

The Unstructured Data Management Maturity Index

As data management matures, unstructured (file and object) data evolves from being a storage cost center to sitting at the epicenter of value creation. To make use of unstructured data for competitive gain, it's important to develop a strategy for managing it to meet the dual needs of cost efficiency and monetization. This 5-stage maturity model can help organizations looking to modernize unstructured data management practices.



From Data Volumes to Data Value

Enterprise data is growing – no surprise there. It is the current rate of data growth that is truly astounding.

- In 2010, the amount of data created, consumed and stored was 2 zettabytes, according to Statista.
- Firms like IDC have been predicting explosive growth overall in data over the next few years: from 64.2 ZB of data in 2020 to 175 ZB in 2025. That's nearly three times growth in five years.
- Roughly 80% of all data is unstructured: file and object data including documents, medical images, video and audio files, design data, research data and sensor data.

According to **Forrester**, organizations that take a data-driven approach to decision-making **grow more than 30% annually.**



Yet much of this data is not being leveraged – instead straining IT budgets for storing and backing it up.

By some estimates less than 5% of this data is being used for any purpose and enterprise IT teams have minimal visibility into their data and its value. A recent **Accenture study** revealed 68% of companies are not able to realize tangible and valuable benefits from data.

Yet consider the opportunity: from real-time analysis of adverse events to inform patient safety measures and new drug development, early product defect identification in manufacturing, customer sentiment and chat analysis after a new product is released to improve go-to-market strategies or applying machine learning (ML) algorithms to real-time seismic data and satellite imagery to predict natural disasters.

To make use of unstructured data for competitive gain, it's important to develop a strategy for managing it to meet the dual needs of cost efficiency and monetization. Here is a 5-stage maturity model to follow for organizations looking to modernize unstructured data management practices.



