Bloor

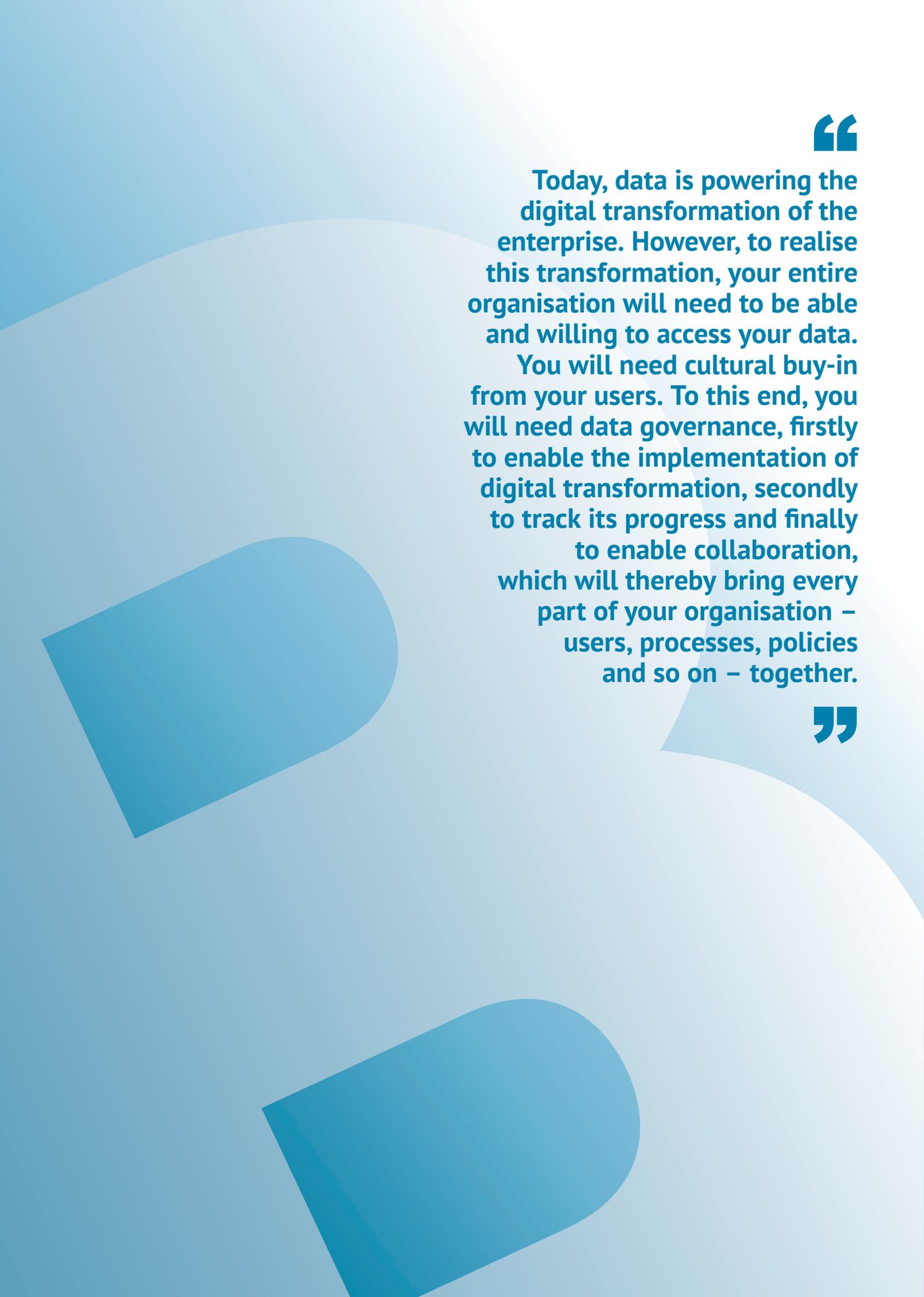## InDetail

InDetail Paper by Bloor
Authors **Daniel Howard and Philip Howard**
Publish date **August 2017**

Informatica Data Governance

"

Today, data is powering the digital transformation of the enterprise. However, to realise this transformation, your entire organisation will need to be able and willing to access your data. You will need cultural buy-in from your users. To this end, you will need data governance, firstly to enable the implementation of digital transformation, secondly to track its progress and finally to enable collaboration, which will thereby bring every part of your organisation – users, processes, policies and so on – together.

