

# SCALABLE, AUDITABLE, AND VENDOR-NEUTRAL UNSTRUCTURED DATA MANAGEMENT

STORAGEMAP GIVES YOU DEEP INSIGHTS INTO YOUR DATA'S KEY CHARACTERISTICS AND ENABLES YOU TO TAKE FAST AND INFORMED ACTION TO OPTIMIZE YOUR DATA STORAGE AND CLOUD ENVIRONMENT.



## KEY BENEFITS

- Provides clear and unambiguous reporting for better stakeholder management.
- Enables hybrid-cloud and avoids storage vendor or cloud lock-in.
- Gets data where it needs to be to extract the most value from it.
- Reduces the risk, cost, and carbon footprint of your storage and cloud estate.

## CHALLENGES OF MANAGING MASSIVE UNSTRUCTURED DATA VOLUMES

Driving new business opportunities at the lowest possible cost, risk, and CO<sub>2</sub> are the challenges of managing unstructured data.

Traditional data storage management solutions are not sufficient to solve these challenges for modern enterprises because they:

- Do not work across products from different storage and cloud vendors.
- Do not span both File and Object storage.
- Provide only basic reporting capabilities.
- Cannot scale to billions of files and hundreds of Petabytes.
- Are not outside of the hot data path.

To meet this growing challenge, organizations are turning to modern data storage management solutions to make sense of their enormous amounts of unstructured data and to ensure that it is bringing value to their organizations at the least possible cost, risk, and CO<sub>2</sub>.

**Actions such as migration, archiving, replicating, data pipelining, and deleting are quickly and easily executed using StorageMAP.**

## COMPLETE VENDOR-NEUTRAL UNSTRUCTURED DATA MANAGEMENT

StorageMAP is a scalable, auditable, and vendor-neutral unstructured data management solution. It is a workflow-driven end-to-end software platform that provides deep insight into the metadata, facilitates organization of the data into manageable data sets, and takes action with a powerful and advanced data mobility engine.

Installation is simple and results are fast. Reports are generated for all stakeholder roles and can be tailored to any criteria specific to the unique needs of the company and industry.

The result is a highly optimized hybrid-cloud environment that helps to ensure unstructured data is realizing its value to the organization.

**A Gartner Strategic Assumption is that 'By 2027, at least 40% of organizations will deploy data storage management solutions for classification, insights and optimization...'**



## SCALE-OUT PLATFORM-AGNOSTIC METADATA INTELLIGENCE AND DATA MOBILITY

StorageMAP's foundation is a dedicated and robust scale-out architecture designed to rapidly scan, index, aggregate, and query thousands of storage servers and services, hundreds of petabytes of data, and billions of files and objects.

StorageMAP is built upon four powerful engines to manage large volumes of unstructured data.

- The *metaData Scanning Engine (mDSE)* collects and aggregates metadata using a parallelized workload that can be distributed across an arbitrary number of "workers" in the storage environment.
- The *metaData Query Language (mDQL)* analyzes and presents insights into the large amounts of metadata collected during the scanning.
- The *unstructured Data Workflow Engine (uDWE)* is where actions are configured, scheduled, and dispatched to the processing pipeline to execute and log results.
- The *unstructured Data Mobility Engine (uDME)* moves, copies, replicates, and verifies large and complex unstructured datasets based on trends and characteristics derived from the metadata intelligence.

As the amount of unstructured data under management grows, StorageMAP increases the parallelism resulting in a performance profile that scales with the data and that is only limited by the storage systems themselves.

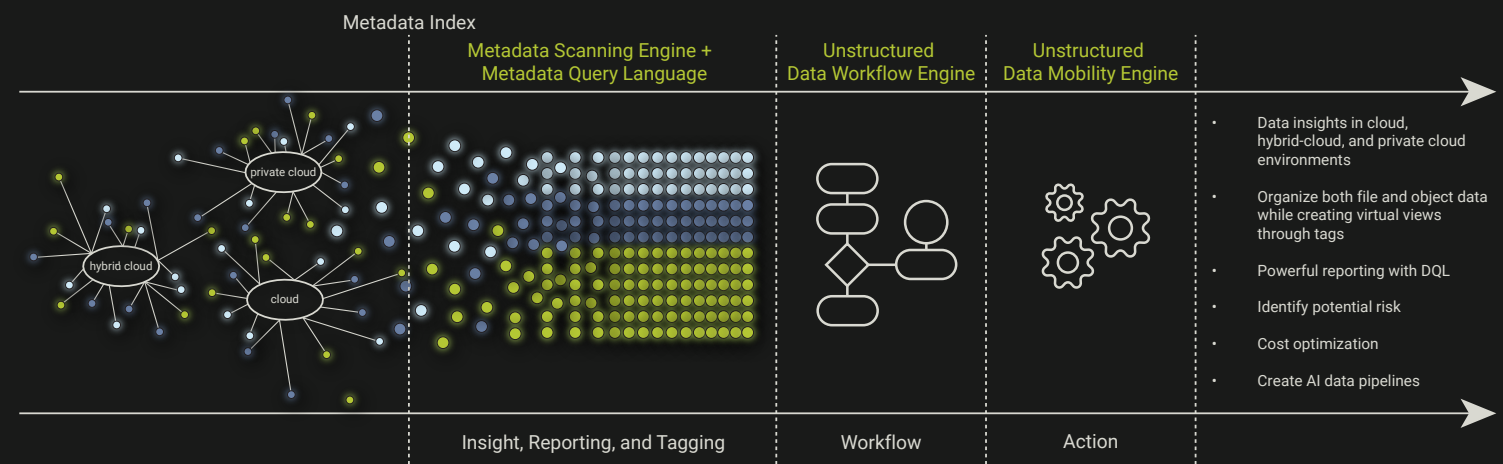
StorageMAP is platform and vendor-agnostic since it uses standardized NAS and Object storage protocols to access the data. It does not require agents on the storage systems, nor does StorageMAP reside in the data path between the application and the storage server eliminating vendor lock-in.

