

Executive View Informatica Cloud Data Governance and Catalog

Informatica Cloud Data Governance and Catalog is a comprehensive solution for managing data and its metadata in a SaaS solution, providing the data catalog and data governance across all corporate data, regardless of where the data resides. It utilizes the AI capabilities of Informatica CLAIRE and is a feature-rich solution, fully integrated into the Informatica Intelligent Data Management Cloud.



By **Martin Kuppinger**
mk@kuppingercole.com

Content

1 Introduction	3
2 Product Description	5
3 Strengths and Challenges	8
4 Related Research	10
Content of Figures	11
Copyright	12

1 Introduction

Data is essential to the Digital Business. It fuels business decisions at all levels, from the strategic to the operational. Data is essential in marketing automation. It is essential for automating manufacturing processes. It is essential for everything around AI (Artificial Intelligence), given that AI and the related ML (Machine Learning) build on large amounts of data.

With the evolution of the past years, the need for managing data has changed. Relational SQL databases or, even earlier, hierarchical databases, have shifted away from the center of attention. Cloud databases of many kinds are common today, with large CSPs (Cloud Service Providers) such as AWS (Amazon Web Services) or Microsoft Azure providing not just one, but sometimes more than 10 kinds of databases. Big data approaches and analytics solutions have become the focus of attention.

However, with the immense growth of both data and data stores, and the multitude of technical solutions, another challenge became apparent: Data that is not known to the users and data that cannot be accessed is of little value. Furthermore, with regulations such as the EU GDPR (General Data Protection Regulation), the need for knowing where certain type of data such as PII (Personally Identifiable Information) resides, has become essential. From a business perspective, though, it is important to note that this is not just a regulatory mandate, but a business mandate. Data can only be used when it is known, and data can only be protected when it is known.

Another trend over the past few years has been what some call "data democratization". Behind that is the fact that there is a need for and concrete use of data by more people than ever before. Making data available helps people at all levels in the business to use that data for their job.

The broadened and the new use cases around data also require new categories of solutions: Data Catalogs, Metadata Management, and Data Governance. These three terms are closely related, increasingly resulting in integrated solutions.

- Metadata Management refers to solutions that enable organizations to manage data across a range of systems, maintaining metadata of the data. This also includes capabilities such as data lineage, i.e., analyzing and documenting the flow of data between various systems.
- Data Catalogs are where this metadata is stored and managed. A data catalog is the central repository that provides a view on the enterprise data across all managed data stores. It enables the use of such data and delivers the capabilities for "data democratization".
- Data Governance builds on Metadata Management and enables control of the use and flow of data.

For businesses, it is becoming increasingly important to have solutions for Data Governance, Data Catalogs, and Metadata Management in place. Only then will organizations succeed in utilizing the potential that is in the data they have and collect. Only then will they succeed in having the data on hand that is needed for business applications, decision support, automation, and AI/ML. And only then, will organizations be able to control the sprawl of data, with new cloud data stores, new analytical solutions, and AI.

Management of cloud data stores; governance of data and fulfilment of regulatory compliance requirements; governance and explainability of AI/ML models; a common understanding of data and its usage; and the efficient utilization of data: All that will only work with strong data management in place. Data Catalogs and Data Governance are the cornerstones of the modern, digital business.

2 Product Description

Informatica is a leader in the data management space. They offer a broad range of solutions for all areas of Data Management, including Data Catalogs, Data Lineage, Data Quality, and Data Governance.

The newest addition to this portfolio, and a logical step following the introduction of the Intelligent Data Management Cloud by Informatica, is their Cloud Data Governance and Data Catalog solutions. Informatica is already present in both areas with their on-premises solutions, and takes a market-leading role in this space. However, Informatica Cloud Data Governance and Data Catalog is not just a lift-and-shift of their existing technology, but a cloud-native SaaS offering. As with the Intelligent Data Management Cloud, Informatica has invested in leveraging the broad set of capabilities they already offer, but extended these and delivered them tailor-made to the specifics of SaaS services.

Cloud Data Governance and Catalog is part of the Informatica Intelligent Data Management Cloud, thus tightly integrating with the broad range of other services therein. Pricing is, as with the other services, consumption-based.

The goal is to deliver a comprehensive, cloud-based solution for managing metadata, Data Governance, and the Data Catalog, that leverages Informatica CLAIRE, its AI/ML engine, to deliver intelligence and automation. Another important capability of this new offering is the integrated approach to governance of data and AI models.

