white **paper**

# Bringing Order to Shared Drives: Introducing Privacy Aware Governance

**Presented by ASG Technologies, BigID, and Zia Consulting**

## De-Risking Cloud Collaboration and Protecting Sensitive Data Across SharePoint, Box, Microsoft 365, and Teams

## MARKET DRIVERS BEHIND PRIVACY AWARE GOVERNANCE

### SENSITIVE DATA

A key market driver behind Privacy Aware Governance (PAG) is sensitive data impacted by recent regulations such as the General Data Protection Regulation (GDPR) in the EU and the California Consumer Privacy Act (CCPA). Governments are seeking to institute safeguards for consumers and protect their information. These regulations mandate control over sensitive data no matter where it resides within a company's ecosystem: databases, website, documents, etc.

This white paper concentrates on managing sensitive data within unstructured data sources, such as SharePoint, file shares, and Box. It's probably no surprise that finding, classifying, and managing personal information is a major challenge. This is true even for documents and records that were originally indexed and put into a content platform, or managed by line-of-business applications. A decade ago, we didn't imagine the need for these identifiers, or foresee that all this data would fall under the personal information umbrella. Now, organizations can automate the handling of personally identifiable information (PII) with machine learning and artificial intelligence.

### MANAGING UNSTRUCTURED CONTENT

Unstructured content is the second key market driver. The vast majority of data is currently unstructured, or from unstructured data sources. It is important to understand that when unstructured content isn't properly identified, it can't be properly managed. Shared drives, SharePoint, Microsoft365, and Box are sources of the largest concern for most organizations when it comes to unstructured data and documents that contain unidentified sensitive information. This problem is made worse when end users try to create a more user-friendly working environment by storing data outside of structured and approved line-of-business applications.

There are a couple issues with using technology that was not designed with these regulations in mind. First, most organizations are playing defense when it comes to their data. This results in them being reactive to new regulations, and spending a lot of time, effort, and money trying to understand their risk and fix problems. Second, most organizations understand the value of digitally transforming their business, but also realize that they will have to deal with technological debt and legacy systems. Finally, stored data should provide key insights that help a company perform better, understand customer needs, and gain market share. Unfortunately, this is very difficult to do when they don't understand the data they have, the risks associated with it, or whether the quality of the data drives these insights. As a result, organizations often wait because they do not have a clear path forward. In the end, what should be a great asset becomes a liability.

Many organizations are at the beginning of their journey around structured data and they're now looking for an all-encompassing, enterprise-wide solution that will cover both structured as well as unstructured data residing on shared drives, Microsoft, and Box. Zia Consulting has worked hand-in-hand with ASG Technologies to create a solution called Privacy Aware Governance (PAG) to assist companies in this journey. Specifically, they've integrated software from BigID, ASG Data Intelligence, and ASG-Mobius Content Services to provide a unified solution that can quickly implement processes to manage documents and records with sensitive data from discovery and access to disposition.

### DECENTRALIZED IT

Decentralized IT is the third market driver. This is where business units and departments purchase solutions without review and approval from IT and Security within their organization. Typically, these business units focus on solving the business problem at hand with little consideration for governance, regulatory considerations, and proper management of consumer and personal information.

## SENSITIVE DATA SPRAWL IN SHARED DRIVES AND ARCHIVES

Scanning and identifying information presents a major challenge when it comes to sensitive data on shared drives. The tools we have historically used to do this rely on looking for patterns in data and are not well suited to handling the variety, complexity, and amount of data we have on share drives and archives.

**Reducing associated risks boils down to understanding two factors:**

**Factor I:** Regulations are based on customer location. For example, businesses in Europe must deal with GDPR, while those in California must be compliant with CCPA. In Brazil, they must follow Lei Geral de Protecção de Dados (LGPD), and so on. Privacy regulations have proliferated in recent years, and businesses are obligated to follow different sets of rules. To reduce this risk, businesses need to understand where a given customer's data resides within the organization so that actions can be taken to comply with location-based regulations.

**Factor II:** The second way to reduce risk is to treat data like an asset and not a liability. An organization's data is one its most valuable assets and organizations cannot maximize on their data if it is locked up and centralized.To avoid compliance issues and start using data like an asset, an organization must answer these questions:

- What data do we have?
- What kind of data is it?
- Whose data is it?
- Where is this data?
- When was the last time the data was cleaned or deleted?

