THE ERA OF UNSTRUCTURED DATA IS HERE

Datadobi

EXECUTIVE SUMMARY

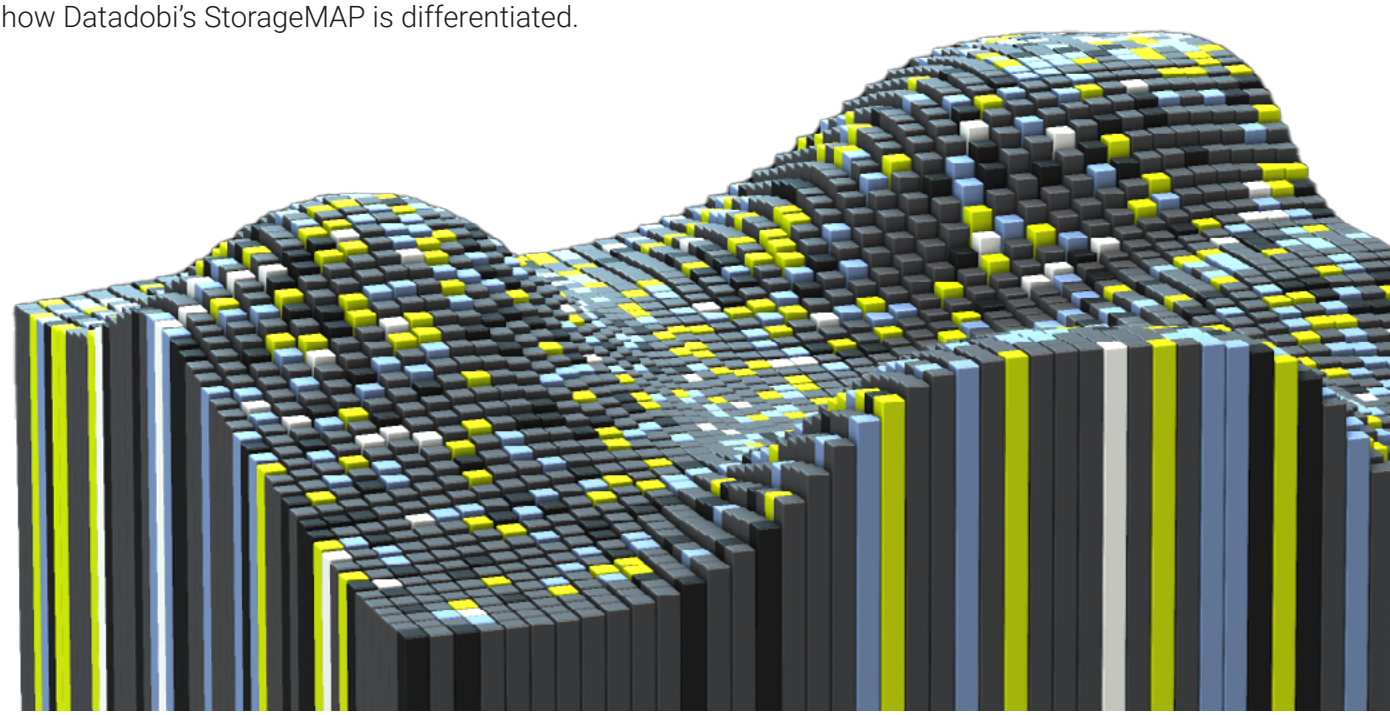
It's no surprise that data growth, whether from human or machine-generated sources, is accelerating. So far, businesses have consistently been able to find ways to store data – initially on-premises, increasingly in cloud environments. Today though, the pace at which data – especially unstructured data – is growing has stretched these strategies to the limit, in large part due to the following trends:

- Costs are rising.
- The environmental impact of corporate decisions such as how and where to store data is widely accepted as good business practice and in some places regulated.
- The growing number of risks associated with data.
- Internal and external customers have higher expectations around availability.

These pressures are forcing organizations to seek solutions that move beyond simply storing data toward truly managing it for business value. There are many dimensions to the challenge, but dynamic data management is finally possible. Effectively addressing the challenge of unstructured data management requires a strategy built on two tenets. First, it is essential to understand what data is stored, no matter how complex and dispersed it is. Second, while keeping cost down is always important, the intrinsic value of data to all aspects of the business requires that the objective of increasing the ROI of data storage be core to any approach.

