BigID

Discovery-In-Depth

The Path to Data Intelligence

—————

Privacy-centric data discovery (a must for data privacy and cybersecurity in today's environment) requires a multi-pronged strategy to identify all types of sensitive & personal data in an organization – and that strategy starts with discovery-in-depth.

BigID

INTRODUCTION

Data discovery is a foundational element of any type of data management: from cybersecurity to data privacy to data governance.

Discovery is at the core of data intelligence, insight, and analysis – and needs to be both scalable and automated in order to successfully address the volume (and type) of data that organizations collect.

Effective (and sustainable) privacy, security, and governance programs require discovery in-depth: empowering organizations to scratch more than just the surface of their data.

That means not only finding and identifying more types of sensitive and personal data with greater accuracy, but being able to apply context, insight, and perspective to that data – which then helps inform policy and controls.

It's no longer enough to only be able to identify regular expressions and common types of sensitive data (like credit card numbers or social security identifiers).

Privacy regulations like the CCPA and GDPR have transformed the very definition of personal data – extending it to a much broader set of data, taking into consideration things like geolocation, friendly names, online activity, and more.

"

Privacy-centric data discovery (a must for data privacy and cybersecurity in today's environment) requires a multi-pronged strategy to identify all types of personal and sensitive data - and that strategy starts with discovery-in-depth.

"

Unlike earlier regulations, today's data privacy initiatives focus on data that can be related to an individual, which means that data discovery solutions need to be able to identify personal data not just by type, but from contextual clues and relationships to other data points.

Furthermore, organizations are now responsible for not only protecting that data, but monitoring and reporting on whose data it is, where it came from, and where it's going.

Privacy-centric data discovery (a must for data privacy and cybersecurity in today's environment) requires a multi-pronged strategy to identify all types of sensitive & personal data in an organization – and that strategy starts with discovery in depth.

Discovery-in-depth:
How it Works

BigID leverages discovery-in-depth to provide deep data intelligence, combining multiple modes of discovery and context around sensitive data.

By applying machine learning and correlation, organizations can more accurately identify personal and sensitive data – and can understand data context and relationships (rather than looking at a data point in isolation).

The first layer of a discovery-in-depth approach is being able to find and identify sensitive data by regular expression (RegEx): these are typically sequences of characters that define a specific pattern. It's a technique that revolves around pattern matching and knowing the exact format of the sensitive data that you're trying to find: traditional identifiers from bank account numbers to email address formats to ID numbers.

BigID goes a step further with this approach to include common security attributes like explicit passwords, encrypted private keys, security tokens, and more.

> **Leverage discovery-in-depth to provide deep data intelligence, combining multiple modes of discovery and context around sensitive data.**

The next layer goes broader: discovering personal information (PI) that's traditionally more difficult to define – information like date of birth, voting trends, first names, last names, residency and more.

Machine learning techniques and context-based classifiers are able to uncover this type of data – discovering and inventorying a broader set of personal and sensitive data.

On top of that, deep data discovery requires an identity and entity-based approach: revealing data relationships, identities, inferred data, and associated data.

By adding correlation, organizations are able to not only uncover dark data, but are able to surface relationships between sensitive data – inferring new data attributes and extending visibility to all sensitive & personal data that they collect.

BigID

HOW TO APPROACH
Discovery-In-Depth

BigID approaches discovery-in-depth with 4 Cs: catalog, classification, cluster analysis, and correlation – all working together in a privacy-centric approach for unmatched data discovery and context around personal and sensitive data.

▶ **Catalog**
Automatically catalog and map sensitive & personal data with deep data insight, incorporating active metadata and classification. Gain additional privacy, security, and business insight – all within a single pane of glass.

▶ **Classification**
Classify data by type, identity, attributes, patterns, category, & policy. BigID goes beyond RegEx and applies different layers of classification to identify and analyze a more extensive set of attributes.

▶ **Cluster Analysis**
Leverage cluster analysis to rapidly and accurately identify file content and type, and label clusters of data for policy and enforcement. Cluster analysis is a machine learning technique to impose structure on unstructured data at scale.

▶ **Correlation**
Add context to classification and surface relationships between data points. Build identity and entity profiles, associate whose data it is, and visualize how data is interconnected across data sources.

BigID empowers organizations to know their data – and apply privacy, protection, and perspective to that data. A discovery-in-depth approach gives 360° visibility to sensitive data, along with deep data intelligence across all types of data, across all data stores – click here for a demo to see how BigID's discovery-in-depth approach transforms data privacy and protection.