As already mentioned above, Informatica Cloud Data Governance and Catalog is not just a revamp of the on-premises solutions Informatica offers in this area, but a cloud-native solution. It operates using a serverless model, i.e., in a deployment model that is built for scale and does not require prior deployment of servers. Being serverless also allows for massive scaling of the solution. While being provided as SaaS solution, it can catalog and govern data in a multi-cloud environment, i.e., across multiple clouds, and integrated with data sources both in the cloud and on-premises. It is integrated with the various other modules of the Informatica Intelligent Data Management Cloud, and it comes with a comprehensive set of APIs (Application Programming Interfaces) to access the capabilities of the solution programmatically. Last, but not least, it is not only built to manage and govern AI models, but also utilizes AI to increase end-user productivity, e.g., in discovering useful data.

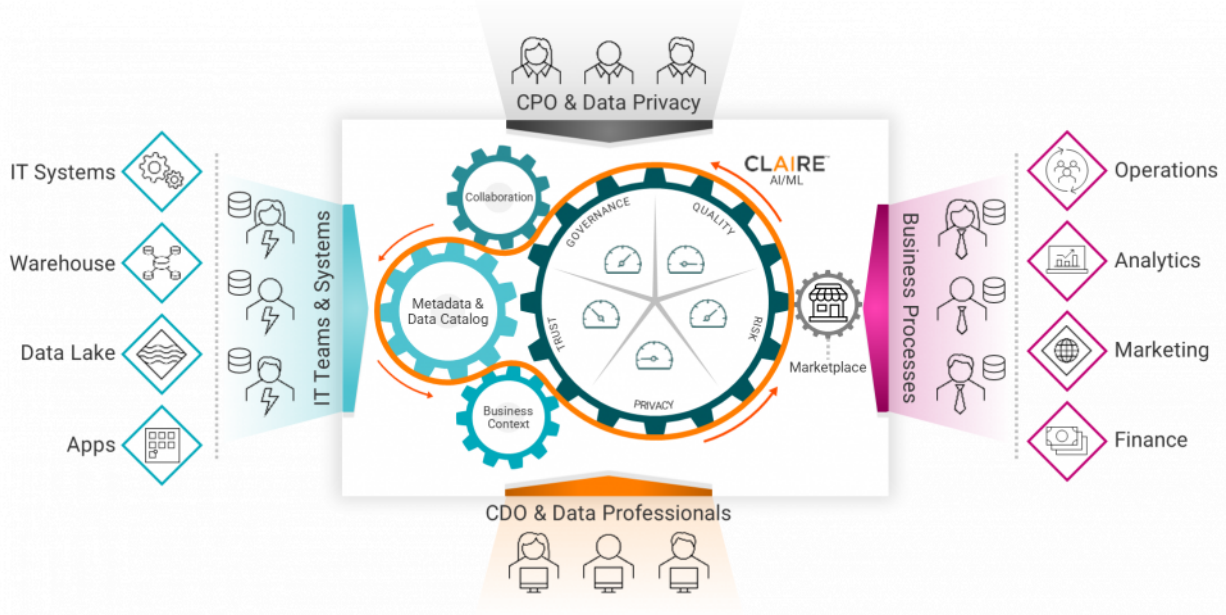


Figure 1: Informatica Cloud Data Governance and Catalog, tightly integrated with the Informatica Intelligent Data Cloud and Informatica CLAIRE AI platform, provides a unified perspective on all data to the various groups of users. [Source of graphic: Informatica.]

Informatica Cloud Data Governance and Catalog builds on a graph representation of the data. All metadata is held in a "connected knowledge graph", i.e., a structure that links data and allows the users to navigate through the data, understanding the relationship of different metadata and data. To populate the graph, Informatica builds on the proven capabilities for metadata extraction, data lineage, data profiling, data classification and linkage to business glossary, and data quality management.

The solution is able to extract and qualify data from a very broad range of sources, ranging from legacy on-premise solutions, to modern cloud storage. Thus, it can build a comprehensive representation of data and its lineage.

Users have a range of options for either navigating through the data catalog, or searching for data. Search, amongst other options, includes capabilities such as natural language-like search, query generators, previews of data assets, and delivers perspectives for both technical and business users. Informatica Cloud Data Governance and Catalog also creates and displays key metrics on data that help users to better understand the value and usefulness of data, e.g., by ratings for the data quality.

Additionally, there is a range of different components in the user interface, from the graph representation and search capabilities to dashboards, task lists, and workflows. Workflows, e.g., can be used to manage approvals to business glossary and terminology, or for triggering the human tasks required in managing and improving data quality for newly-added data sources.

As mentioned before, Informatica Cloud Data Governance and Catalog also makes extensive use of the AI capabilities provided by Informatica CLAIRE as an integral part of the Informatica Intelligent Data Management Cloud. Beyond that, the solution also supports the management of the AI models by adding

governance capabilities for AI models. AI models can be represented as assets in the data catalog & governance tool, inputs and outputs can be documented, key metrics related to performance, bias and data drift can be tracked and monitored, and parameters for alerting can be set, e.g., if drift and bias scores are outside of defined thresholds.

