active (Maturity Levels 3 and 4 on the maturity model [see Figure 1 and Note 2]). It's early in the active metadata lineup, but it could be delayed in preference to analyzing the passive metadata as a graph compared to use-case and utilization graphs.
- **Content analysis:** A technique used to make replicable and valid inferences by interpreting and coding textual material. By systematically evaluating text (e.g., documents, oral communication and graphics), qualitative data can be converted into actionable quantitative data. Possibly start using externally developed content analytics at Maturity Level 3 (maybe even 2) but mandatory at Maturity Level 4. This capability includes coordination between data quality, master data management and data governance functions.
- **User/use-case clustering:** Includes grouping together users with similar viewing patterns to recommend similar content, and anomaly detection (fraud detection, detecting defective mechanical parts, data outliers). This is proactive or Maturity Level 3. This is the gateway to combining DataOps with active metadata and critical to maintaining a data fabric, whether done manually or with automation (see [Quick Answer: What Is Data Fabric Design?](#)).

- **Resource allocation metrics:** Resource allocation may be decided by using computer programs applied to a specific domain to automatically and dynamically distribute resources (for example, storage or protoapplicants). This is preactive or Maturity Level 3. This is the gateway to combining DataOps with active metadata and critical to maintaining a data fabric, whether done manually or with automation (see [Quick Answer: What Is Data Fabric Design?](#)).
- **Alerts and recommendations:** List alerts (recommendations) metadata information based on value of identifier parameter.
- **ML by case example with trend, with usage:** Practically all data management and data utilization software solutions track users, processes and data assets. Patterns of usage form a graph pattern, and this implies a social graph of users and use cases, which can be used as a form of crowdsourcing to teach an enterprise how to better use data. Additionally, these same patterns can be used to identify data and analysts relative to the outcomes they achieve and the data they are using.
- **Orchestrate recommendation and response:** In many ways, this follows the model above for alerts and recommendations. The additional capability here is registering and accessing the design and runtime metadata from adjacent data management platforms (data integration, data quality [DQ], master data management [MDM], data governance, enterprise integration platform as a service and cloud DBMS) to compare it in the data fabric use case through the alignment and exception engine.
- **Use case to new asset inference:** Focus on metadata inference and discovery as well as continuous comparisons to passive metadata. This creates a more complete and accurate description of data assets than relying on passive metadata alone. It is likely first encountered experimentally at Maturity Level 3, but becomes operational at Level 4, and exists in identified and targeted trends and use-case patterns in Level 5 (see Figure 1).

Figure 1: Metadata Management Technology Maturity

**Metadata Management Technology Maturity**



Source: Gartner  
736613\_C

**Market Direction**

The market is shifting in relation to two major dimensions: market adoption and the business value of metadata sharing.

