Bloor InBrief

Philip Howard – Research Director, Information Management
Daniel Howard – Analyst

Bloor MUTABLE AWARD 2020 GOLD

Informatica

www.informatica.com

2100 Seaport Blvd
Redwood City, CA 94063, USA
Tel: +1 (800) 653 3871

Informatica Data Privacy Management

The company

Informatica was founded in 1993, initially focusing on data integration (extract, transform and load: ETL). In 1999 it floated on the NASDAQ but subsequently went private again in 2015, its shares having been acquired by Permira and several other investors (including Microsoft and Salesforce). Over the years, the company has grown both organically and through acquisition. Today, Informatica has evolved from a company that specialised in ETL to one that covers a broad range of data management capabilities, including data quality, data governance, master data management, data privacy and protection, and more. The company has more than 9,500 customers, operates world-wide and across all industries. It has revenues in excess of $1.2 bn (the actual figures are not disclosed).

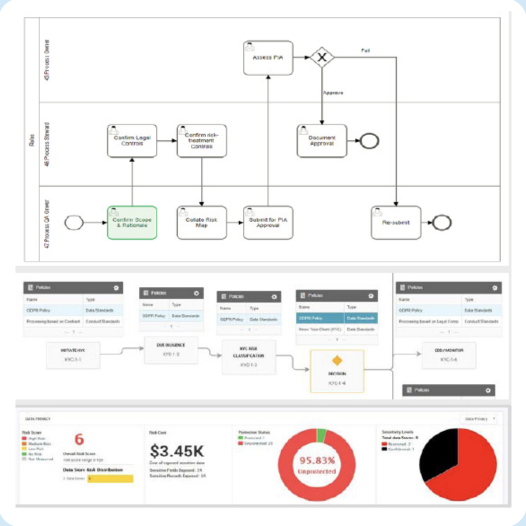
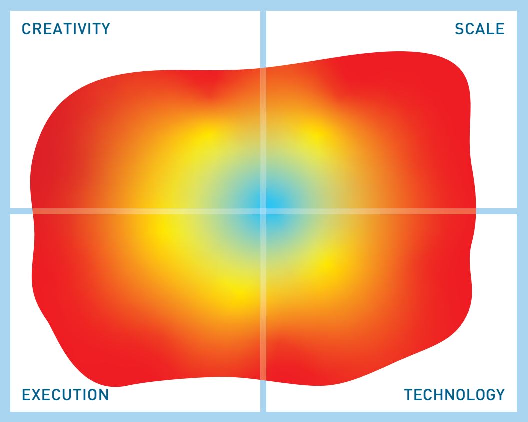


Figure 1 – Informatica Data Privacy Management integration with Axon

What is it?

Informatica Data Privacy Management is a data-centric privacy, governance and security solution that is focused on discovering and classifying sensitive data to understand how it moves around the organisation, where it may be located from a geographical perspective, who owns the data, and which people and processes access that data. In short, to manage privacy risks in a comprehensive, integrated, solution. The product shares a common metadata platform with Informatica Enterprise Data Catalog (EDC) as well as the Informatica Axon data governance offerings. These solutions have many of the same cataloguing capabilities as Informatica Data Privacy Management. Informatica also provides format preserving encryption as well as both static and dynamic data masking, which are used to protect sensitive data. The company also offers encrypted archival capabilities. For consent management, you can use the company's MDM offering for consent mastering, and Informatica also partners with both OneTrust and TrustArc.



The image in this Mutable Quadrant is derived from 13 high level metrics, the more the image covers a section the better. Execution metrics relate to the company, Technology to the product, Creativity to both technical and business innovation and Scale covers the potential business and market impact.

Specifically for discovering sensitive data the product supports relational databases in the cloud or on-premises, applications such as Salesforce and SAP R/3, Amazon S3, ETL processes (limited to Informatica, Microsoft SSIS and Cloudera Navigator and Atlas), file systems and both SharePoint and OneDrive. However it is less of a primary focus area – it is not alone in this – when it comes to unstructured and NoSQL data sources. For example, it is limited to supporting Hive and HDFS at present, though the company plans to support Cassandra in version 6.0 (the current release is 5.1) and BigQuery.

What does it do?