"

Executive summary

Data itself has been relatively constant throughout the history of enterprise computing. However, the way it has been used has changed dramatically, and continues to change now. In the past, data was provided primarily to be consumed by specific applications. More recently, data has been used to support enterprise-wide processes, such as business intelligence and analytics. Data is more valuable than ever, as evidenced by the rise of big data, and consequently there is vastly more data available in any given organisation than ever before. Not only that, there are a great many more types of data available and, moreover, many more users would like to access this data, the majority of whom will not be technical.

Today, data is powering the digital transformation of the enterprise. However, to realise this transformation, your entire organisation will need to be able and willing to access your data. You will need cultural buy-in from your users. To this end, you will need data governance, firstly to enable the implementation of digital transformation, secondly to track its progress and finally to enable collaboration, which will thereby bring every part of your organisation – users, processes, policies and so on – together. The accuracy of the data you are using, and trust in it, as well as compliance with an ever-increasing number of regulations, are additional considerations that cannot be provably handled without a data governance solution. Informatica's data governance platform is one such offering, which consists of a collection of products that have mostly been built on top of the Informatica Intelligent Data Platform. These products – Axon, Secure@Source, and Enterprise Information Catalog, plus the company's well-known data profiling and cleansing capabilities – come together to form a complete data governance solution that addresses the core governance issues of data quality, data cataloguing, compliance and privacy, and policy management.

Fast Facts

The three main products discussed in this paper are Axon, Secure@Source and Enterprise Information Catalog. The last two of these were developed directly by Informatica but Axon became part of the Informatica portfolio when the company acquired the original developer, Diaku, in early 2017. As might be expected, Axon is therefore not as tightly integrated as other Informatica products not supported by the same platforms, although the integration is ongoing through 2017. Briefly, Axon provides a collaborative business environment to support data governance, though organisational understanding might be a better way to put it. It also has capabilities designed to support and monitor compliance with specified regulations. Of the other two products, Enterprise Information Catalog is exactly what its name suggests; while Secure@Source provides a way to monitor and manage data-centric security, which can be enabled through the addition of Informatica's data masking (both static and dynamic) solutions.

Key Findings

In the opinion of Bloor Research, the following represent the key features of Informatica's data governance platform:

- A major focus across these products is on visual discovery. For example, Enterprise Information Catalog features fully navigable lineage and asset relationship views (the product has an underlying graph engine to enable exploration) that allow you to explore your system via the relationships between the data inside it. Axon (which also uses a graph database) has a similar approach.

- Collaboration is another strong point. For instance, Enterprise Information Catalog promotes collaboration by allowing users to relate business terms and technical assets, then display these relationships via annotations. This makes it much easier for non-technical users to understand the assets within the catalogue. Again, when using Axon, all stakeholders are collaboratively involved in the governance it supports.

> **Data is more valuable than ever, as evidenced by the rise of big data, and consequently there is vastly more data available in any given organisation than ever before.**

> **The biggest strength of Informatica's data governance solution is that it is complete... You can't say that about almost any other vendor.**

- Secure@Source is a data security and privacy tool. In addition to providing security analytics, it will continually monitor your system for suspicious or anomalous behaviour, automatically alerting your security team and starting a workflow that updates the data protection and access controls on the asset or assets in question. It is not only extremely rare to offer data-centric security in conjunction with data governance, we do not know of any other vendor that can offer the degree of integration between these two that Informatica does.

- Axon has policy hierarchy capabilities, including the ability to consider policies that relate to a project (for example, compliance with GDPR). Policies can then be connected to its business use, value and objectives via the various business facets of Axon – including Glossary, Process, Capability, Quality, Project, Client Types and Regulation. The way policies are linked to these facts is customisable, so the link between policy and business can be described in a detailed, flexible way.

- The individual products discussed here apply machine learning principles where appropriate. Axon has these built-in while Informatica's other products leverage the company's CLA/RE unified metadata engine. CLA/RE gathers and stores the metadata from your system, inferring additional metadata where it can, using machine learning techniques. CLA/RE can be used to provide a consistent metadata store across the Informatica environment.

- The biggest strength of Informatica's data governance solution is that it is complete. There might be individual features that we would like to see, and we might wish for tighter integration between individual products (especially between Axon and other products, though this will come in due course anyway) but Informatica does offer all the elements that you might look for in a data governance solution. You can't say that about almost any other vendor.