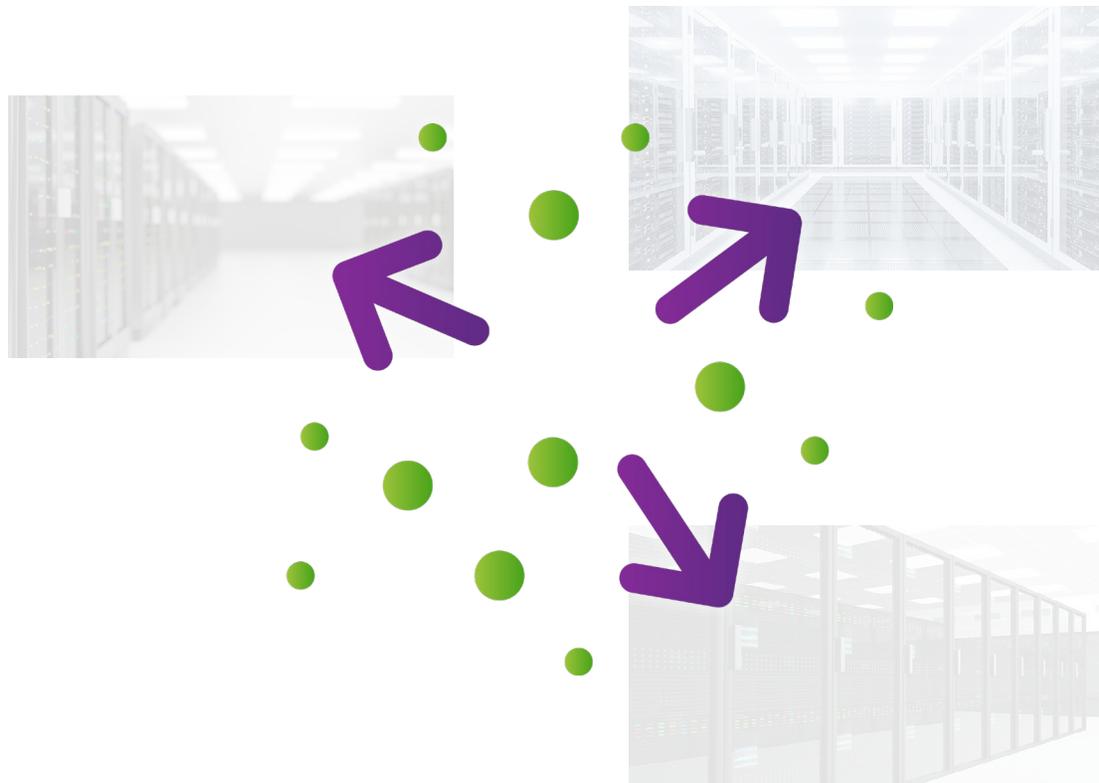
Unmanaged Unstructured Data

In this first stage of maturity, unstructured data volumes are large and distributed across on-premises, edge and cloud silos, resulting in minimal visibility and few if any insights across the entire data storage ecosystem.

In many cases, data is treated the same way: most or all data is on expensive primary storage and not being managed appropriately to save money nor meet the needs of distinct groups and workloads. Meanwhile, there is pressure from above to manage costs, moving away from data center sunk costs on hardware/maintenance to more flexible, on-demand cloud storage. Yet without the proper visibility into data assets, requirements and value, it's difficult for IT and storage professionals to plan and manage effective cloud data migrations. **Many will opt for a basic lift and shift approach, which may actually drive up costs further.**

Key Characteristics:

- Disconnected storage silos limit visibility into data assets.
- Storage, backup and DR costs are high as percentage of IT budget.
- Tension between storage IT professionals and users/department heads regarding data management decisions.
- Lack of expected ROI from cloud storage migrations or tiering.



Storage-Centric Data Management

The second phase is characterized by a move to better control data storage costs by using the storage vendor's own data management capabilities for unstructured data migration, replication and tiering.

Storage-centric data management may be effective in environments with only one storage vendor, but most environments include multiple sites, additional vendors plus cloud deployments. Storage administrators are required to use disparate tools to migrate, replicate and analyze data within these storage silos.

This approach achieves some cost savings but may not lower complexity, reduces flexibility and still leaves money on the table. If an organization wants to access the data after it's been moved to the cloud through the storage vendor's tools, **IT must retain storage capacity and pay egress fees.**

Key Characteristics:

- Unclear strategy for migration to lower-cost storage.
- Multiple tools in use for migration and other data management tasks.
- Hidden costs from storage vendor tiering to the cloud.
- Planned migration to new platforms often behind schedule or delayed due to complexity.



Independent Unstructured Data Management

As an enterprise's unstructured data reaches into the petabytes and beyond and hybrid cloud IT infrastructure dominates, the need to separate data management from storage management becomes apparent. Storage teams will look to adopt an independent data management approach—*sometimes called a data fabric*.

Teams rely on analytics to look across storage silos and identify opportunities for savings. For instance, moving “cold” data not accessed in a year or longer to cheaper storage (such as in the cloud) frees up space on expensive, high-performing NAS storage.

Key Characteristics:

- Consolidation of data management tools.
- IT can manage data apart from the storage technology or service.
- The ability to cut 70% or more of storage and backup costs by identifying and moving cold data to secondary storage.
- The unstructured data management solution should not affect the performance of end-user data access.



Policy-Driven Unstructured Data Management

Organizations in this phase move beyond cost savings to better support security, compliance and research requirements. Data policies and open data formats are critical. **Organizations are automatically and continuously moving data to the right storage based on business priorities, cost, or monetization opportunities.**

For instance, an electric car manufacturer wants to understand how its vehicles perform under different climate conditions and so creates a data management policy to continually pull trace files from cars at regular intervals into data lakes and analyze them. Once the study is over, that policy retires and the moved data is deleted or moved to deep archive storage.

Key Characteristics:

- Storage teams have moved from storage-centric operations to a focus on managing data appropriately throughout its lifecycle with self-service capabilities for users.
- Increased automation to move data to the appropriate storage at the right time, expanding use cases for unstructured data management.
- Data management policies run automatically until changed or deleted, eliminating manual policy management that is error prone.



