WHITEPAPER

COMPARING DATA DISCOVERY & FILE ANALYSIS SOFTWARE SOLUTION

Data and privacy management (DPM) technologies are now expected to be available from the same product suite to protect enterprises from rising penalties for failing to meet privacy and compliance requirements. As the types, categories, locations, and processors of sensitive data continue to proliferate, every organization needs a comprehensive DPM solution that provides protection across all data stores. In addition, every DPM solution should leverage existing IT investments that support meeting governance and compliance demands.

The DPM platform itself should also provide constant and consistent classification and discovery capabilities to ensure continuous protection. The move to remote/hybrid work requires that effective discovery/classification be in place for endpoints, email, all cloud providers, and SaaS application solutions.

Ensuring data privacy also requires a consistent platform to protect the data, with centralized policy management applied to all data and with policy changes applied quickly across all data sets. The key: automating these processes.

This overview will explore and compare Data443's Data Identification Manager to other vendor solutions by BidID and Varonis.

BIGID ENTERPRISE

Getting the DPM solution choice right is necessary to protect the business and not all solutions are created equal. In many ways, tasks such as discovery, governance and classification, form the foundation for DPM, and the quality and breadth of this functionality are critical to success. The chart and text below identify how Data443 Data Identification Manager and BigID Enterprise compare across the most important aspects of this functionality.

	Data443	BigID
Cloud data drives	✓	?
Databases	✓	✓
Email	✓	
Endpoints	✓	
Data in flight	✓	✓
Data at rest	✓	✓
Structured data	✓	✓
Unstructured data	✓	✓

DATA SECURITY & PRIVACY MANAGEMENT

is critically important for protecting businesses from risk as well as compliance and legal issues due to loss or misuse of private or sensitive data. However, effective data privacy management requires a platform that has the capabilities to handle these complexities.

First the solution must identify data across the entire IT "estate"—that is, look across cloud services, endpoints, servers, cloud applications and all other infrastructure that can store data. The business data estate also includes data at third-party providers.

KEY FINDINGS

- Data443 supports all activities across the entire data estate, whereas BigID lacks the visibility into endpoints and email systems, which can contain plentiful and highly sensitive data
- Data443 can engage directly with data owners and stewards. BigID has less ability to support this direct engagement. Engagement with these individuals improves results and leverages data knowledge to deliver better results.
- Data443 can support interactions at the user/device level, which will be increasingly important as the move to hybrid/remote work continues, and that will include key stakeholders.
- Only Data443 supports classic enterprise use cases such as assigning share and folder owners and determining data ownership based on classification and not simply on usage.
- The Data443 platform engages other important parts of enterprise infrastructure such as Microsoft MIP, SIEM tools/appliances, CyberArk access management, and ShareFile. Data443 works well with other components of the technology stack and is a team player, not a silo

CLASSIFICATION & GOVERNANCE

Best-in-class data stewardship is delivered with active classification and governance processes that, ideally, occur continuously. These major activities comprise many individual tasks that enable effective data privacy management throughout the entire data lifecycle. Automating these tasks is the only effective way to complete the substantial amount of work requires, and this should be augmented with human management as needed. These management tasks will be jointly completed by IT and other teams such as the compliance office.

	Data443	BigID
End user classification/governance	✓	
Data access governance	✓	
Governance alerts	✓	
Live classification	✓	
Manual classification	✓	?
High-risk classification	✓	✓
Data audit with classification information	✓	✓

KEY FINDINGS

- Data443 provides broad classification capability across all data sets, regardless of location, including end users. It also enables data stewards to provide input.
- Data443 has strong orchestration functionality.
- Data443 is built with a “classify all the time” approach that also enables an organization to move at its own pace and learn as it goes. BigID lacks live classification which can result in data gaps that put sensitive data at risk.
- One of the most important capabilities of Data443’s solution is the ability to take governance actions on data, not just classify it. This results in better enforcement of existing IT data security policies to reduce risk.
- Data443 provides alerts for governance problems as they are found, either to SIEM or other alert types. BigID does not have this functionality.
- BigID misses the key functionality to manage the governance and remediation of access. Data443 includes all identity repositories for this capability.
- Data443 leverages 900+ active taxonomies included with the product, available in 14 languages and powered by both fuzzy logic and machine-learning technologies.