The bottom line

Axon forms the core of this solution, by both making your data accessible and providing a great deal of quality and privacy information around it. Moreover, it enables collaboration, via change requests and the workflows that are subsequently started, and policy management, via the creation of hierarchies to support the implementation of data quality and privacy-oriented policies. In fact, Axon by itself is a reasonable data governance solution. Even so, Enterprise Information Catalog and Secure@Source provide invaluable services which we would be loath to dismiss. Without Enterprise Information Catalog, there is no easy way to relate the assets present in Axon to your actual system – if this became necessary, it would likely require the intervention of a specialist, such as a database administrator – and without Secure@Source, you are missing a vital security component to your governance solution. Of course, there are other tools that fill these niches, but the integration is unlikely to be as smooth as with three products from the same vendor. In particular, you would need to find a way to utilise the same metadata store across your products, since you would not be able to leverage CLA/RE. The end result is that, together, Enterprise Information Catalog, Secure@Source and most importantly Axon form a powerful and effective data governance solution.

The products

Informatica's products are made available via the Intelligent Data Platform, and several are relevant to data governance. This paper will primarily focus on Axon, Secure@ Source and Enterprise Information Catalog as tools for monitoring data quality, data privacy and data cataloguing respectively. That said, it is worth bearing in mind that Informatica's product range is extremely broad, and there are products available on its platform that are often used in conjunction with the products and solutions we are discussing in this paper. For example, Data Quality, Data Masking, Master Data Management, and the Intelligent Data Lake are all Informatica products that are potentially relevant to data governance. Leveraging this platform across your organisation has an obvious benefit compared to trying to integrate multiple tools from multiple vendors. From the perspective of data governance, the chief advantage is that it means your users – both technical and not – will be using the same platform, which will mean that your solution could potentially be used to enable collaboration across your entire organisation.

The products we will focus on – and several others besides – are mostly built on CLA/RE, Informatica's enterprise-grade unified metadata engine. CLA/RE forms the backbone of these applications, gathering all the metadata available in your system and inferring additional properties – lineage, data quality, data sensitivity, and so on – via data profiling. This core of metadata is then used across all of your Informatica products, allowing them to remain consistent with each other and eliminating the need to re-gather metadata for each application.

Data Cataloguing

Informatica's data cataloguing capabilities come in the form of Enterprise Information Catalog. Available in the cloud, on-premises and for big data (the latter usually in conjunction with its sister product, Intelligent Data Lake), Enterprise Information Catalog automatically scans and indexes data assets from within your enterprise, then makes them available for

discovery – via a data catalogue – by your users. This has all the usual benefits of a data catalogue, namely that it allows your users to get self-service access, thereby eliminating any bottlenecks that might otherwise occur. Its basic features are a global search of your data assets and data profiling of those same assets, but it also includes data lineage, impact analysis, the ability to trace how your data is moving, and a graph-based map of the relationships between data in your enterprise (this is shown in **Figure 1** below). This map can be navigated around, allowing you to explore your system via the relationships between your data.

Figure 1: The relationships view of a data asset in Enterprise Information Catalog
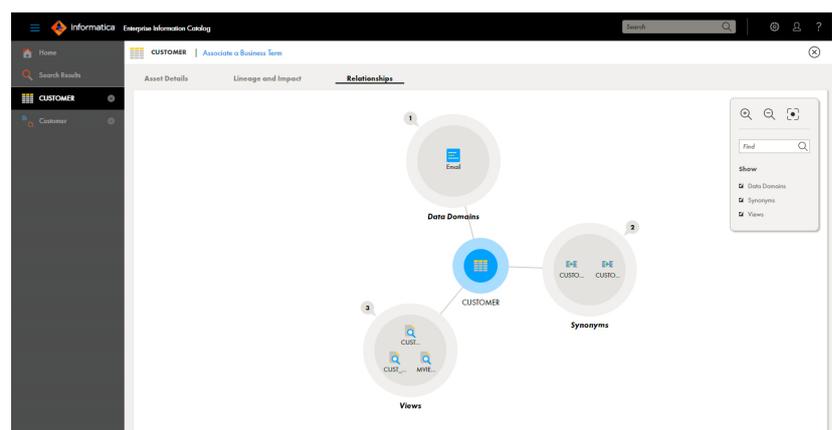
One of the most useful features of data catalogues is their capacity to support collaboration, particularly between technical and non-technical users, and Enterprise Information Catalog is no exception. Notably, it allows users to create business classifications and relate them to technical data assets as annotations, making it much easier for non-technical users to find what they're looking for. It also features recommendations for additional data that is similar to what you're already looking at. These recommendations are backed by a similarity engine powered by machine learning, meaning that they will improve over time as the product 'learns' how you (and your organisation) use your system.