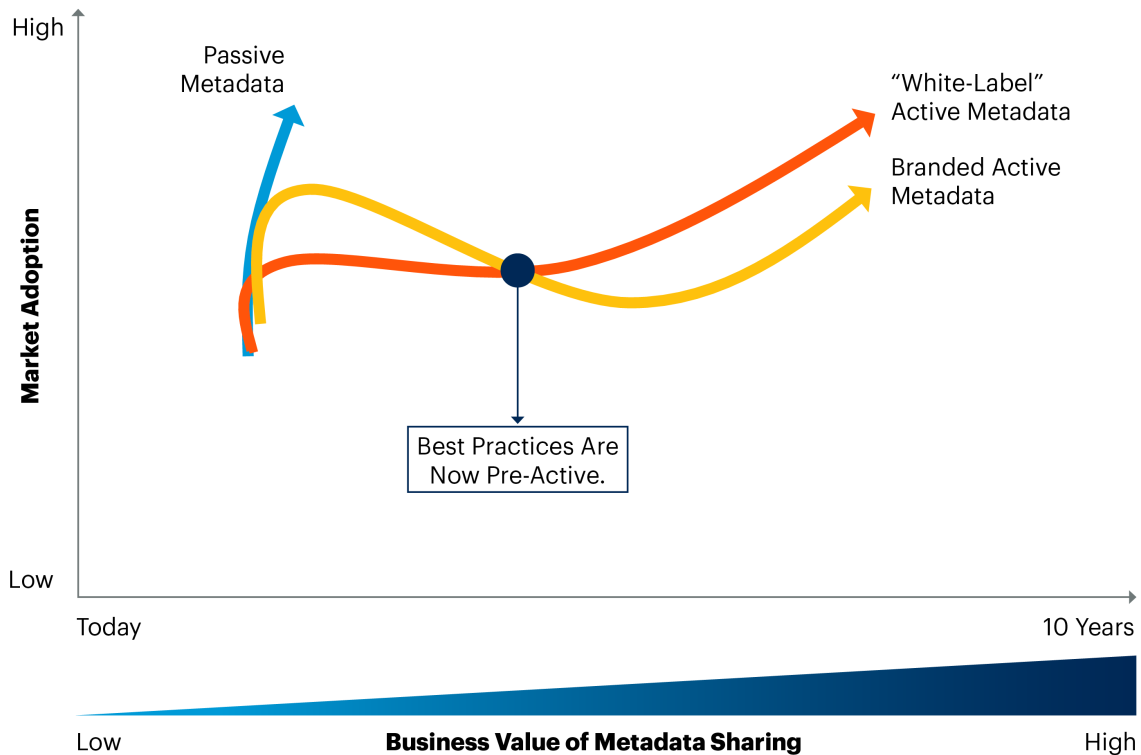
Passive metadata will continue to drive growth as current metadata management capabilities continue to drive maturity in metadata management practices and investment in technology. Overall, the metadata management software market grew at 21.6%, reaching \$1.54 billion in U.S. dollars. This is one of the highest growing markets within data management software overall, following the DBMS market growth of 22%, although from a much smaller revenue base. The top five vendors by revenue (IBM, Informatica, Collibra, SAP and Microsoft) collectively grew their metadata management software revenue by 25.7%, marginally above total market growth. These top five vendors have an aggregated market share of 45%, and since they've grown at near the same speed as the total market, this share has been fairly stable over the last three years. Smaller vendors like data.world, Smartlogic (acquired by MarkLogic), Solidatus and Alex Solutions continue to move the market forward with above-market growth (see [Market Share: Metadata Management Software, Worldwide, 2021](#)).

A “white-label product” is a product or service produced by one company that other companies rebrand as part of their own solution. In this context, active metadata management (practices and techniques) is leveraged within the evolution of established metadata management solutions. This market adoption will evolve at a slow pace compared to the established passive metadata-driven market but is expected to accelerate over the next five years, driven by the maturity of data fabric design and the realization of shared metadata value.

However, “branded active metadata management” is going to show significant adoption from the early hype generated by data fabric. As the data fabric architecture matures and shows success, branded active metadata management will be consumed by established metadata management players that drive all aspects of metadata (passive and active) management and contribute to an overall growth in alignment to the data management market's transformation (for example, DQ, MDM and data integration). Figure 2 describes the different directions.

Figure 2: Market Shift

**Market Shift**



Source: Gartner  
756612\_C

**Gartner.**

Figure 2 shows how the business value of metadata sharing will improve with the adoption of active metadata. First, leveraging passive metadata, the focus will be on branded active metadata; in the long term, on white label active metadata.

**Market Analysis**

In the first six months of 2022, interest in active metadata grew more than 100% compared to the same period in 2021 (based on Gartner inquiries). Existing purpose-built metadata management tools and solutions are increasingly challenged as the primary metadata asset in the enterprise by adjacent data management platforms such as databases, data integration, data quality and data governance tools. As a result, the metadata management capabilities required in the enterprise will be distributed across many markets. The stand-alone metadata management platform will be refocused from augmented data catalogs to a metadata “anywhere” orchestration platform.