Unstructured Data Management Value

Some data sets contain value beyond the original application that created it. With advances in scalable, affordable services such as cloud-based data lakes and machine learning, business leaders are eager to see what their troves of stored data might deliver in terms of new insights benefiting R&D, operations and customer relationships.

At this ultimate level of unstructured data management maturity, **the new prize is managing data for long-term value.** Capabilities include the ability to search across storage and cloud silos to find precise data sets and then move the data into cloud analytics environments for access by analysts and data scientists.

Mature organizations can tag files with additional metadata throughout the lifecycle, enhancing possibilities for search and query. Storage teams work closely with business/departmental stakeholders to understand data needs for proper planning and long-term objectives.

Key Characteristics:

- Unstructured data management tools allow for the fluid movement of data into external data analytics platforms and services.
- End-to-end workflow automation eliminates the steps in discovering and delivering unstructured data to the platforms of choice.
- Storage administrators elevate their role from configuring and managing storage technologies to managing data for marketplace gains.
- Data management becomes a flexible framework that future proofs data for new applications and business use cases as they evolve.
- IT can measure increased revenue generated from unstructured data insights.



Key Takeaways

Enterprises have more data than ever to fuel data-driven initiatives across internal processes, R&D and customer relationships. Yet research has shown that less than 5% of this data is being used for any purpose and nearly 70% of organizations are not realizing tangible benefits from their data.

Unmanaged Unstructured Data

Disconnected storage silos limit visibility into data assets and storage and backup costs are high as a percentage of IT budget.

Storage-Centric Data Management

A move to better control data storage costs by using the storage vendor's own data management capabilities, yet complexity reigns and costs are still too high.

Independent Unstructured Data Management

Rely on analytics to look across storage silos, create a data services layer and identify opportunities for savings.

Policy-Driven Unstructured Data Management

IT moves beyond cost savings to better support security, compliance and research requirements and applies policy automation to continuously move data to the appropriate storage for its intended use. Data mobility is now a core principle.

Unstructured Data Management Value

Fluid movement of data into external data analytics platforms and services as well as end-to-end workflow automation to discover and deliver unstructured data to the platforms of choice. Governed self-service helps storage teams align with LOB stakeholders.

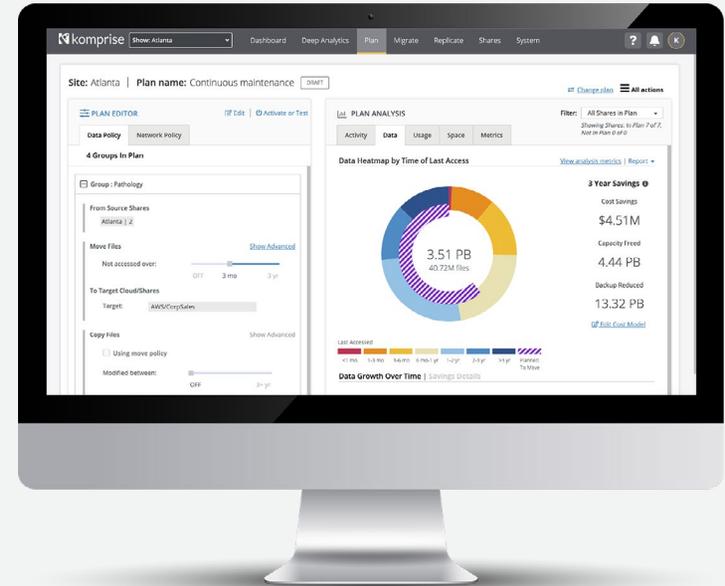


About Komprise

Komprise is a provider of unstructured data management and mobility software that frees enterprises to easily *analyze, mobilize, and monetize* the right file and object data across clouds without shackling data to any vendor.

With Komprise Intelligent Data Management, you can **cut 70% of enterprise storage, backup and cloud costs** while making data easily available to cloud-based data lakes and analytics tools.

[Learn more at komprise.com](https://www.komprise.com)



[SCHEDULE A DEMO](#)