Unidentified sensitive data residing in documents and records can be challenging to discover, classify, and correlate to the correct customer or consumer. Having software that can manage the petabytes of accumulated unstructured data within an organization is a big step in the right direction. Moreover, the software can learn what sensitive data means to the specific organization and find it within documents on shared drives, as well as archives within ASG-Mobius and other content management platforms. This makes it possible for the software to tag, classify, and correlate the data according to the individual or organization, rather than root terms alone.

Leading software solutions, such as BigID, search across structured, unstructured and semi-structured data. The data can be at rest or in-motion, allowing companies to classify sensitive information and assign a risk score. This is a powerful new tool for organizations that solves a very difficult problem.

## DE-RISKING DARK DATA: SHAREPOINT, MICROSOFT 365, TEAMS, AND BOX

Dark data is defined as documents that users leverage outside of their IT team's purview or awareness. Dark data can reside in many locales, with some of the most popular including SharePoint, Microsoft 365, Teams, and Box. Discovering it, and determining what to do with it, is the first step to de-risking this kind of data. Dark data can be deleted or migrated into a secure place like Mobius, which has a rich governance services suite including rules-based redaction, federated redaction, time and geography-based tagging, user permissions, records management, and history tracking.

De-risking IT requires actionable intelligence on the data. It is important to understand which parts of your organization leverages SharePoint, Box, Teams, and Microsoft 365, as well as how the data is being used and stored. Many companies are using outdated strategies that can't keep pace with data or unstructured data proliferation. Without the right tools, the vast majority of an organization's data is never analyzed.

## UNDERSTANDING YOUR OPTIONS

Key benefits and differentiators of a unified BigID, ASG Technologies, and Zia Consulting solution:

- ✓ Data identification and classification
- ✓ Maintaining governance and compliance of data within documents and records across the enterprise
- ✓ Managing unstructured data with the privacy implications of sensitive data
- ✓ Discovering sensitive information residing outside company databases and in data lakes
- ✓ Integration with data intelligence and content services solutions

## STEPS TO PROTECT AND ACT ON REGULATED DATA

The idea of centralizing and securing content is outdated. Organizations thrive on collaborating across functions and business units. The following steps will help protect regulated data without hampering core business functions:

1. Understand the company's regulations,
2. Set a unified strategy, and
3. Define data retention, data masking, and classification policies.

To determine the best place to start the automation journey, look for where the organization has the most sensitive data, such as the finance, legal, or HR departments. Consider whether the department is consumer-facing or has a substantial amount of customer information. Then, take a holistic look at the data that is collected, stored, and protected. Automation is key to this process because no organization has the manpower required to scan and classify all documents via manual efforts. Using an automated solution can ensure that governance steps are aligned with policies for discovering and acting on the data.

The policies within applications like Microsoft Information Protection don't apply across all the shared drives and repositories within an organization. No matter where data resides, Microsoft customers can develop clear and consistent compliance and regulation processes by utilizing BigID solutions. BigID creates a privacy lens of data discovery and looks across all types of data, including structured or unstructured, on-premises or in the cloud. It can identify sensitive information, target and classify that information, and create a metadata catalog. Data stewards then have actionable insights and can set policies for remediation, masking, deletion, or duplication. BigID follows an automated process to de-risk dark data across all your systems, like SharePoint, repositories, and shared drives.

## AUTOMATING ACTION BASED ON CLASSIFICATION

BigID can be implemented to conduct data discovery once the organization unifies a strategy for managing their unstructured data. BigID identifies personal information and sensitive data as defined by the company. It can also perform a cluster analysis, providing a wider perspective on the data. It can discover data duplicates, redundant data, and data that can be more quickly classified if grouped together for action. This means organizations can delete duplicate or redundant data before any movement or migration. From there, information can be fed automatically into a content services platform like ASG-Mobius. Mobius takes action based on the organization's policies. This saves organizations a substantial amount of time on manual efforts and properly adhering to compliance guidelines.

Classification provides an indication that data requires action. An example of this would be to determine if the document includes a social security or a credit card number.  If so, action would be required immediately. BigID assesses the risk based on a sample of the data that it's already accumulated within the organization. If BigID determines that the data needs to have action taken immediately, like the example above, a flag will go up to an administrator. More importantly, this happens in real time, so BigID can continually scan new documents and records that end users add to these monitored unstructured data sources.

