Intelligent Data Management in an Era of Explosive **Unstructured Data** Growth

How to access, analyze, and continuously take action on unstructured data.

dryv**IQ**

Table of Contents

Why **unstructured data management** is essential

THE IMPORTANCE OF DATA QUALITY IN A DATA-DRIVEN WORLD

The enterprise data landscape has evolved significantly in recent years, driven by the shift to remote work in 2020 and the rise of generative AI in 2023. A 2024 Wavestone report highlights a sharp increase in data and analytics initiatives in the past year, with a 30% surge in projects aimed at driving business innovation with data and a 24% increase in managing data as a business asset.

As organizations increasingly depend on data to make better decisions and drive business outcomes, the need for trustworthy and reliable data has grown. With the advent of groundbreaking technologies like GenAI, data accuracy and reliability have become non-negotiable.

With 90% of data being unstructured – practically doubling each year – and spread across multiple repositories, effectively managing and maintaining the quality of this data has become progressively challenging, impeding business transformation and creating risk.

But with intelligent data management, organizations can turn the unstructured data challenge into an opportunity, unlocking the full potential of their data to drive sustainable growth and innovation.

WHAT IS THE RISK OF IMPROPERLY MANAGED UNSTRUCTURED DATA?

At its core, data risk is business risk. If organizations cannot access or analyze knowledge worker content, they become susceptible to various risks, from sensitive data exposure and compliance violations to wasted technology investments and missed opportunities to extract value from their data.

With a full understanding of unstructured data, enterprises can reduce their risk and unlock critical business insights.

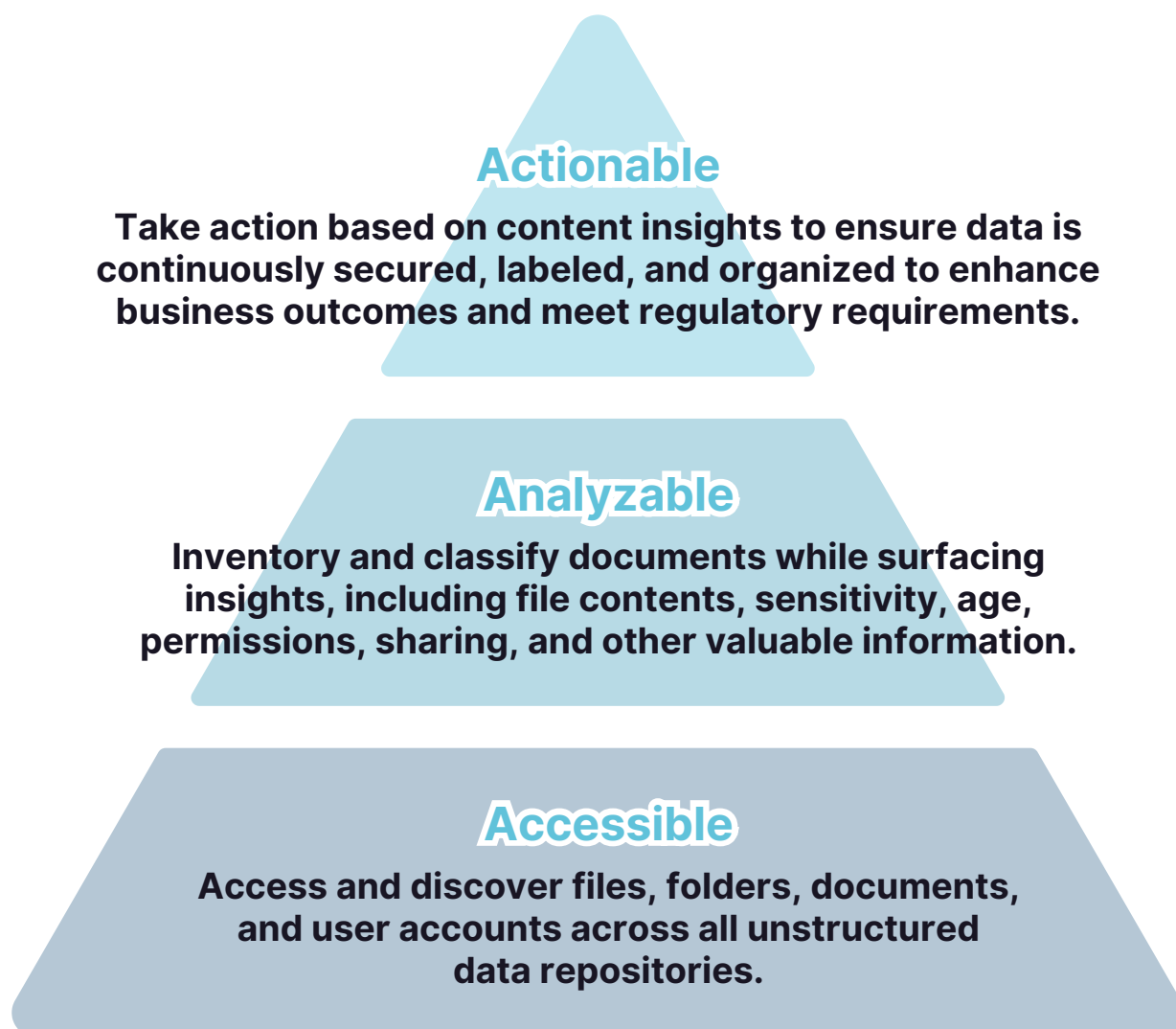# WHAT IS THE RISK OF IMPROPERLY MANAGED UNSTRUCTURED DATA?