DISCOVERY & COMPLIANCE

The output most organizations focused on is discovery and compliance information. Given the amount of data stored by modern organizations, automation of these activities is mandatory, and it enables full lifecycle data privacy management. With these capabilities the organization will find it easier to meet changing compliance demands. The emergence of inconsistent and overlapping global and local compliance rules demand a comprehensive automated solution.

	Data443	BigID
All-enterprise search and discovery	✓	
Customer e-discovery	✓	✓
Sensitive-data detection	✓	✓
Workflow for compliance response and reconciliation	✓	
Compliance with GDRP, HIPAA, PCI and other standards	Explicit and fuzzy	Explicit only

IDENTIFYING IMPORTANT DIFFERENCES

- Data443 can deliver the global discovery that is required, using both “crawl” and indexed approaches.
- BigID is limited in this area of functionality in terms of customer e-discovery and sensitive-data detection features.
- The Data443 solution has a built-in privacy request panel with workflows for compliance response and remediation, including for FOIA, GDPR, CCPA, e-discovery, litigation sup- port, and retention management demands.
- Data443’s product has the ability to work with “fuzzy” logic, which drives many new and emerging compliance demands. Without this functionality, it is likely that the business will fail to meet some compliance and governance demands. BigID can work only with explicit demands.
- Support for full archiving management is another key benefit that Data443 provides that BigID does not. This feature supports archiving for email, unstructured data, and many ECM platforms, and actions can be based on both condition and time.
- Identifying data that is duplicated or not necessary is essential for efficiency, and only Data443 has next- generation ROT (redundant, obsolete, or trivial) identification capabilities, including fingerprinting, obfuscation, and other display options.

SUMMARY & KEY TAKEAWAYS

The Data443 solution, in comparison, contains a more modern approach to help organizations meet the challenge of current and future data privacy management – including capabilities around breadth, reporting and remediation.

The Data443 better positions to support the move to hybrid and remote work, with the ability to find and protect data on endpoints and in email systems. With the inclusion of classification and discovery abilities, Data Identification Manager proves the stronger solution – BigID’s lack of live classification functionality within the offering can further put organizations at risk and the absence of alert functionality is highly problematic.

The Data443 solution provides build-in workflows for compliance response and reconciliation, substantially reducing the number of resources necessary to complete projects, enabling faster completion.

VARONIS SECURITY PLATFORM

Choosing the wrong solution for this critical initiative can leave an organization with a big data knowledge gap. What follows is a competitive analysis of the solutions from Data443 and Varonis, examining their capabilities across three critical areas of functional capability. The better a solution can perform these tasks, the better the organization can protect its sensitive and private information. This chart shows how Data443 and Varonis compare.

	Data443	Varonis
Cloud data drives	✓	Limited
Databases	✓	
Email	✓	
Endpoints	✓	
Data in flight	✓	
Data at rest	✓	✓
Structured data	✓	
Unstructured data	✓	Limited

KEY FINDINGS

- Data443 can find and protect data in many more locations, such as every major cloud provider, laptops and desktops, and hundreds of SaaS services/applications.
- Varonis has very limited data identification capabilities and a comparatively small library of patterns. Data443 has more than 900 built in, in 14 languages, enabling a better false detection rate and more-accurate reporting and governance controls.
- New data types in new locations are increasingly important, and only Data443 can find them.
- Varonis has a substantial blind spot for user data on end user devices, and this is increasingly problematic as remote and hybrid work become the norm.
- Risks created when information is sent to others (both inside and outside the organization) make Data443's ability to discover sensitive information in all major email plat- forms (and archives if needed) a major advantage.
- Varonis lacks interaction with data owners, whereas Data443 engages data owners and stewards actively and continuously training its machine learning (ML) with your own data.
- Data443 assigns data ownership based on classification accuracy and other forward generation vectors (including ML), not only legacy methods such as who uses the data.
- In terms of integration, Data443 was first in many cases and continues to natively integrate with Microsoft MIP, SIEM tools, CyberArk access management, Dropbox and others.

