

GET YOUR *UNSTRUCTURED* DATA GenAI READY AND REAP MANY OTHER BENEFITS WITH STORAGE MAP



UNSTRUCTURED DATA CAN REPRESENT 90% OR MORE OF ALL DATA IN AN ENTERPRISE, SO YOU SHOULD BE GETTING VALUE FROM IT AT THE LEAST POSSIBLE COST, RISK, AND CARBON FOOTPRINT.

Unstructured data contains extremely valuable information such as customer data, financial reports, call recordings, and legal agreements that can all help an organization to become more efficient and competitive.

Until now, organizing and extracting value from billions of files that have formats as diverse as PDFs, sensor data, sound files, and images has required overwhelming manual effort which meant unlocking its value remained elusive.

With the rise of GenAI, the ambition of unlocking this value is now within reach. But GenAI is not auto-magic. It requires training on quality data and that means weeding out the garbage in billions of files.

With StorageMAP's unsurpassed ability to visualize and manage large scale unstructured data, the path to unlocking its hidden value is now clear.

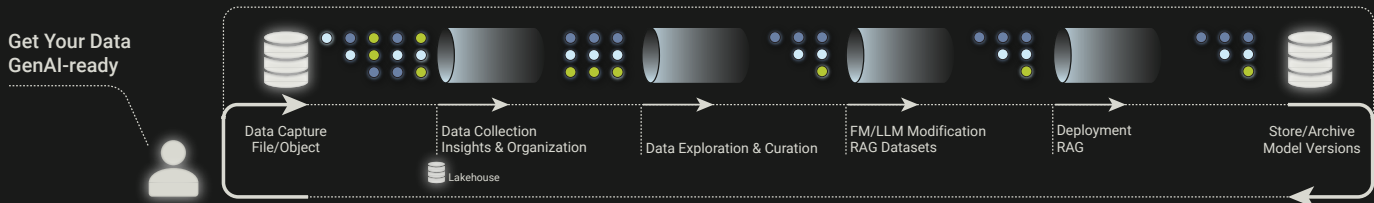


Figure 1 | The GenAI Pipeline For Unstructured Data

GET YOUR DATA GenAI READY

The GenAI pipeline starts with feeding the data lake with quality data. So, the key to getting GenAI ready with unstructured data, and reaping many other benefits, is to get an enterprise-wide view of all data to determine what has potential value.



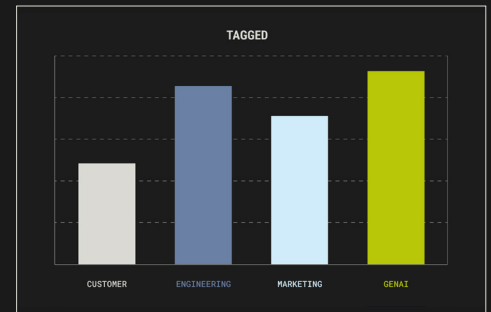
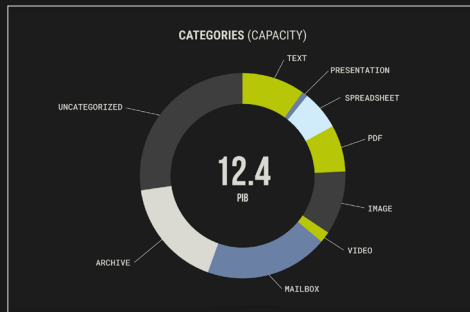
Figure 2 | Get Unstructured Data GenAI Ready

SOLUTION BRIEF

The detailed picture created by StorageMAP (see below) can be used to communicate with data scientists and other stakeholders speeding up the identification of the right data to train GenAI models.

Some decisions will involve the governance of data such as whether it is appropriate for use in GenAI applications. In other cases, more granular criteria are required such as specific applications, projects, age, ownership, and other characteristics that data scientists would like to focus on.

Once the data is visualized, identified, and tagged, it is then copied by StorageMAP to the data lake to begin the exploration and curation stage of the pipeline.



MOVING AND PROTECTING DATA THROUGH THE GenAI PIPELINE

As data moves through the GenAI pipeline, easy and safe data mobility is essential. Millions or billions of files will need to be regularly moved or copied to the systems used to further process the data and get it ready to train GenAI models.

StorageMAP is built on the world's leading unstructured data mobility engine used in thousands of enterprises and government agencies in the most demanding environments to make large-scale data mobility simple. StorageMAP contains intelligent connectors to over 60 unstructured data storage and cloud platforms that use disparate and incompatible configurations.

Alongside the pipeline, there is a need to protect data throughout its lifecycle. Curating data, creating models, or training existing models can be expensive so protecting them is paramount. StorageMAP helps to create a golden copy of this data on the heterogeneous systems needed in any healthy data protection strategy.

THE MANY OTHER BENEFITS OF GETTING GenAI READY

As unstructured data continues to mount, in many cases exponentially, most organizations are still grappling with the basic challenges of keeping costs under control, reducing risk, and managing their carbon footprint all with shrinking resources.

Luckily, getting GenAI ready with StorageMAP is also the first step in achieving these fundamental challenges.

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 insights, 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.