IN THIS WHITEPAPER, WE:

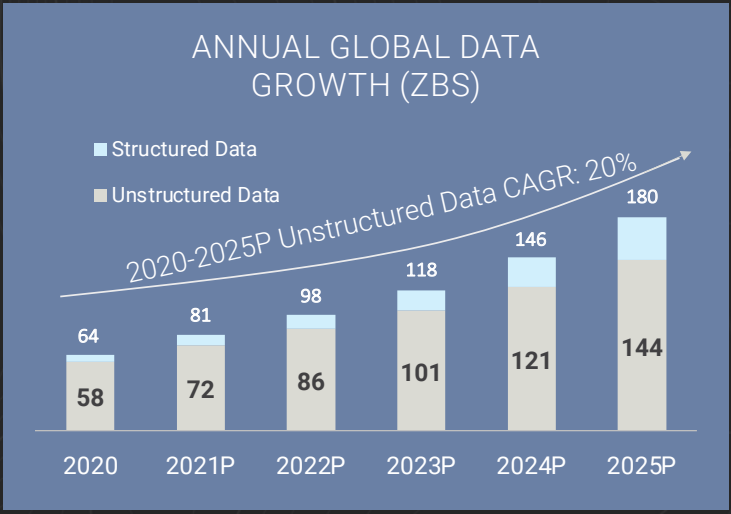
- Explore data growth and the challenges presented.
- Establish what modern data management requires.
- Define the requirements of an effective data management solution.
- Explain how Datadobi's StorageMAP is differentiated.



DATA GROWTH FUELS COST AND COMPLEXITY

Although storage capacities have increased around 1000x every 20 years, capacity—even cloud capacity—is not infinite. One corollary to this rule is that capacity will never be able to keep up with storage requirements. To break out of this cycle of adding capacity, filling it, adding more capacity only to fill it, and so on, a new approach that allows companies to create and store the real data they need is essential.

While structured data stored in databases continues to grow, the growth of unstructured data (file and object) is truly astounding.



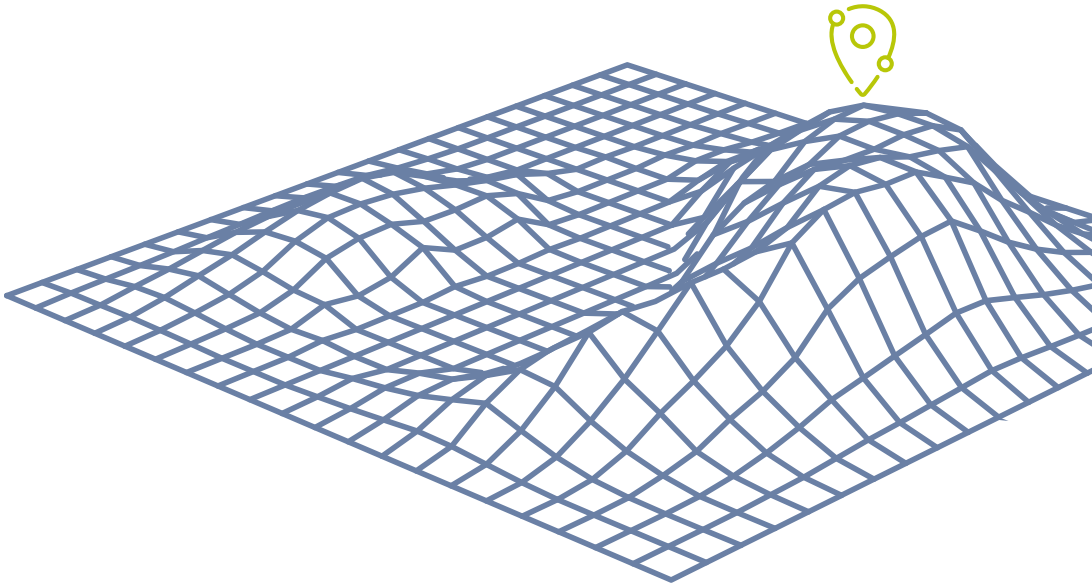
Unstructured data now represents as much as 90% of all data being stored. This data has become mission critical for businesses and as such, requires robust data protection strategies and capabilities. Factor in the protection copies required to ensure that vital information is not lost and in many cases is needed to meet regulatory requirements, and the measurable accumulation of data scales astronomically.