> **Enterprise Information Catalog automatically scans and indexes data assets from within your enterprise, then makes them available for discovery – via a data catalogue – by your users.**

Enterprise Information Catalog also offers a particularly useful integration with Axon that allows you to automatically map the conceptual model of your system that resides inside Axon to the physical model of your system within the catalogue. The same is true of lineage, allowing you to match technical lineage in the catalogue to business lineage in Axon. This means that your users can choose to work entirely within your organisation's conceptual space as captured by Axon. If they need to concretise the data they are looking at, they can very easily move into the data catalogue and get the physical model that they need. The advantage of all this is that your users don't need to understand your physical system to get useful data out of it, which is important for enabling self-service, data discovery and collaboration, particularly for non-technical users.

## Figure 2: The data quality dashboard in Axon

Data Quality and Governance

Informatica's data quality and governance offering comes in two parts, Data Quality and Axon. Axon is concerned with understanding data in your organisation and enabling you to create policies accordingly, whereas Data Quality's focus is on supporting those policies via data transformation and stewardship. As this paper is about data governance specifically, we will concentrate on the former.

As mentioned, Axon's primary goal is to facilitate understanding of the data within your organisation, including the quality of that data. To this end, it provides a data quality dashboard, as shown in **Figure 2**, that monitors data quality over all governed assets. However, the emphasis here is on monitoring data quality at a business level rather than at a technical level, which is what you use Data Quality for. At present, this potentially means that data stewards will have two different dashboards to use: the IT data stewards would use the dashboards within the Data Quality and the business data stewards, the data governance practitioners, would use the dashboards within Axon. However, the dashboards are intended for collaboration within data governance regardless of role, so both IT and business can operationalise their initiatives around both Data Quality and Axon.

Axon allows access to the data sets available in your system, as well as metadata attributes, business terms, processes, and so on, with an emphasis on modelling your data as it relates to your business rather than (as in a data catalogue) at a more technical level. This has the advantage of making your data, and the relationships between it, much easier to understand and utilise, particularly for non-technical users. If needed, it is easy to access the physical data underlying the model you're looking at via the integration with Enterprise Information Catalog described in the previous section. What's more, this access is provided via a multi-dimensional search across your entire system, either globally or for a particular type of asset.
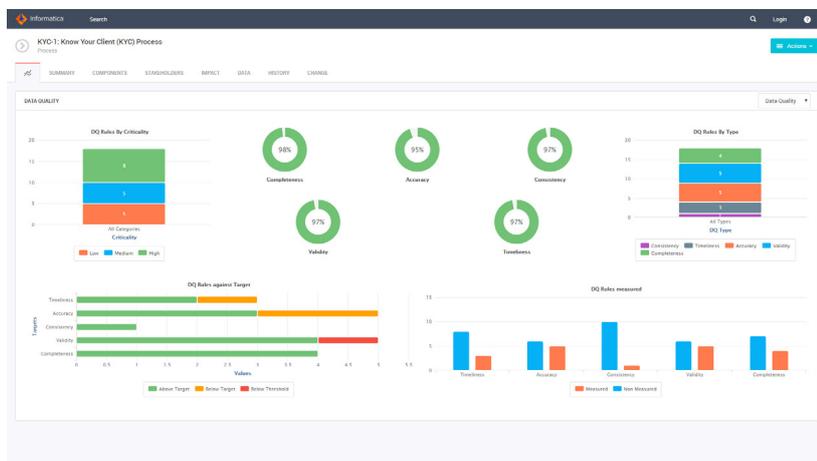
Once you've moved past the dashboard and navigated to a data asset, a detailed data quality assessment is displayed, in similar style to the dashboard but contextualised to the data in front of you. Business lineage data is available and displayed visually, enhanced by optional overlays displaying information about the assets contained within the lineage diagram. If you notice a problem with the asset you're looking at – for instance, the quality is poor, or the asset isn't governed by Axon at all – you can request a change that addresses the problem, either to a stakeholder directly or to the community stakeholders as a group. If a stakeholder approves your request, a workflow is started that moves between the stakeholders important for implementation, allowing each of them, in turn, to make the necessary changes. All changes made are documented and tracked as part of the workflow. The result is that every relevant person in your organisation can contribute to data quality based data governance, either by requesting changes or implementing them.

