



www.globalids.com

182 Nassau Street, Suite 202, Princeton,
NJ 08542, USA
Tel: +1 (888) 514 0192
Email: info@globalids.com

Global IDs Sensitive Data Discovery

The company

Global IDs is a privately funded company that was established in 2001. It provides solutions for enterprise information management and governance, focusing on the largest, most complex environments, which are typically multi-domain and often span multiple geographies. What's more, the company's technology suite is extremely broad ranging and offers a variety of functions, including data profiling, data quality, data governance, data lineage, and master and reference data management.

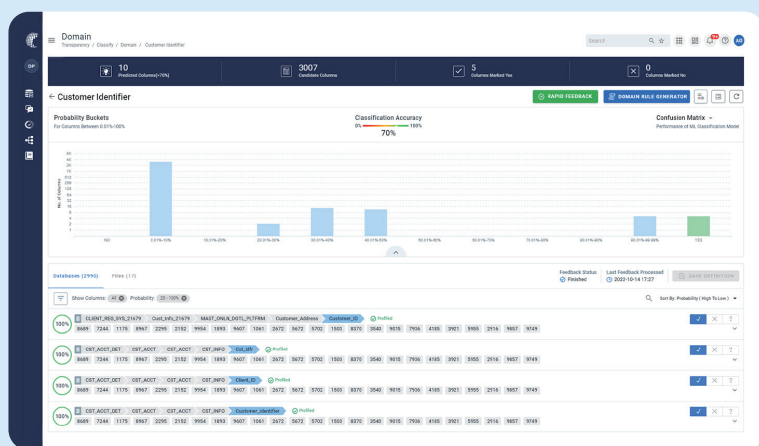
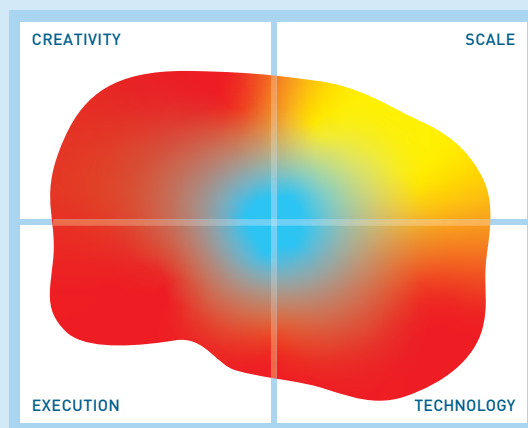


Figure 1 – Global IDs data classification

What is it?

Global IDs offers data discovery and classification as part of DEEP, its Data Ecosystem Evolution Platform. It thereby provides sensitive data discovery and compliance with regulations such as GDPR and CCPA. Four primary capabilities are provided: discovering sensitive data in general, scanning for data associated with a particular individual, locating individuals from an address (a CCPA-specific capability), and creating privacy reports that prove compliance. A number of other relevant capabilities, such as data lineage, are available, as are enterprise-wide visualisations of your entire sensitive data landscape. What's more, as a Global IDs product, it has been designed to do all of this at scale, regardless of the size of your ecosystem.



The image in this Mutability 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.

The product supports a wide range of data sources and file formats, as you would expect from a product designed to support large and therefore often highly varied ecosystems. In particular, it provides support for both relational and NoSQL databases, the latter most notably including MongoDB and Cassandra but, in principle, any data source that can be resolved into a columnar structure. Mainframes are also supported, as is Amazon S3, while unstructured support extends to text files and emails.

What does it do?

For sensitive data discovery, Global IDs leverages its data profiling, classification, and lineage capabilities, as seen in [Figure 1](#). In turn, this process identifies personal data as well as the individual that it corresponds to. For data classification, the product heavily (and increasingly) leverages semantic tagging and machine learning to identify columns and tables containing sensitive data, backed up by more traditional methods such as rule and pattern-based matching. The idea is that it will look at each field holistically – including both associated metadata and the values of the data itself – and make a classification recommendation, which can either be accepted or rejected. Either way, this user feedback will be incorporated into its underlying model, enabling it to automatically adjust to your

Data Privacy Report For		Generated By	
John Smith		Global IDs (India)	
Report Version: 1		Generated On: 25 Oct 2019	
Organised by Domain of Interest			
Domain	Value		
Date Of Renewal	2022-11-08		
Price	150000		
CaliforniaAddress	831 Chapel Dr San Francisco		
California Zip Code	95210		
California Phone No	209-200-2011		
Company Name	Companies Default		
Yearly Premium	1200000		
Owner Last Name	Smith		
Owner Email Address	john.smith@global.com		
California City	Alhambra		
Owner First Name	John		
Policy Number	P12378904		

Figure 2 – A data privacy report in Global IDs

system and become more accurate over time.