Where to keep all of the unstructured data is one piece of the puzzle. Storing it is expensive and even public cloud storage costs can be high, unpredictable, and introduce vendor potential lock-in that makes a cloud-only strategy difficult to justify. Still, public cloud storage offers flexibility and agility that is difficult to achieve in a pure on-premises model. As a result, many organizations choose to leverage a multi-vendor and hybrid multi-cloud approach with the goal being to maximize benefits in the most cost-effective manner possible.

With data stored in multiple environments, the second issue arises: how to manage it all so the information needed can be located, accessed, used, and then restored to its proper place as needed. As unstructured data rapidly accumulates, its location becomes disparate, and its importance to an organization grows, new challenges are brought to IT organizations, especially the need to know more about the data.

DEEP KNOWLEDGE OF THE DATA STORED ENABLES THE VISIBILITY AND CONTROL THAT IS NECESSARY TO:

- Understand the cost of it being stored.
- Place it in the right location depending on its real value and regulations.
- Understand critical aspects of it, such as aging, level of protection, environmental impact, and more.
- Protect it better and against malicious activities.
- Determine if it poses risk.



MODERN UNSTRUCTURED DATA MANAGEMENT DEFINED

The need to find a better way isn't news to most. In fact, research firm IDC's 2021 "IT Infrastructure for Storage and Data Management Report" found that 72% of surveyed respondents said their organizations are pursuing a common data management strategy.

GigaOm's 2021 report "Building a Modern Data Management Strategy" outlines how an effective data management strategy can answer several infrastructure and business needs. Full visibility of the data stored in on-premises systems and cloud environments makes it possible to:

- Provide trend analysis and advising.
- Find unnecessary duplicates.
- Index and search for data reuse.
- Report on system utilization and user activity.
- Identify orphaned and dark data as well as potential threats.
- Identify and dispose of unnecessary and unwanted data.
- Simplify workflows on data ownership claims and security assessments.
- Identify usage patterns and anomalies.
- Classify data for compliance.
- Provide adequate analytics and reporting to help business and application owners to make decisions quickly.

Strategic data management is increasingly reliant on a new paradigm where data mobility and circulation is a much more common occurrence. In this model, the need to manage data regardless of its location becomes critical, as does the ability to protect the data, and manage its lifecycle in order to minimize cost and risk.

SOLVING THE COMPLEXITY OF UNSTRUCTURED DATA MANAGEMENT WITH THE RIGHT PARTNER

So what should you look for in an effective unstructured data management solution? The complexity of unstructured data management requires that an effective unstructured data management solution exhibit three core capabilities:

- **Visibility**
- **Ability to classify data, plan, and take action**
- **Execute at scale**

The key is to find a solution that addresses the whole process.

And therein lies the challenge. Most solutions today claim to solve the unstructured data management challenge, but in reality only address one or possibly two areas.

There is an answer: Datadobi's StorageMAP is an innovation in unstructured data management that turns data into the asset needed for resiliency, compliance, competitive advantage, and more.

VISIBILITY: GAIN A SINGLE VIEW OF THE ENTIRE DATA LAKE

The first step in creating a data management strategy is to create an inventory of all data and visualize its characteristics such as aging, capacity consumed, category, location, and more. You need a solution that assembles the required visualization regardless of the:

- Location of the data
- Format (file or object)
- Type of platform on which it is stored
- Amount of data

WITH VISIBILITY COMES THE CONTROL YOU NEED TO CLASSIFY DATA, PLAN, AND TAKE ACTION

Once you have visibility into the data landscape, the next step is to plan and take action. You need be able to apply tags that act as an extension of the metadata available from the storage systems. For example, user-defined StorageMAP tags allow you to classify data and report on:

- **OWNERSHIP** – the user or group(s) who are organizationally responsible for a dataset.
- **ROLE** – an identifier showing whether a dataset is departmental data, application data, home directory data, or another special user-defined role.
- **LOCATION AFFINITY** – an identifier noting where the data should be (e.g. low tier on-premises storage for arching or high-compute cloud storage to get it closer to cloud-native analysis apps).