Data Security and Privacy

Axon also provides a second dashboard that relates to data privacy, focusing on compliance. And again, by navigating to a data asset you can view privacy information on that asset in more detail. This allows you to see precisely how your data has been used, accessed and reported on. As with many tools of this nature, this means that you can both ensure that you are complying to a regulation and prove that you are doing so.

Informatica also provides a second product to address data privacy concerns, namely Secure@Source. Unlike Axon, Secure@Source is primarily concerned with security and privacy: where Axon seeks to aid in understanding your data, Secure@Source is designed, first and foremost, to discover and analyse sensitive data, monitor and protect it.

Secure@Source's overview screen provides a dashboard for profiling sensitive data across your organisation, as can be seen in **Figure 3**. From here, you can analyse sensitive data as well as the risk to that data, determined automatically via several factors,

including protection status, user access and prior activity. You can navigate to additional screens in which you can view system-wide protection status in several different ways, such as via a geographic or heat map. You can also monitor data flows through your organisation, enabling insight into both where your data is and where it's moving to. Coverage includes semi-structured data (either CSV, XML or JSON) on either the Hadoop Distributed File System (HDFS) or Amazon S3, unstructured data (100+ formats including Office, PDF, text, XML, CSV, JSON and so on) as well as Microsoft SQL Service Integration Services (SSIS) proliferation information.

Secure@Source also has a continuous detection and alarm system, whereby it alerts the necessary users (such as your security team) if it identifies a high-risk condition or anomalous behaviour. This could be a suspicious pattern of behaviour, unauthorised use of a data asset, or violations of privacy or industry regulations. In addition, it will initiate workflows to update data protection and data access controls to cope with these conditions. These workflows can either be completely automated or route through an appropriate user or users, as desired.

> **Axon's primary goal is to facilitate understanding of the data within your organisation, including the quality of that data.**

Figure 3: The overview screen in Secure@Source

Policy management

In terms of generalist policy management, Informatica provides a business glossary within Enterprise Information Catalog that acts as a central place to store and manage business data assets, including policies. In addition, Axon provides a significant level of support for the creation and implementation of policies oriented around data quality and data privacy. Specifically, it allows you to design and build a policy hierarchy (see **Figure 4**), beginning with an abstract description of your policy, and ending with concrete business rules that can be executed on your data assets. Policies will typically be organised by project although, of course, policies may be reused across projects. Projects may be purely internal or they may be related to regulations such as the General Data Protection Regulation (GDPR). As far as the reuse of projects is concerned, it is important that Axon has similar graphical capabilities as Enterprise Information Catalog, so that you can visually explore relationships both across policies and between policies and projects.

Figure 4: Process map of policies in Informatica Axon
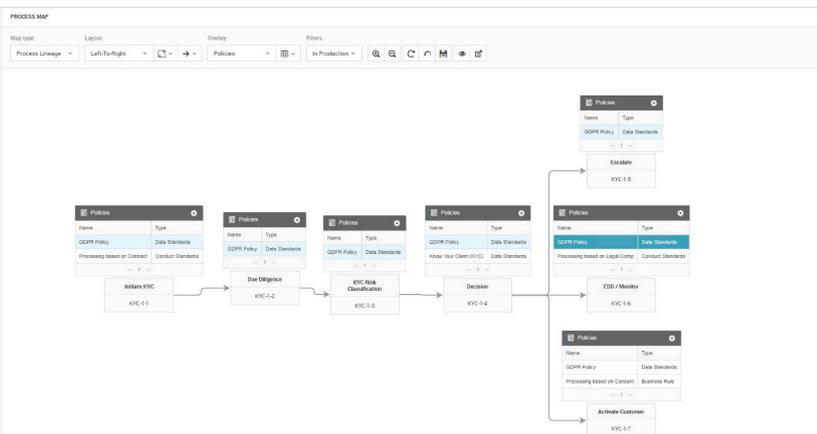
The vendor

Informatica was founded in 1993 as a services company specialising in helping its customers to migrate to a client/server environment. It was not until 1996 that it introduced its first product, Informatica PowerMart, which was followed by Informatica PowerCenter in 1998. The following year the company floated on NASDAQ but in 2015 the company went through a leveraged buyout whereby Permira and the Canada Pension Plan Investment Board acquired Informatica for a price of $5.3bn. Microsoft and Salesforce were also investors. In the year prior to the acquisition the company had revenues in excess of $1bn and net income of over $100m.