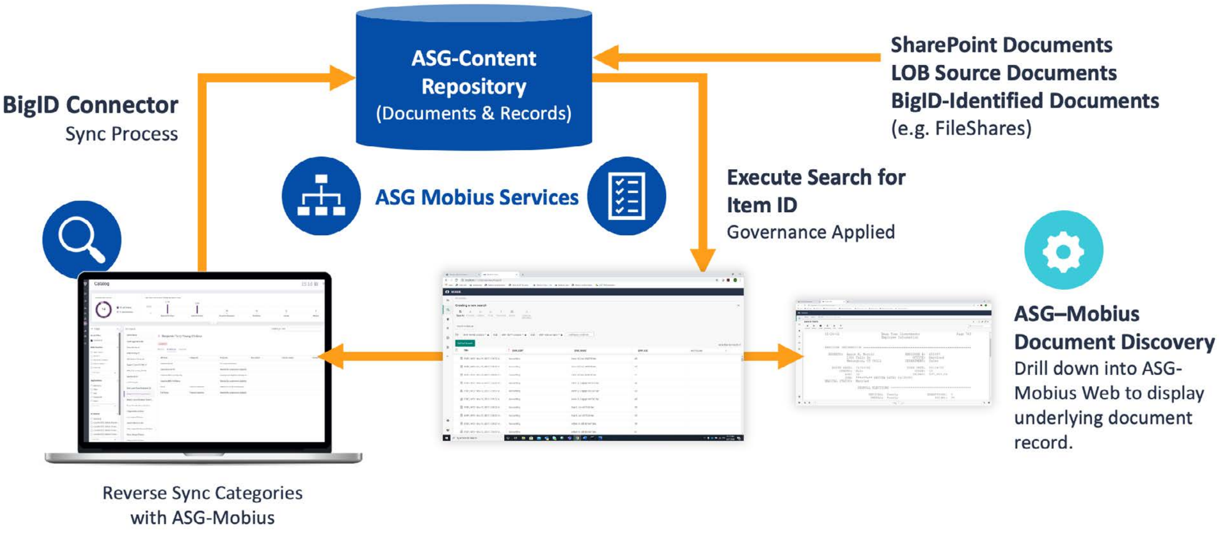
Below, we have laid out the process for identifying, classifying, and cataloging data. We also illustrate the need for migration and governance so that you do not have sensitive data floating around unmasked, unprotected, and unmigrated to a more secure area. Our goal is to highlight both the existence of unmanaged data and the importance of using automation to act on it immediately.

## LOOKING ACROSS ON-PREMISES AND CLOUD SOLUTIONS

Organizations, particularly those in a highly regulated industry, will always have some on-premise data. With the proliferation of cloud applications and storage options, it is important to have a consistent policy across the entire enterprise. Organizations should survey the data landscape, while also creating and enforcing policies that support where the data resides. Further, it should be possible to tag data in order to track when it moves from on-premise to the cloud. This helps organizations understand the data lifecycle management for sensitive personal information.

At times, we see business units unwilling to give up control over existing hardware, which leads to friction with the architects and compliance teams. We offer a solution that is flexible enough to work inside each business unit while aligning with architects' vision and compliance teams' guidelines. This allows the organization to leverage the cloud and on-premises environments, where and how it makes the most sense for their needs.

## A UNIFIED SOLUTION FOR PRIVACY AWARE GOVERNANCE


ASG-Content Repository
(Documents & Records)
SharePoint Documents
LOB Source Documents
BigID-Identified Documents
(e.g. FileShares)
BigID Connector
Sync Process
ASG Mobius Services
Execute Search for Item ID
Governance Applied
ASG–Mobius Document Discovery
Drill down into ASG-Mobius Web to display underlying document record.
Reverse Sync Categories with ASG-Mobius

This scenario illustrates documents and records that reside inside Mobius. Imagine you've been archiving documents and records into Mobius for years, or even decades. Certain pointers with topic indexes and metadata were put in place when the documents and records were originally archived. Later, you realized that the original metadata was not comprehensive enough to address the growing definitions of sensitive information.

Zia's solution grants BigID access to scan and discover sensitive data within documents and records managed by Mobius.