This means:

---

*Metadata management is becoming a critical functionality in practically all data-enabling technologies and metadata analytics, augmented and automated design, and even deployment of data management platforms.*

---

Providers are tending toward becoming metadata analysis and orchestration platforms. Their new competition will be any data management platform that can absorb and refactor metadata assets in a highly flexible and adaptive manner using metadata-driven access and delivery. Nearly every data management market is undergoing a massive, user-driven disruption. Data management has begun a shift from managing the data content to focusing on metadata. At the same time, the concept of metadata platforms and tools serving as metadata collection and storage solutions is being challenged in terms of both their completeness and utility.

The broad demand and distribution of metadata capabilities is demonstrating that metadata utilization to support data management, operational applications and analytics is more important than ever. The shift in demand and expectations for even more capabilities means that any network management, infrastructure administration solution, software tool or application platform that lacks significant metadata capability will be rendered archaic and eventually retired.

Importantly, metadata platforms themselves will evolve to become coordinating partners with other data management tools. The very best metadata management offerings will be able to present their capabilities as technology partners or as embeddable and coordinated services instead of becoming diffused or diminished. They will contribute to the composable business by relieving other data management tools from localized metadata burdens and simultaneously enabling standardization of metadata management, utilization and analysis. This is the focus of delivering on the data fabric (see [Emerging Technologies: Critical Insights on Data Fabric](#)).

The new requirements for delivering metadata capabilities are distributed across a primary set of providers. These providers are continuing to offer metadata search, identification, cataloging and metadata analysis tools. Additionally, they will incorporate a diffused capability across existing data-management-capable platforms or tools. Metadata management solutions have new competitors that are not even identified as metadata management solutions because they use metadata in inordinately powerful ways in their own platform.



---

*Increasingly, the metadata-focused solutions will harvest, analyze and evaluate metadata internally, but more importantly, they will begin orchestrating metadata processes across a broad spectrum of other data management tools.*

---

At the same time, existing data management, application development and analytics platforms or tools will evolve. Their functions and features will be increasingly capable and reliant upon metadata-driven instructions from other tools, even replacing their internal, direct-control processes; in short, metadata-meets-packaged-metadata capabilities.

Metadata management itself will begin to separate into patterns for design, implementation and orchestrated delivery. Market drivers and demand for change include:

- **Openness will become mandatory.** Existing metadata standards are often varied and incomplete when compared to each other. Current metadata market offerings may provide a good opportunity to teach what is needed for broader standards and specific “forks” in those standards.
- **Metadata ownership transfers from the CIO to the chief data officer (CDO)** and line-of-business teams will increase the requirements for layers of metadata to enhance broad comprehension of data management for less technical leaders.
- **Extent and variety of metadata are expanding** well beyond lineage, content and glossaries, requiring an ever-increasing level of automated analysis of users, use cases, usage and utilization of familiar and unfamiliar data assets.
- **Machine-to-machine capabilities** to enhance the scope of metadata through automation (machine learning) and through automated enrichment by semantic search capabilities, standard processes and crowdsourcing.
- **Semantic formalism** (also known as formal ontologies such as Financial Industry Business Ontology [FIBO], Manufacturing’s Semantics Ontology [MASON], etc.) to improve interoperability.
- **Taxonomic resolution** to develop shared understanding of data assets across multiple domains.
- **Collaborative utilization** will require new ways to capture and visualize metadata (driven by data preparation for analytics). Included is the capability of rating, ranking, tagging of data and ability to communicate within the metadata solutions.

- The crowdsourcing capability of multiclient, multi-use-case solutions will enable pattern detection for any affinity between abstract classes of data and logical arguments or statistical operators – even across horizontal business functions and industry verticals.

## Representative Vendors

*The vendors listed in this Market Guide do not imply an exhaustive list. This section is intended to provide more understanding of the market and its offerings.*

### Market Introduction

The list of vendors is not exhaustive, and it represents vendors that leverage both passive and active metadata management at different levels of maturity (see Table 1). These vendors are not identified as branded active metadata management solutions.

