

Intelligently Classify Sensitive Data & Files Everywhere

ML Driven Data Classification

The definition of sensitive data is no longer readily encapsulated in a regular expression. Increasingly, companies need to classify data that is sensitive based on context to a person, or a thing like patent or account. This requires a new approach to classification that can identify contextually sensitive data across all modern data stores - unstructured, structured, Big Data, Cloud and enterprise applications like SAP. BigID provides a first of its kind approach that combines Machine Learning and Contextual Intelligence to deliver on advanced data classification, categorization, cataloging and correlation for the privacy

ML-Driven Classification

Traditional pattern matching approaches to discovery and classification still struggle with accurately identifying contextually sensitive data like Personal Information (PI) and disambiguating similar looking information. Moreover, regular expression based classifiers which predominate in data loss prevention, database activity monitoring, and data access governance products tend to operate on a limited number of data sources, like relational databases or on-prem unstructured file shares. BigID leverages patented machine learning to classify, categorize and compare data and files across structured, unstructured, semistructured and Big Data in the cloud or on-prem. BigID can resolve similar looking entities and build association graphs to correlate data back to a specific entity or person - essential for meeting emerging privacy use cases like personal data rights.



Correlation plus Classification

Even with AI and ML classification approaches like clustering or random forest, classifiers can improve accuracy through smarter matching and comparison analysis - but lack the context to understand who the data relates to. This is a common problem for privacy requirements and regulated industries. The capability to build a graph of connected or relevant data can be characterized as a correlation problem. Correlation helps an organization find sensitive data because of its association to other sensitive data. BigID provides a first of its kind model that can, not only match similar data withing the same class based on ML analysis, but also match connected data of different classes based on relevancy and connectedness. This correlation-based classification is critical to privacy.



Intelligently Classify Sensitive Data & Files Everywhere

Cataloging plus Classification

BigID's ML-based classifiers use advanced AI techniques to match data within a class and also correlate data of different classes that have a common sensitivity level owing to a shared association. But, there is a third way sensitivity can be measured. Most data also has certain attributes associated with it, such as date of creation, last modification of ownership and access details. Unlike traditional classifiers, BigID can also integrate meta-data analysis to provide a richer view of the data and its usage. This meta-data input can be used to better and more automatically catalog data for easier discovery via search as well as measure sensitivity risk. The combination of intelligent classification, correlation and cataloging give organizations the unique ability to find, inventory and map sensitive data by additional dimensions than just data class or category. These include finding data by person, residency, application and ownership.



Intelligent Labeling and Tagging

Enforcement of security protection and privacy compliance requires data risk and sensitivity knowledge. BigID helps organizations understand data sensitivity through advanced ML-based classification, correlation and cataloging to provide a complete view of data. To simplify enforcement on classified data, BigID enables customers to automatically assign data tags for files and objects. These classification tags can be consumed through Microsoft's Azure Information Protection framework as policy labels, BigID's labeling APIs or additional frameworks like Box. Using these labels, organizations can classify or categorize data - such as Highly Sensitive, as well as Personal Data based on privacy, health or financial services compliance mandates. These tags can then be used for more granular policy enforcement actions by DLP, information rights management, database activity monitoring or other enforcement products.



How BigID Can Help

BigID is redefining personal data protection and privacy in the enterprise. Organizations are facing record breaches of personal information and proliferating global privacy regulations, with fines reaching 4% of annual revenue.

BigID gives enterprises software to automate the security and management of structured and unstructured PI across datacenters and cloud. Using BigID, enterprises can better steward their most vital assets: their customer, employee and client data.

For more information, email info@bigid.com or schedule a demo at bigid.com/demo