This is particularly relevant given the size of ecosystem that Global IDs will generally be working with: with such a large system, creating appropriate classification rules and keeping them up to date will be difficult and time

consuming. Automating this process is therefore highly valuable. Notably, Global IDs also leverages disambiguation and validation as part of its process for discovering personal data, which can help to eliminate false associations.

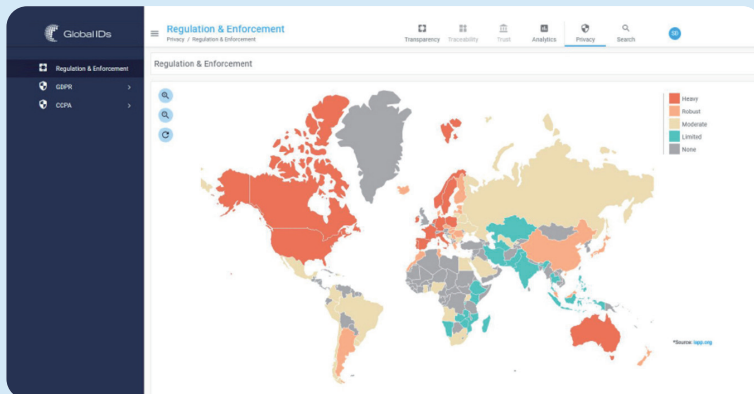


Figure 3 – Global regulation and enforcement view in Global IDs

In addition to the discovery of sensitive data, Global IDs also allows you to search for and track all personal information related to a specific individual within your system, going so far as to create a data privacy report to that effect, as shown in [Figure 2](#). The reports themselves can be produced and accessed through a privacy dashboard, with any sensitive values contained therein masked automatically. They are highly useful for addressing Data Subject Access Requests (DSAR), and particularly impressive when you consider the scale that the product operates on: across a large, possibly pan-global data ecosystem, it is likely that data relating to an individual will end up distributed far and wide across your system. Global IDs allows you to reconsolidate that information by providing a centralised view of it.

As an extension to the above, the platform is designed to help you create a comprehensive view of your (sensitive) data landscape. This includes the creation of privacy domains (which is to say, types of personal information), as well “*semantic objects*” that consist of multiple such domains. These objects can then be treated as singular entities for the purposes of, say, search, allowing you to locate specific groupings of personal information across your enterprise.

The product also provides full data traceability and data lineage, allowing you to see where a given individual’s data is being used within your system. It does this by generating hypotheses regarding where said data is being (or has been) used, based on both human and machine input. It is thus able to recommend probable flows that can be confirmed or rejected by your users by validating them against your actual system. Further complementary capabilities, such as data masking, are provided through partnerships.

Finally, to aid in understanding the geography of regulations like GDPR, Global IDs provides an interactive map of the world that displays each country alongside the regulations that apply there. For the USA specifically, there is a separate map that does the same on a state-by-state basis. The first of these views is shown in [Figure 3](#).

Why should you care?

The foremost reason to use Global IDs is its capacity to operate at extremely large scales. As with many of the spaces it operates in, there are few, if any, vendors that can compete in this regard. Global IDs is able to operate on huge numbers of data sources, which may themselves be geographically distributed, and not only locate personal data across these sources, but associate that data to a specific individual (who can then be searched for to return this data) and locate where in your system it is being used. These are essential capabilities for GDPR compliance – suppose a customer asks, as is their right, for all of the information you are storing on them – and without Global IDs you will be hard pressed to provide them at such scales. Moreover, it is worth noting that privacy regulation requires both compliance and convincing proof of said compliance. Global IDs’ reporting functionality can be instrumental for the latter.

Global IDs also supports a much broader range of data sources than many competing products in the space, including mainframes, relational databases, multiple types of NoSQL, S3, and others. This is a notable capability in a space which, for the most part, has very limited database support. Although support for one or two of these ranges of technology is common, supporting all of them within a single product is much less so.

The Bottom Line

If you need sensitive data discovery at very large scale, Global IDs should be your first port of call. Even at somewhat less extreme scale, it is well worth looking at.

[FOR FURTHER INFORMATION AND RESEARCH CLICK HERE](#)