Considered holistically, Informatica's approach is that you start by creating actionable data privacy policies (integration with Axon – see ***Figure 1*** – from which you can overlay policies into Informatica Data Privacy Management). Then, discover and classify your sensitive data; uncover and map *"identities"* to data

Automation & ease of use	★★★★	Performance & scalability	★★★★
Discovery	★★★★½	Relational data	★★★★½
DSAR, masking & compliance	★★★★	Security & classification	★★★★½
NoSQL/cloud object storage	★★★½	Unstructured data	★★★

> “Before we embarked on this journey, we didn't have a clear view of sensitive data. Now, with Informatica, we can see and manage our entire universe of information. This capability is a game-changer, and it's enabling us to take a proactive approach to data protection that is helping to strengthen customer trust in our services.”
>
> Financial Services Company

(that can be used to support data subject rights requests under regulations such as the GDPR and CCPA); analyse the risks posed by sensitive data so that you can prioritise your protection plans; protect the data (masking and other techniques); respond to rights and consent requests; and, finally, be able to track and report on all of this.

The facilities for discovering sensitive data, which may be run against samples of the data, if required, are extensive and can be automated using ML and AI. You can match on the metadata via patterns, regular expressions and rules and you can also introspect SQL – both SQL queries and any SQL used for data movement purposes – though not stored procedures. Distance constraints (for example, post code needs to be near city name) can be used and you can define white (always sensitive) and black (never sensitive) lists. In the latest version, the underlying engine can make recommendations about what should be in these lists. The leverage of primary/foreign key relationships is planned for the next release. For unstructured data the product uses AI to look for parts of speech and otherwise relies heavily on the use of reference data. When potentially sensitive data is discovered you can set your system up to automatically agree that this is sensitive or that it is not sensitive or that it needs human validation, according to thresholds that you define.

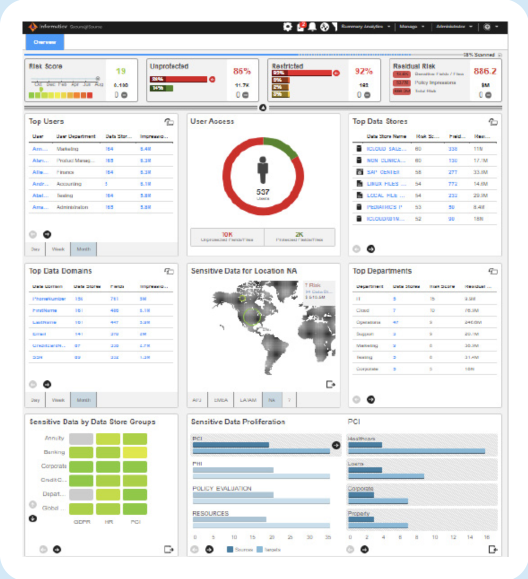


Figure 2 - *Risk simulation planning in Informatica Data Privacy Management*

The use of identity mapping, is interesting because it allows you to support rights requests: for example, where is Philip's data? To discover and map identities, the product uses fuzzy matching and the product ships with various pre-built classification policies such as PCI, GDPR and so forth. This is augmented by support for domains (name, email address and so forth) which are provided out of the box and can be combined.

Finally, we should mention the risk scoring (see ***Figure 2***). In addition to providing risk analytics and key performance indicators there is also risk simulation planning. This allows you to see the impact of using different approaches to say, masking.

Why should you care?

While sensitive data isn't only about personal data, it is issues over complying with new privacy regulatory mandates that are driving the market for sensitive data discovery and the subsequent protection of that data. Informatica is well-known as a market leader in the data management space and this is where its strength lies. The company has very strong credentials when it comes to structured data and it has focused its discovery capabilities in this area, where it has significant strengths. We particularly like the company's support for managing identities, which makes a lot sense within the context of GDPR, CCPA and similar regulations to determine data access.

Conversely, Informatica has respectable rather than comprehensive capabilities when it comes to discovery in unstructured environments. But, and this is a big but, companies that have focused on discovery for unstructured data tend to have very limited structured capability, typically limited to just Oracle and SQL Server. Most large enterprises are not limited to just these providers which would mean having two different sensitive data discovery solutions which, to our minds, does not represent any sort of solution.

The Bottom Line

Organisations should be aiming to have a single solution for sensitive data discovery that enables a data privacy governance strategy across a global enterprise. Organisations with multiple heterogeneous database implementations, as well as file systems, that they need to secure, would do well to shortlist Informatica as one of only a few companies that offers significant structured data discovery along with unstructured support.

FOR FURTHER INFORMATION AND RESEARCH CLICK HERE