---

*Further details on these providers can be found in the Gartner research [Tool: Vendor Identification for Active Metadata Management](#).*

---

**Table 1: Representative Vendors in Active Metadata Management**  
 (Enlarged table in Appendix)

Vendor	Headquarters Location	Product/Solution Name(s)
Alation	United States	Alation Data Catalog, Alation Data Governance App, Alation Cloud Service, Alation Cloud Service for Snowflake
Alex Solutions	Australia	Alex Augmented Data Catalog, Alex Intelligent Connectors
Alteryx	United States	Alteryx Intelligence Suite, Alteryx Designer
Ataccama	Canada	Data Governance Suite. Ataccama supports version 11 and above. Next version of Ataccama ONE will be available in 1Q23
Atlan	United States	Atlan
BigID	United States	BigID Data Intelligence Platform, SmallID
Cinchy	Canada	Cinchy Dataware Platform
Claudin	Denmark	Claudin
Collibra	United States	Collibra Data Intelligence Cloud
data.world	United States	data.world Enterprise
Datactics	United Kingdom	Datactics Platform, Legal Entity Identifier (LEI) Match Engine
DataGalaxy	France	Data Knowledge Workplace
Denodo	United States	Denodo Platform
DTA Healthcare Solutions	United States	Compendium Data Catalog
Global Data Excellence	Switzerland	Data Excellence Management System (DEMS)
Global IDs	United States	Data Ecosystem Evolution Platform (DEEP)
IBM	United States	IBM Cloud Pak for Data, IBM InfoSphere Information Governance Catalog, IBM InfoSphere Information Analyzer, IBM Industry Models, IBM Data Observability by Databand
Ilumex	Israel	Ilumex
Informatica	United States	Intelligent Data Management Cloud
Intenda	South Africa	Fraxses – Chameleon
Inron	Italy	Inron EDM Data Platform, Inron EDM Application Backbone
Ketch	United States	Asset Manager – Data Discovery – Ketch Data Control Platform, Data Classification – Ketch Data Control Platform
MANTA	United States	MANTA
MarkLogic	United States	MarkLogic Server, MarkLogic Data Hub, MarkLogic Semaphore
Modak	United States	Modak Nabu
The Modern Data Company	United States	DataOS
Nexla	United States	Nexsets, Data Integration with Nexla
Octopai	Israel	Data Lineage XD, Data Discovery, Data Catalog
OneTrust	United States	Data Discovery, Data Catalog, Data Mapping Automation, Assessment Automation
Orion Governance	United States	Orion Enterprise Information Intelligence Graph (EIG)
Precisely	United States	Precisely Data360 Govern
Qlik	United States	Qlik Catalog, Qlik Sense Enterprise SaaS
Quest Software	United States	erwin Data Intelligence by Quest, erwin Data Catalog by Quest (also included as part of erwin Data Intelligence by Quest)
SAP	Germany	SAP Data Intelligence Cloud – Cloud, SAP Information Steward – OP, SAP PowerDesigner – OP
Semantic Web Company	Austria	PoolParty Semantic Suite, PowerTagging for SharePoint and Microsoft 365
Solidatus	United Kingdom	Solidatus
Stratio	Spain	Stratio Augmented Data Fabric – Scorpio
TopQuadrant	United States	TopBraid EDG
Zeenea	France	Zeenea Data Discovery Platform, Zeenea Studio and Zeenea Explorer, Zeenea's SOC 2, Type II

Source: Gartner (October 2022)

## Market Recommendations

- Assess current data management tools and platforms by evaluating their capability to share internal metadata to support broader platform-to-platform orchestration.
- Replace systems that share only design-based metadata by identifying tools and platforms that can export their own metadata and import “foreign” metadata relative to processing instructions and optimization strategies.
- Select offerings that accept parameterized instructions to alter design inputs such as job flow, resource allocation, and manage precedent/dependent processing calls to third-party tools.
- Enable support for automated system changes in adjacent data management systems with tools that support metadata analytics workflow management capabilities, including collaborative design capabilities.

## Note 1: Representative Vendor Selection

The list of vendors is not exhaustive, and it represents vendors that leverage both passive and active metadata management at different levels of maturity. These vendors are not identified as branded active metadata management solutions.