Here are the most common risks and consequences that can be caused by a lack of unstructured data management:

Data risk	Business risk
IMPROPERLY SECURED SENSITIVE DATA	If this data is stolen, corrupted, or encrypted it can cause significant financial damage including: business downtime, regulatory fines, loss of reputation/revenue, and costs to repair breach.
DATA EXPOSED BY EMPLOYEES, WHETHER INTENTIONALLY, UNINTENTIONALLY, OR THROUGH GENAI OUTPUTS	A significant threat to data risk comes not only from malicious actors but from internal users. Although they may not intend to cause harm by sharing a payroll list with a third party, for example, the result could be a privacy violation that can lead to reputational damage, financial loss, and regulatory fines. Exposure of sensitive data, whether through employee actions or outputs generated by generative AI, poses a serious risk to the organization.
INABILITY TO LOCATE PERSONALLY IDENTIFIABLE INFORMATION (PII) ACROSS SYSTEMS	Locating data on an individual to comply with regulations like the California Consumer Privacy Act (CCPA) can be a significant drain on resources. Lack of visibility into where or how this data is stored also raises the risk of fines for non-compliance. Even unintentional violations could cost organizations $2,500 per infraction.
HAVING TOO MUCH STALE OR DUPLICATE DATA	There are several risks when it comes to stale data. First, with the explosion of data production, storage costs are increasing both on-premises and in the cloud. Second is the category of data breach and regulatory non-compliance. Just because you aren't using the data, doesn't mean it can't hurt your business if it is exposed—particularly if it's sensitive. And finally, duplicate and stale data can lead to noise, biases, and hallucinations in AI training, compromising the relevance and accuracy of GenAI outputs.
FAILURE TO PRESENT A CLEAR AND COMPREHENSIVE OVERVIEW OF DATA PROTECTION TO AUDITORS	The number of active data security and privacy regulations is growing rapidly and enterprises need to ensure auditors that they're taking steps to maintain compliance. Failure to do so can lead to more significant fines. Since it went into effect in 2018, the GDPR has assessed more than 2,000 fines, worth more more than $4.9 billion cumulatively.

HOW TO OVERCOME THE UNSTRUCTURED DATA CHALLENGE

The sheer volume and variety of unstructured data make it difficult for organizations to access, analyze, and take action on it effectively. From knowledge worker content to invoices to sensitive human resources documents, unstructured data is varied and often scattered across multiple repositories, in the cloud and on-premises. The exponential growth of this data – 50% year-over-year – exacerbates the challenge, as it becomes increasingly difficult to locate and organize data in a meaningful way.