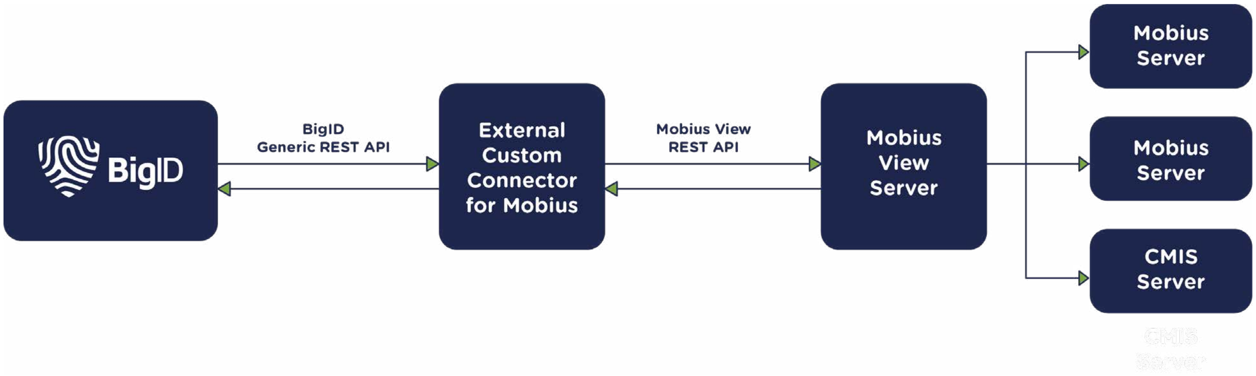
Next, look at SharePoint and Microsoft 365 documents that can be managed immediately within line-of-business applications. Identification and classification of these documents brings awareness to the need for migrating and utilizing event-based retention, redaction, and encryption within Mobius.

Let's take a look at a few pre-built solutions.

**Solution #1: Privacy Aware Governance**
***Scanning Mobius Documents for Sensitive Information***

Our solution for PAG offers the ability to scan within Mobius documents and archives for personal or sensitive information, as defined by your organization. On the right side of the diagram, you'll see a simplistic view of the application and how it can connect to multiple instances of Mobius repositories, or other data sources within your enterprise. As shown in the middle of the diagram, we're connecting the interface through the BigID generic REST API, allowing it to map directly to the Mobius View REST API. This enables connectivity to all the data sources you may have configured throughout your infrastructure.

An architecture like this is relatively straightforward when laid out in a diagram. However, your environments are far from straightforward. Typically, environments evolve over time and become more complex. It is important that you have the ability to configure this solution based on your deployment. Where to deploy a connector, or whether it will sit with proximity to your Mobius deployment, must be considered. We've taken numerous requirements into account as we've built out this solution, making data available to the BigID scan operation and discovery process.

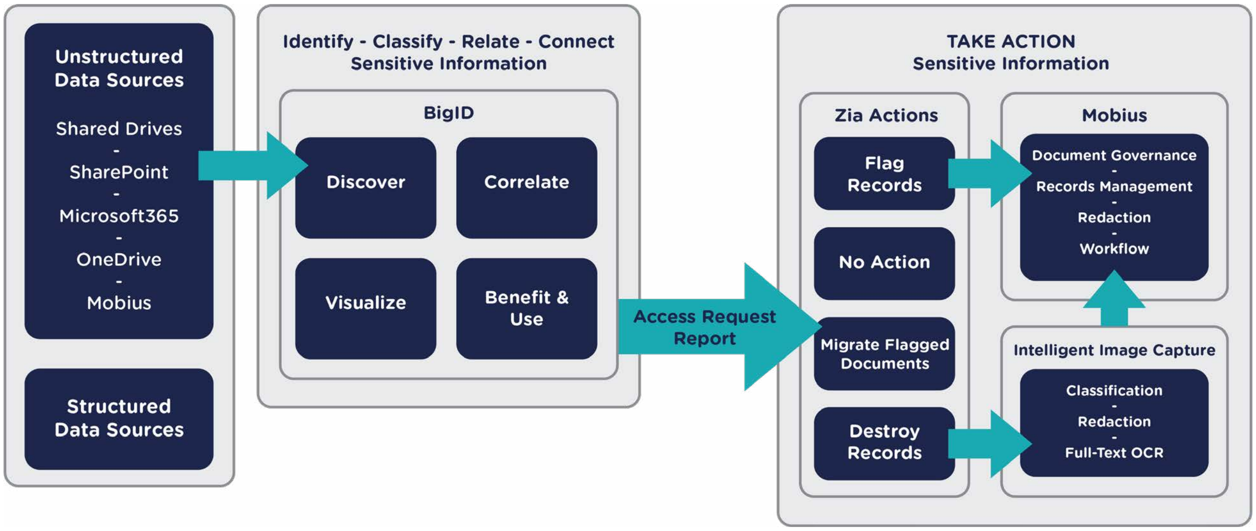


This is commonly where "decentralized IT" is seen. When installations of Mobius were implemented, governance requirements such as GDPR or CCPA may not have been top of mind for its license holders.  Instead, search and retrieval were a priority. Business owners leveraged the technologies to streamline processes and access but didn't intend to address compliance challenges. This is partially because systems were adopted to support departmental initiatives, not a cohesive IT process. We're excited to bring the capability to look into repositories and find inadvertently placed PII and sensitive data.