## ADVANCED INTEGRITY PROTECTION (AIP)

The StorageMAP product team considers the integrity and security of data the highest priority. Extremely defensive software coding has generated technologies and methods that work in concert with each other to protect data under management:

- **File Level Verification** – Verifies the integrity of every file or object copy through a strong cryptographic content digest comparison.
- **Advanced Integrity Protection** – Prevents any form of data corruption by detecting unexpected data modifications during data migration or data copy actions.
- **Chain-of-Custody** – Generates a "Flight Recorder" for the data copies that proves that the destination content is the same as the source.
- **User Error Mitigation** – Detects many common user mistakes and misconfigurations in order to inform users about actions which could lead to unwanted results.

## SOFTWARE ARCHITECTURE

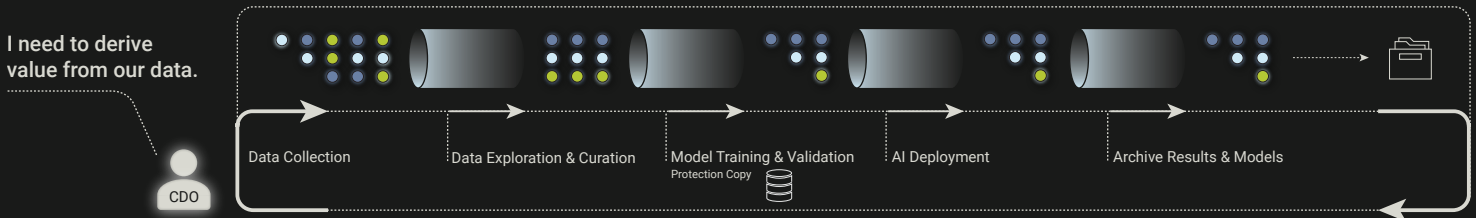


## USE CASES

StorageMAP is used by enterprises and government agencies for a range of business and technical use cases such as full or partial system migration, cloud migration, data repatriation, replication of Golden Copies, search-based moves, and search-based copies. Illustrated below are three examples.

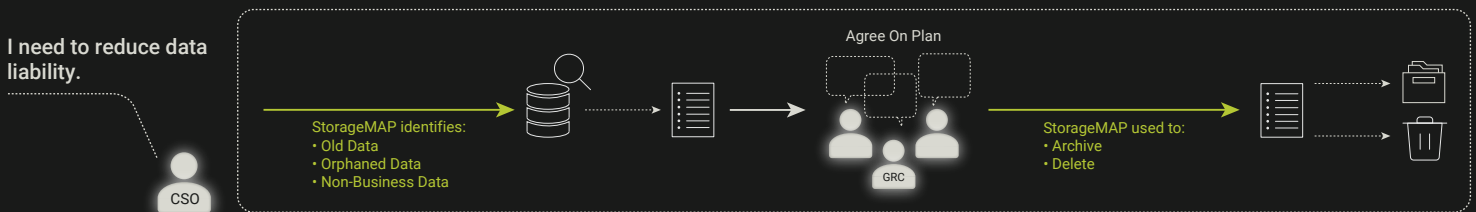
### Data Pipelines

Create data pipelines to allow maximum value to be derived from data (such as with AI/ML). Data pipelines first analyze and then visualize the massive amounts of unstructured to help determine which data should be sent through the pipeline. The data is then relocated to a new location where it can be analyzed or processed further. For example, data to be used for generative AI processing can be relocated and verified through a pipeline enabling curation, training set creation, and LLM creation.