By making complex unstructured data accessible, analyzable, and actionable at scale, organizations can ensure their enterprise data is always ready to support a variety of mission-critical business initiatives. The ability to continuously discover, classify, and manage this data is crucial to increase efficiency, decrease risk, reduce costs, and unlock transformative business value.

Actionable
Take action based on content insights to ensure data is continuously secured, labeled, and organized to enhance business outcomes and meet regulatory requirements.

Analyzable
Inventory and classify documents while surfacing insights, including file contents, sensitivity, age, permissions, sharing, and other valuable information.

Accessible
Access and discover files, folders, documents, and user accounts across all unstructured data repositories.

FOUR KEY BENEFITS OF UNSTRUCTURED DATA MANAGEMENT

1. Access and manage data stored within any system

According to Splunk, 55% of an organization's data is dark data – the unknown data stored across various systems. While data owners are aware of the existence of this dark data, they don't know how much they have, where it's stored, who has access to it, and whether any of it contains sensitive information. Enterprises need better visibility into this mass of largely undiscovered data to properly protect their information and extract key insights.

With the ability to access, identify, and organize data within any repository, organizations can easily refine enterprise content into business-ready data to support data minimization, access control auditing and remediation, data classification for sensitive data protection, retention rule audit and compliance, AI readiness, intelligent migrations, and other data transformation projects.

2. Enhanced data quality and security for GenAI success

Generative AI is quickly changing how we work, and organizations are eager to adopt this technology to boost productivity, growth, operational excellence, and profitability, ultimately delivering more value to customers. But to fully benefit from GenAI without risk, it's vital to ensure trust in the quality and security of the organization's data.

Discovering, classifying, and managing unstructured data across the enterprise is key to preparing data for GenAI. Without tools to automatically manage data accuracy and relevance while securing sensitive information, the integrity of AI training data (and the value GenAI provides) will be at risk. Intelligent data management platforms can help tailor unstructured data to a specific AI model and use case, delivering precise, up-to-date datasets that reduce the risk of using stale, irrelevant, or inaccurate data in AI outputs.

3. Reduced data storage costs and risk

As unstructured data volumes grow, so do storage and management expenses, including server maintenance, storage fees, and data protection software. With cloud vendors (including Microsoft, Box, Google and Dropbox) raising prices and implementing storage limits, storage costs are becoming a concerning line item in IT budgets.

With better visibility, organizations can easily identify stale and duplicate data and implement strategies such as data minimization, file migration, and continuous data classification and management to significantly reduce storage costs. Eliminating unnecessary data also reduces the risk of sensitive information exposure. Unstructured data management gives organizations the ability to proactively manage data growth, optimize storage costs and enhance data security and compliance.

4. The ability to seamlessly handle reorganization events

In large enterprises, reorganizations are common and require their own level of data risk management, but few things are more disruptive than a merger or acquisition. Acquiring organizations need to quickly scan and classify the data from the merged organization to prevent taking on unknown risks. They also need the ability to automate data migration into their environment, especially in a way that can assess and mitigate risk along the way.

In the case of a divestiture, it is critical to identify the enterprise content that stays with the business to prevent the loss of data and intellectual property. In organizations and industries with prevalent M&A activity, implementing intelligent unstructured data management solutions to help with restructuring can decrease risk and increase efficiency with minimal end-user disruption.

What is **intelligent data management?**

Intelligent data management utilizes advanced technologies such as artificial intelligence and automation to discover, classify, and manage enterprise data across an organization's entire landscape. By analyzing files within all storage repositories, these platforms eliminate information silos and provide complete visibility into unstructured data from a single pane of glass. This approach helps organizations make informed decisions, extract value from their data, reduce costs, and improve operational efficiency.

THE THREE KEY FUNCTIONS OF INTELLIGENT DATA MANAGEMENT

1. Unstructured Data Discovery & Classification

To better understand an organization's data, intelligent data management platforms will automatically discover file properties such as:

- Type (like pdf or jpg)
- Size
- Age
- Location (storage system and file hierarchy)
- Internal and external access and sharing permissions
- Custom metadata fields
- Whether the file contains sensitive information

The platform must be integrated across all enterprise content repositories and storage systems to do this discovery. This provides an overview of what type of data you have, where it's stored, and who has access.