## Note 2: Metadata Management Maturity Model

The metadata management technology maturity model is a Gartner original maturity assessment aimed at identifying the application of metadata management capabilities over technology maturity. The maturity model is fully explained in [Leverage the Metadata Management Maturity Model to Support Augmented Data Management](#).

## Document Revision History

[Market Guide for Active Metadata Management - 27 July 2021](#)

---

## Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

[Hype Cycle for Data Management, 2022](#)

---

© 2022 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by [Gartner's Usage Policy](#). Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "[Guiding Principles on Independence and Objectivity](#)."

Table 1: Representative Vendors in Active Metadata Management

Vendor	Headquarters Location	Product/Solution Names (s)
<a href="#">Alation</a>	United States	Alation Data Catalog, Alation Data Governance App, Alation Cloud Service, Alation Cloud Service for Snowflake
<a href="#">Alex Solutions</a>	Australia	Alex Augmented Data Catalog, Alex Intelligent Connectors
<a href="#">Alteryx</a>	United States	Alteryx Intelligence Suite, Alteryx Designer
<a href="#">Ataccama</a>	Canada	Data Governance Suite. Ataccama supports version 11 and above. Next version of Ataccama ONE will be available in 1Q23
<a href="#">Atlan</a>	United States	Atlan
<a href="#">BigID</a>	United States	BigID Data Intelligence Platform, SmallID
<a href="#">Cinchy</a>	Canada	Cinchy Dataware Platform
<a href="#">CluedIn</a>	Denmark	CluedIn
<a href="#">Collibra</a>	United States	Collibra Data Intelligence Cloud
<a href="#">data.world</a>	United States	data.world Enterprise
<a href="#">Dataactics</a>	United Kingdom	Dataactics Platform, Legal Entity Identifier (LEI) Match Engine

DataGalaxy	France	Data Knowledge Workplace
Denodo	United States	Denodo Platform
DTA Healthcare Solutions	United States	Compendium Data Catalog
Global Data Excellence	Switzerland	Data Excellence Management System (DEMS)
Global IDs	United States	Data Ecosystem Evolution Platform (DEEP)
IBM	United States	IBM Cloud Pak for Data, IBM InfoSphere Information Governance Catalog, IBM InfoSphere Information Analyzer, IBM Industry Models, IBM Data Observability by Databand
illumex	Israel	illumex
Informatica	United States	Intelligent Data Management Cloud
Intenda	South Africa	Fraxses – Chameleon
Irion	Italy	Irion EDM Data Platform, Irion EDM Application Backbone
Ketch	United States	Asset Manager – Data Discovery – Ketch Data Control Platform, Data Classification – Ketch Data Control Platform
MANTA	United States	MANTA
MarkLogic	United States	MarkLogic Server, MarkLogic Data Hub, MarkLogic Semaphore
Modak	United States	Modak Nabu

The Modern Data Company	United States	DataOS
Nexla	United States	Nexsets, Data Integration with Nexla
Octopai	Israel	Data Lineage XD, Data Discovery, Data Catalog
OneTrust	United States	Data Discovery, Data Catalog, Data Mapping Automation, Assessment Automation
Orion Governance	United States	Orion Enterprise Information Intelligence Graph (EIIG)
Precisely	United States	Precisely Data360 Govern
Qlik	United States	Qlik Catalog, Qlik Sense Enterprise SaaS
Quest Software	United States	erwin Data Intelligence by Quest, erwin Data Catalog by Quest (also included as part of erwin Data Intelligence by Quest)
SAP	Germany	SAP Data Intelligence Cloud – Cloud, SAP Information Steward – OP, SAP PowerDesigner – OP
Semantic Web Company	Austria	PoolParty Semantic Suite, PowerTagging for SharePoint and Microsoft 365
Solidatus	United Kingdom	Solidatus
Stratio	Spain	Stratio Augmented Data Fabric – Scorpio
TopQuadrant	United States	TopBraid EDG



Zeenea

France

Zeenea Data Discovery Platform: Zeenea Studio and Zeenea Explorer. Zeenea's SOC 2 Type II

Source: Gartner (October 2022)