Since going private Informatica has not only been aggressive in introducing new products but has also been transitioning away from an on-premises, traditional licensing model to a more cloud-based subscription-oriented approach. In a number of the markets it serves, Informatica is also having to evolve from a company that traditionally marketed itself to IT and technical departments, to one that is more involved at business levels. This is arguably more difficult to do than a transition to the cloud.

Website: **www.informatica.com**

Conclusion

nformatica's goal with its data governance platform has been to enable the digital transformation of the enterprise by providing universal access to data, and enabling collaboration between all parts of the organisation to govern that data. In many ways, it has succeeded admirably: Axon alone provides a capable governance solution oriented towards the business, and Enterprise Information Catalog and Secure@Source only sweeten the deal.

However, we would like to see the offering strengthened in terms of the policy management features offered. That said, competitors that offer more advanced capabilities in this area are few and far between, and they lack the breadth and depth that Informatica can offer. All in all, Informatica's data governance platform is far ahead of the curve, and an excellent aid to digital transformation.

> "
> **Axon alone provides a capable governance solution oriented towards the business, and Enterprise Information Catalog and Secure@Source only sweeten the deal.**
> "

FURTHER INFORMATION

Further information about this subject is available from
www.BloorResearch.com/update/2345

About the authors
DANIEL HOWARD
Information Management

Daniel started in the IT industry relatively recently, in only 2014. Following the completion of his Masters in Mathematics at the University of Bath, he started working as a developer and tester at IPL (now part of Civica Group). His work there included all manner of software and web development and testing, usually in an Agile environment and usually to a high standard, including a stint working at an 'innovation lab' at Nationwide.

In the summer of 2016, Daniel's father, Philip Howard, approached him with a piece of work that he thought would be enriched by the development and testing experience that Daniel could bring to the table. Shortly afterward, Daniel left IPL to work for Bloor Research as a researcher and the rest (so far, at least) is history.

Daniel primarily (although by no means exclusively) works alongside his father, providing technical expertise, insight and the 'on-the-ground' perspective of a (former) developer, in the form of both verbal explanation and written articles. His area of research is principally DevOps, where his previous experience can be put to the most use, but he is increasingly branching into related areas.

Outside of work, Daniel enjoys latin and ballroom dancing, skiing, cooking and playing the guitar.

PHILIP HOWARD
Research Director/Information Management

Philip started in the computer industry way back in 1973 and has variously worked as a systems analyst, programmer and salesperson, as well as in marketing and product management, for a variety of companies including GEC Marconi, GPT, Philips Data Systems, Raytheon and NCR.

After a quarter of a century of not being his own boss Philip set up his own company in 1992 and his first client was Bloor Research (then ButlerBloor), with Philip working for the company as an associate analyst. His relationship with Bloor Research has continued since that time and he is now Research Director, focused on Information Management.

Information management includes anything that refers to the management, movement, governance and storage of data, as well as access to and analysis of that data. It involves diverse technologies that include (but are not limited to) databases and data warehousing, data integration, data quality, master data management, data governance, data migration, metadata management, and data preparation and analytics.

In addition to the numerous reports Philip has written on behalf of Bloor Research, Philip also contributes regularly to *IT-Director.com* and *IT-Analysis.com* and was previously editor of both *Application Development News* and *Operating System News* on behalf of Cambridge Market Intelligence (CMI). He has also contributed to various magazines and written a number of reports published by companies such as CMI and The Financial Times. Philip speaks regularly at conferences and other events throughout Europe and North America.

Away from work, Philip's primary leisure activities are canal boats, skiing, playing Bridge (at which he is a Life Master), and dining out.

Bloor overview

Technology is enabling rapid business evolution. The opportunities are immense but if you do not adapt then you will not survive. So in the age of Mutable business Evolution is Essential to your success.

> *We'll show you the future and help you deliver it.*

Bloor brings fresh technological thinking to help you navigate complex business situations, converting challenges into new opportunities for real growth, profitability and impact.

We provide actionable strategic insight through our innovative independent technology research, advisory and consulting services. We assist companies throughout their transformation journeys to stay relevant, bringing fresh thinking to complex business situations and turning challenges into new opportunities for real growth and profitability.

For over 25 years, Bloor has assisted companies to intelligently evolve: by embracing technology to adjust their strategies and achieve the best possible outcomes. At Bloor, we will help you challenge assumptions to consistently improve and succeed.

Copyright and disclaimer

Bloor