Once connected to your repositories, the intelligent data management platform will automatically classify your content by scanning the documents and extracting the text and content for analysis to determine the document type and whether it contains sensitive information. Pre-trained AI classifiers are used to review and compare your unstructured data to known data types, automatically identifying the type of document and file format, assessing its sensitivity, and applying a classification label.

Intelligent data management solutions can easily discover and classify the following:

- Document type (resume, W-2, invoice, etc.)
- PII, including names, ages, addresses, dates of birth, phone numbers, social security numbers, banking information, etc.
- Standard government forms
- Foreign language detection
- Any custom data attributes unique to an organization

This information and other custom insights can be tailored to support any data transformation initiatives.
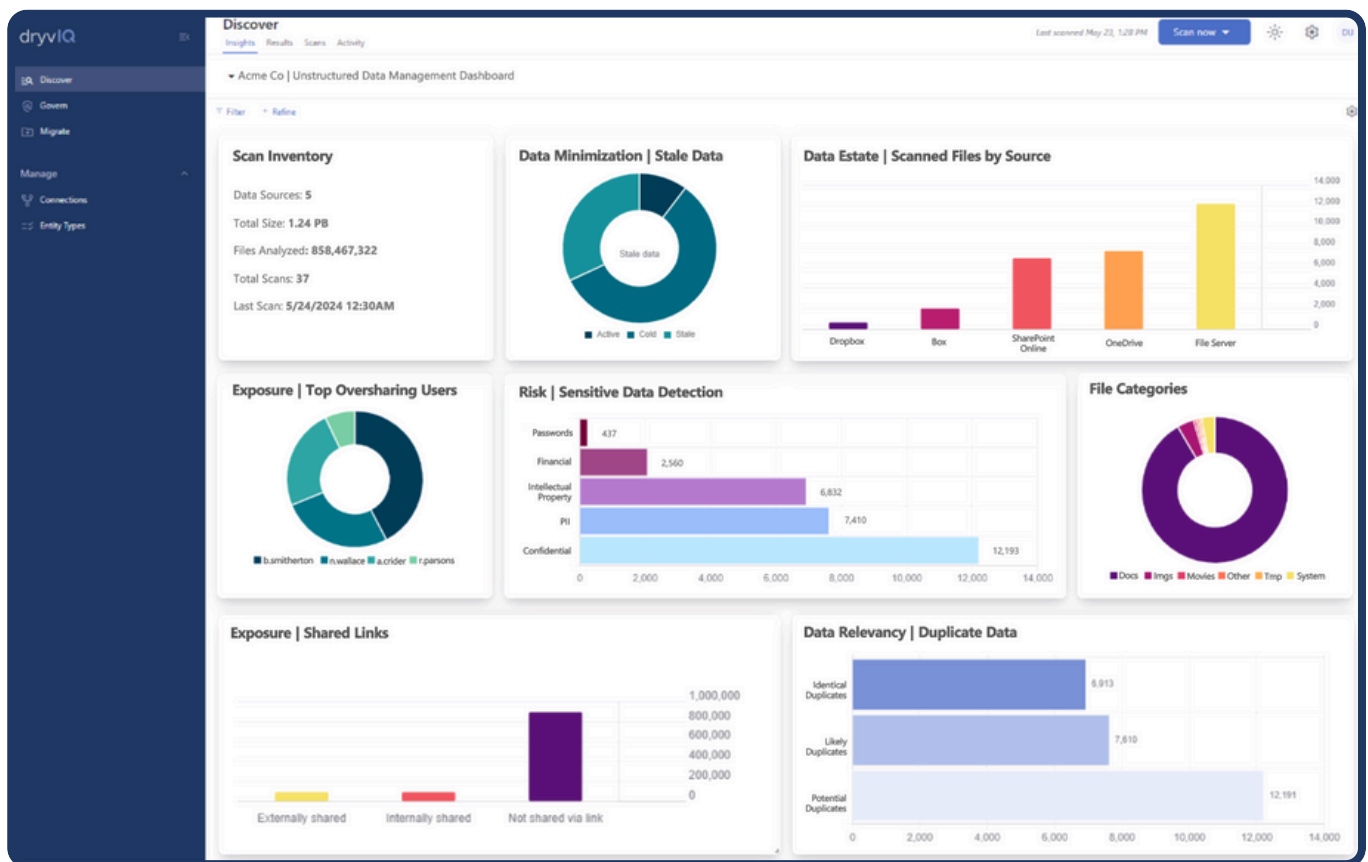
2. Continuous Data Management

Intelligent data management solutions go beyond discovery and classification by automatically and continually performing actions to manage and protect enterprise data.