CLASSIFICATION & GOVERNANCE

Effective data stewardship means that classification and governance are active tasks that must happen regularly or even continuously. Classification has many new subtasks that are essential to ensuring effective data privacy management as data moves through the lifecycle. The best setup is automated engines with human supervisory management – and the burden should not always be directly on IT.

	Data443	Varonis
End-user classification/governance	✓	
Data access governance	✓	✓
Governance alerts	✓	✓
Live classification	✓	
Manual classification	✓	
High-risk classification	✓	
Data audit with classification information	✓	✓

MAJOR DIFFERENCES

- Varonis has very limited classification functionality, with very few built-in sensitivity lists to pick from.
- Data443 can obtain visibility and feedback from non-datacenter data sets (end users, SaaS platforms, unstructured data sets) to support more accurate decision-making on classifications, a capability Varonis lacks.
- Varonis classification functionality often delivers false positives from regular expressions, which increases costs and results in bad execution.
- Data443 has more than 900 mature prebuilt classification schemas in 14 languages that support a "classify all the time" mentality, enabling organizations to take an iterative approach and reducing the demands and limitations of doing it all at once.

DISCOVERY & COMPLIANCE

Active and automated discovery and policy compliance functionality is essential to completing a full lifecycle data privacy deployment. This enables better response to changing compliance demands, and the emergence of mismatched and overlapping global and local compliance rules requires a modern solution that can respond to them.

	Data443	Varonis
All-enterprise search and discovery	✓	
Customer e-discovery	✓	✓
Sensitive data detection	✓	✓
Workflow for compliance response and reconciliation	✓	
Compliance with GDRP, HIPAA, PCI and other standards	Explicit and fuzzy	Explicit only
IT Governance actions	✓	

Data443 has numerous advantages here:

- Varonis' solution focuses on the discovery part of the process only and is limited by the basic capabilities of its tools, namely in on-premises and file systems only.
- Data443 provides a "global" discovery, including all cloud, 200+ database types and 300+ SaaS leaders, including Salesforce, Zoom, WebEx, Snowflake and others.
- The automated governance action capabilities of Varonis are limited and provide little value in automated compliance management
- Data443 offers strong support for compliance reports and remediation, supporting key regimes such as FOIA, GDPR and CCPA, plus workflows surrounding e-discovery and retention management – including archiving, retention and request portals with over 40K customers using it today.
- Many new compliance and governance standards are still being developed, thus introducing vague requirements, but Data443 can adapt to these nonexplicit directives as they develop.
- Data443 solves the problem of archiving data across all data types and sources based on conditions such as policy, sensitivity, classification, age, and other factors.
- Strong policy enforcement and reporting are key capabilities of Data Identification Manager – this technology detects governance failure conditions (permissions, access control and data sensitivity) and executed actions (SIEM, email alert, data encryption, permissions removal, etc).
- Identifying data that is duplicate or not necessary is essential for efficiency, and Data443 has next-generation ROT – Redundant, Obsolete or Trivial – identification capabilities, including fingerprinting, obfuscation and other options.

SUMMARY & KEY TAKEAWAYS

The Data443 solution is designed with current and future data privacy demands in mind. Unlike Varonis' solution, which is focused primarily on legacy on-premises reporting use cases, the Data443 solution supports all aspects of the data privacy management life cycle.

Data443 possesses a substantial advantage based on the ability to look across the IT estate for private or sensitive data, not just in on-premises file shares. The inability of the Varonis solution to find private data in cloud services, SaaS solutions, or endpoints is a substantial drawback and can leave organizations with large gaps in their data inventory and potentially noncompliance risks for privacy laws.

Organizations need a comprehensive data privacy platform that not only locates all sensitive data but also provides auto- mated and effective tools for protecting it. The system must also be dynamic and meet new or changing compliance, legal, and governance demands. Finally, the privacy platform that meets forthcoming regulatory requirements must also seamlessly integrate with existing IT infrastructure—not be a new technology silo to manage.