### Redundant, Obsolete, Trivial (ROT) Data

Find ROT data and remove or archive. ROT data results in sub-optimal use of assets along with an increase in risk due to data being stored in perpetuity. Quickly find and tag data that is not relevant to the business, data that is orphaned, data that has aged beyond reasonable usefulness, or compliance data that no longer needs to be retained per industry regulations. Quickly deal with ROT data through actions such as moving the data to lower cost storage, moving data to a quarantine or review area, making archive copies, or even deleting the unnecessary data.



### Financial Cost and CO<sub>2</sub> Emission Reporting (Chargeback/Showback and Sustainability)

Track and report financial costs and carbon emissions associated with on-premises and cloud data storage. Financial costs illustrate the monthly cost to the organization while the carbon reporting indicates the CO<sub>2</sub> emissions associated with the data being stored. Reporting is available at any desired level of detail such as the organization's global storage infrastructure, individual file servers and S3 endpoints, file systems within a server, S3 buckets within an endpoint, and even specific portions of file systems and S3 buckets.



## TECHNICAL SPECIFICATIONS

ANALYTICS & REPORTING	
Search	Interactive metadata search with downloadable results
Advanced Data Querying	Yes, using meta-Data Query Language (mDQL)
Managing Logical Data Sets	Yes, using Tags
Executive (C-level) Reporting	Yes
Cost & Carbon Reporting	Yes
Storage Admin Reporting	Yes
Technical Reporting	Yes
ACTIONS	
Data Mobility	Migration, Move, Copy, Delete, Pipeline
Data Protection	Replication, Third Copy
Multi-Vendor	Yes, source and target can be different storage platform types
Hybrid Cloud	Yes, both datacenter and cloud
Multi-Protocol	Yes, for data accessed simultaneously using SMB and NFS
Protocol Change	SMB-to-S3 and NFS-to-S3
Data Under Retention	Failsafe workflows for handling and migrating WORM data
Data Identity Mapping	Yes, both for SMB SIDs and NFS UIDs/GIDs
DATA INTEGRITY	
Advanced Integrity Protection	Yes
Chain of Custody	Yes, available for all data mobility actions
SCALE OF DATA UNDER MANAGEMENT	
Storage Capacity	Unlimited
File Count and Object Count	Unlimited
Storage Systems	Maximum 1,000
STORAGE PROTOCOLS	
Data Center Storage	Any SMB/NFS file server or S3 object server, including vendors such as Dell, Hitachi, HP, IBM, Microsoft, NetApp, Pure, VAST, etc.
Cloud Storage	Any SMB/NFS/S3 storage cloud. Including AWS, Azure, Google Cloud, Wasabi, etc.
SMB Protocols	CIFS, SMB 2.0, SMB 2.1, SMB 3.0, SMB 3.0.2, SMB 3.1.1
NFS Protocols	NFSv3, NFSv4.0, NFSv4.1, NFSv4.2
Object Protocols	S3
Protocol Conversions	SMB and/or NFS to S3
DEPLOYMENT	
Software Deployment	Linux RPM running on VM (1 core and up to 32 proxies)
One-Click Upgrade	Yes
GUI	Yes (using Web Browser)
API	Yes (REST API)
Agents	No, agents are not required on the storage systems
Can Run on Isolated Network	Yes, no cloud or internet connectivity required
No Unmanaged Data Copies	StorageMAP never stores a copy of data in transit or under analysis
Authentication	Built-in Users, Active Directory or LDAP Users, API Tokens
Authorization (roles)	Admin, expert, user
Full Audit Logs	Yes

## UNSTRUCTURED DATA MANAGEMENT SOFTWARE THAT REDEFINES STORAGE

StorageMAP gives insight into your unstructured data and helps you to get the most value at the least possible cost, risk, and CO<sub>2</sub>.

With StorageMAP, your enterprise or cloud will be efficient, agile, and resilient – always, automatically, and continually. Best of all, with StorageMAP, your entire data landscape adapts dynamically at the speed of change, directly in line with your business needs.

## SUPPORT AND SERVICES

DatadobiDriven is a comprehensive support and services offering designed to ensure a positive outcome for every Datadobi customer.

DatadobiDriven integrates system installation and setup, administrator training, 24x7 technical support, and professional services.

The DatadobiDriven team of enterprise unstructured data and storage experts works day and night to help you optimize your storage environment.

## ABOUT DATADOBI

Datadobi, the global leader in vendor-neutral unstructured data management, brings order to heterogeneous unstructured storage and hybrid-cloud environments via its StorageMAP platform.

Datadobi helps enterprises manage unstructured data growth through the power of visualization, organization, and action in a single pane of glass. Founded in 2010, Datadobi is a privately held company headquartered in Leuven, Belgium, with subsidiaries in New York, Melbourne, Dusseldorf, and London.