A series of actions can be taken based on the data discovery and classification results. These actions can be automated as organizations continually generate and update their data. Through activities like transferring data to new locations, re-applying classification labels, or other modifications based on sensitivity and risk, continuous data management enables an always-on approach to managing data quality, risk, and growth.
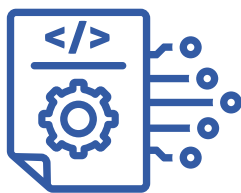
Actions that can be triggered to automatically manage and protect unstructured data include:

- Apply or update classification labels
- Adjust permissions and sharing rights
- Apply or edit metadata
- Quarantine the data for a set period of time
- Archive or delete data
- Any other set of custom actions

Classification

Apply Labels

Apply Metadata

Permissions

Fix Permissions

Remove Shared Links

Data Management

Exclude Stale Data

Exclude Duplicate Data

Learn how AI-driven sensitive data discovery and classification delivered a stronger data security posture at Nicolet National Bank.

Read More

3. Intelligent File Migration

Intelligent migration is a data migration methodology that uses AI-driven data discovery and classification to analyze and organize files before, after, or while transferring them to a new storage platform. This can be done to eliminate redundant, obsolete, or trivial (ROT) data, reduce storage costs, enhance security, and improve data quality. This ability to discover, classify, migrate, and manage data from a single system is a core function of intelligent data management platforms.

When data discovery and classification are integrated into the migration process, the organization can use the discovery insights to make informed migration decisions, and even leverage built-in automation to easily migrate the data based on this analysis.

Simply put, intelligent migration involves surfacing deep insights from your unstructured data to execute a more efficient and lower-risk migration.

Before moving any data, intelligent file migration helps answer questions like:

- How much content are we managing? How old is it?
- What types of documents do we have and are they sensitive?
- Who can access my content? Has it been shared externally?
- Have end users have been properly labeling content manually?
- Do we actually need to migrate all this data?

How to start with **intelligent data management**

IT'S TIME TO GAIN CONTROL OF YOUR ENTERPRISE DATA

In today's data-driven landscape, the challenges posed by the explosive growth of unstructured data are mounting. A strategy to effectively manage this data is no longer "nice to have" – it's essential.

By adopting intelligent data management solutions, companies can transform complex, unstructured data into their most reliable business asset, ready to support crucial initiatives that boost efficiency, mitigate risks, lower costs, and unlock substantial business value.

Make your unstructured data accessible, analyzable, and actionable. Overcome information silos, drive better business outcomes, and unlock the untapped potential of your fastest-growing data set.

Data Discovery	**Data Classification**	**AI-Readiness**
Stale Data Management	**Sensitive Data Protection**	**Data Quality and Security**
File Labeling	**Intelligent Migration**	**File Synchronization**

Not sure where to get started? DryvIQ has helped more than 1,100 organizations worldwide solve some of their most complex unstructured data challenges, and we can help with yours too.

Contact us to take the first step toward effective unstructured data management.

Analyze, Access & Take Action on Your Unstructured Data

DryvIQ's intelligent data management platform transforms unstructured data to drive tangible business outcomes. Trusted by over 1,100 organizations, DryvIQ's patented AI-driven platform ensures data is always business-ready for critical initiatives, from AI readiness to audit preparation, compliance, intelligent migrations, and more.
By making unstructured data accessible, analyzable, and actionable at scale, DryvIQ unlocks the value of unstructured data across 40+ cloud and on-premises systems, driving transformative results and making data the catalyst for business growth and innovation.

Learn More

dryvIQ