3 Strengths and Challenges

Informatica has leveraged the proven strength in Data Governance, Data Catalogs, and Data Management for its new SaaS solution, the Information Cloud Data Governance and Catalog. The new solution provides not just a parity set of features, but adds capabilities specifically for AI Governance, and for utilizing AI to improve its core feature set. Furthermore, the solution has been architected as a real SaaS service, being serverless and highly scalable. With the multi-cloud support, it also supports the common business needs of managing data already sprawling across multiple clouds, plus the connect back to data still held on-premises.

Due to the very broad set of capabilities and the inherent challenges of Data Management and Data Governance, customers are strongly advised to thoroughly plan and execute their data journey. A significant advantage that the new SaaS solution brings is that consumption of services can commence gradually, and is backed by a pay-per-use data model across capabilities within Cloud Data Governance and Catalog, and also across all services in the Intelligent Data Management Cloud. That allows for a step-wise roll-out of new capabilities. In any case, however, there needs to be a clear plan, and there need to be defined responsibilities, processes, and policies for Data Management in the organization. Only then will businesses succeed in leveraging the potential of all the data they own.

Informatica Cloud Data Governance and Catalog is an interesting solution for organizations moving towards a centralized, efficient, and comprehensive approach to Data Management and Data Governance, that helps in making data available for analytics, AI and other business use cases. We strongly recommend including this solution in the tools evaluation for that market segment.



Strengths

- Comprehensive set of capabilities across all areas of Data Management and Data Governance
- Deployed as SaaS solution, serverless architecture, ready to govern data across multiple clouds
- Pay-per-use pricing model, allowing for flexibility in scaling based on business needs.
- Modern user interface with a variety of options for the users, from dashboard to business and technical views
- Supports a knowledge graph representation for navigation and drill-down into the connected data
- Outstanding capabilities in the area of data lineage
- Good support for a broad range of data sources, including strong legacy support
- Integrates with Informatica CLAIRE AI capabilities, but also allows for governance of AI models

Challenges

- Variety of components within the solution, and in the Informatica Intelligent Data Management Cloud (IDMC), requiring customers to understand their needs and to have a well-thought-out roadmap for implementation, though having integrated data management capabilities in the same platform provides added flexibility.
- Inherent complexity of Data Management and Data Governance requires a well-thought-out approach for the organization's journey towards becoming data-driven or data-backed
- Despite strong capabilities, the effort required for connecting to data sources and adding them to the data catalog and data lineage must not be underestimated, though having these capabilities in a SaaS service simplifies management and operations, and enables faster time to value.

4 Related Research

[Database and Big Data Security \(kuppingercole.com\)](#)

[Enterprise Databases in the Cloud \(kuppingercole.com\)](#)

[Database and Big Data Security \(kuppingercole.com\)](#)

[Getting a Grip on Your Big Data - Informatica Announces Acquisition of Compact Solutions](#)

[\(kuppingercole.com\)](#)

[Fighting Data Friction, One Battle at a Time \(kuppingercole.com\)](#)

Content of Figures

Figure 1: Informatica Cloud Data Governance and Catalog, tightly integrated with the Informatica Intelligent Data Cloud and Informatica CLAIRE AI platform, provides a unified perspective on all data to the various groups of users. [Source of graphic: Informatica.]

Copyright

©2021 KuppingerCole Analysts AG all rights reserved. Reproduction and distribution of this publication in any form is forbidden unless prior written permission. All conclusions, recommendations and predictions in this document represent KuppingerCole's initial view. Through gathering more information and performing deep analysis, positions presented in this document will be subject to refinements or even major changes. KuppingerCole disclaim all warranties as to the completeness, accuracy and/or adequacy of this information. Even if KuppingerCole research documents may discuss legal issues related to information security and technology, KuppingerCole do not provide any legal services or advice and its publications shall not be used as such. KuppingerCole shall have no liability for errors or inadequacies in the information contained in this document. Any opinion expressed may be subject to change without notice. All product and company names are trademarks™ or registered® trademarks of their respective holders. Use of them does not imply any affiliation with or endorsement by them.

KuppingerCole Analysts support IT professionals with outstanding expertise in defining IT strategies and in relevant decision-making processes. As a leading analyst company, KuppingerCole provides first-hand vendor-neutral information. Our services allow you to feel comfortable and secure in taking decisions essential to your business.

KuppingerCole, founded in 2004, is a global, independent analyst organization headquartered in Europe. We specialize in providing vendor-neutral advice, expertise, thought leadership, and practical relevance in Cybersecurity, Digital Identity & IAM (Identity and Access Management), Cloud Risk and Security, and Artificial Intelligence, as well as for all technologies fostering Digital Transformation. We support companies, corporate users, integrators and software manufacturers in meeting both tactical and strategic challenges and make better decisions for the success of their business. Maintaining a balance between immediate implementation and long-term viability is at the heart of our philosophy.

For further information, please contact clients@kuppingercole.com.