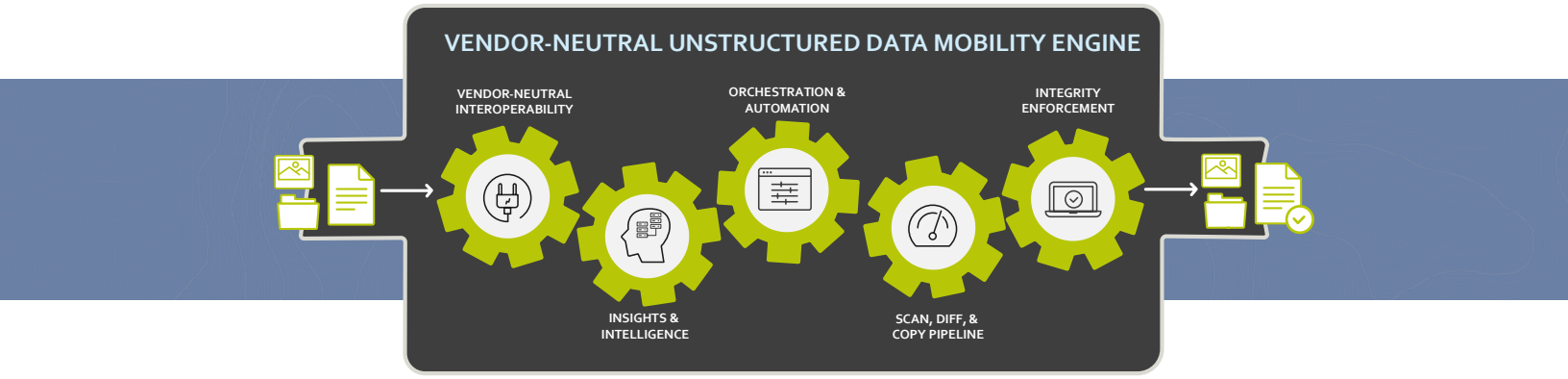
In addition to classification and reporting, you can assign action tags:

- **TYPE OF ACTION TO EXECUTE** – these tags instruct the solution to migrate, replicate, sync, or delete data.

DATA VOLUME REQUIRES THAT YOU CAN EXECUTE AT SCALE

Scale can be viewed in two dimensions: capacity consumed and the number of files/objects stored. While capacity is a simple metric, it is actually the number of files/objects stored in an environment that creates the scaling challenge for unstructured data management solutions. Handling scale is the difference between the ability to execute and the inability to execute in an enterprise environment – and its importance when evaluating and ultimately selecting an unstructured data management solution should not be underestimated.

At the core of StorageMAP lies Datadobi's Unstructured Data Mobility Engine. The engine is built on vendor-neutral interoperability, performance, data integrity verification, and scale. Key intellectual property in the engine empowers StorageMAP to discover, scan, and inventory large datasets with billions of files/objects.



Datadobi's Unstructured Data Mobility Engine provides the basis for data management at scale

Scanning and inventorying extremely large datasets is only the beginning. It is also imperative that actions can be executed at scale. Any data circulation, movement, or protection actions must be performed both at speed and in their entirety. Too many solutions today complete some percentage of the operation and then simply stop, unable to continue and unable to verify the accuracy of its operations. In the case of StorageMAP, even as the engine performs at speed and at scale, it also implements integrity verification algorithms to help protect against data corruption.

THE WAY FORWARD IS HERE

The growth of unstructured data that needs to be managed is not likely to slow down. Nor is the expectation that data be available whenever needed. With Datadobi's StorageMAP, managing unstructured data for business value is possible today – and well into the future.

THE DATADOBI ADVANTAGE: OPERATING AT SCALE, OBSESSED BY DATA INTEGRITY TO STAND OUT

From its inception, Datadobi has focused exclusively on the problem of unstructured data storage and management. Years of research and development have resulted in software solutions proven to work in many of the world's largest and most complex environments. StorageMAP is vendor-neutral and implements a broad body of platform interoperability knowledge gained only through years of intense focus in customer environments.

The software is only part of the story. Datadobi provides customers enterprise-class support via regional-specific resources including a 24/7 help line giving customers direct access to experts who can provide quick and complete help.

Finally, the Datadobi global network of service partners helps customers achieve their data management goal with support that accelerates the implementation, adoption, and effective use of StorageMAP.

VISIT STORAGEMAP.COM TO LEARN MORE

storageMAP