**Solution #2: Privacy Aware Governance**
***Archive to Mobius - Data Flow***

Now you have completed the discovery and analysis of your unstructured and structured data sources, and associated PII. Next, you need to take action to measure and remediate the gaps, data vulnerabilities, and exposures. As mentioned previously, you need an automation platform that can handle this volume of data and documents. We use ASG-Zenith as the platform to migrate, flag, destroy or federate access to the data exposed in the Access Request Report.

In essence, you can use BigID to discover all the vulnerabilities of PII in the variety of data sources. Then, you can build out automated actions using Robotic Process Automation (RPA), or workflow, to perform tasks against the discovered PII data. While BigID offers remediation tools, Zenith provides the ability to scale the remediation in an automated way. Zenith can verify that the process is providing the intended result for the enterprise and comply with external requests and audits.



Organizations often tell us how important it is for them to be able to identify something as a record and then make choices to act, ignore, delete, or move the data. This would apply to the clusters of data and all the other outputs of an Access Request Report. Governance is not always about the actions you take, but how the data can be used. The goal of this is to make data an asset to our customers, not a liability.

This whitepaper discussed automation as a key aspect of governance because the size and scale of the problem is too large for businesses to solve with manual processes. The best way to improve your automation capacity is to integrate your RPA and workflow processes with an intelligent image capture solution. This works for existing images stored in repositories, as well as new data coming into the organization. Knowing your data, where it is stored, and what it contains drives automation, improves data analysis, and reduces the labor cost associated with data entry, validation, and report generation.

## SUMMARY

This white paper started with a look at the market drivers behind Privacy Aware Governance. We highlighted recent compliance regulations around consumer privacy and personally identifiable information. Further, we noted that when unstructured content isn't properly identified, it can't be properly managed. You may be just starting your journey around consumer privacy and looking for an enterprise-wide solution that will address both structured and unstructured data. You'll want to minimize, and ideally eliminate, shadow IT. The problem is that oftentimes businesses purchase solutions, and leverage hardware or cloud, without central IT being involved.

Next, we shared that vast amounts of sensitive data resides in shared drives and documents archived in Mobius and other content services platforms. Software like BigID looks across all types of data (i.e., structured, unstructured, semi-structured) and in all points of processing (i.e., at rest or in-motion), allowing companies to classify sensitive information and assign a risk score. This is the first time organizations have had this power internally.

Then, we outlined how to de-risk dark data in SharePoint, Microsoft 365, Teams, and Box. Dark data is data that is leveraged by users outside of IT's awareness. We showed the steps an organization can take to protect and act on regulated data, without hampering core business functions. This included understanding the organization's regulations, creating a unified strategy, and developing data retention, masking, and classification policies.

Finally, we noted that our unified solution for Privacy Aware Governance is widely available and outlined two potential solutions. First, we looked at scanning Mobius documents for sensitive information. Second, we discussed archiving in Mobius and using ASG-Zenith's automation capabilities to take action. Both solutions provide data flow that allows for flexibility and customization to suit your organization's needs. We're happy to set up an initial consultation with your organization to gather important details and customize a solution for you. Contact us when you're ready to bring order to your organization's shared drives.

## ABOUT ASG TECHNOLOGIES:

ASG Technologies is an award-winning, industry-recognized and analyst-verified global software company providing the only integrated platform and flexible end-to-end solution for the information-powered enterprise. ASG's Information Management solutions capture, manage, govern and enable companies to understand and support all types of information assets (structured and unstructured) and stay compliant. ASG's IT Systems Management solutions ensure that the systems and infrastructure supporting that information lifecycle are always available and performing as expected. ASG has over 3,500 customers worldwide in top vertical markets including Financial Services, Healthcare, Insurance and Government.

## ABOUT BIG ID:

BigID redefines data privacy and protection by helping organizations know their data for privacy, protection and perspective. BigID uses advanced machine learning and identity intelligence to help enterprises better protect their customer and employee data at petabyte scale. With BigID, enterprises can better manage and protect their customer data, meet data privacy and protection regulations like the CCPA and GDPR, and leverage unmatched coverage for all data across all data stores.

## ABOUT ZIA CONSULTING

Zia Consulting is a provider of industry-leading solutions that automate business processes through streamlined content management. Zia is committed to providing measurable business results on time and within your budget. By connecting you with industry-leading technology partners such as ASG Technologies and BigID, you get the benefits of open source and open standards. In 2020, Zia was awarded Partner of the Year by ASG Technologies for the second year in a row.

## CONTACT US

sales@ziaconsulting.com
ziaconsulting.com
888-732-4101

5525 Central Avenue #200, Boulder, Colorado 80301