



White paper

Digital pathology storage: strategies and considerations



Summary

Digital pathology offers many benefits, but image storage creates a major challenge. Healthcare organizations are weighing concerns about storage cost, scalability, security, and accessibility. Here, we examine three approaches and measure them against key considerations. We detail how a trusted cloud storage service provider can help healthcare organizations simplify, secure, scale, and optimize their digital pathology operations with confidence.

Contents

- 3/ Executive summary
- 3/ Unlock value from your data
- 4/ Sizing up the storage issue
- 5/ Three patterns in storage strategy
- 5/ Option 1: An on-premises approach
- 5/ Option 2: A hybrid model
- 6/ Option 3: Outsourcing to professionals
- 8/ Understanding storage tiers
- 9/ Consider Iron Cloud digital pathology storage

Executive summary

Technology innovations are keeping pace with and, in some cases, driving the digital transformation of pathology image analysis.

Healthcare professionals have noted the [significant benefits](#) of this new way of practicing pathology – including faster access to important information, accelerated collaboration, and greater opportunities for consultation.

Organizations that have already adopted digital pathology have reduced their handling of glass slides and enhanced their workflows with algorithm-based assistance. Those with advanced implementations have also been able to monetize data that was once trapped in large sets of stored glass slides. With each stage of digital pathology adoption, the potential to support research and education grows.

But there's one major challenge. ***How and where will all the data be stored?***

The answer isn't simple, and there's no one-size-fits-all approach. However, some facts do apply to every digital pathology practice:

- Healthcare organizations need storage models with enough scalable capacity to manage growing volumes of data.
- The data must be available as needed without too much complexity in the storage system's pricing or access.
- The storage environment must be secure to control access and protect personal health information.

Unlock the value from your data

Once in digital form, digital pathology image data can be used to train artificial intelligence models, and identifying specific patterns for analysis, verification, and patient care collaboration.

Smaller sets of images, even individual images, can be made available anonymized with consent to other entities. In these instances of approved data sharing, there is a monetization opportunity.

Compensation can be monetary, or it can take other forms. For example, researchers may trade access to data sets to further discovery, and data can be used to enhance education.

Storage strategies emerging now range from on-premises to hybrid to the cloud. Here, we investigate these solution options and identify key considerations for digital pathology storage.

Sizing up the storage issue

Industry [experts estimate](#) a typical digital pathology practice will need to consume at least a petabyte of storage per year. To [fathom a petabyte](#), imagine 20 million filing cabinets filled with 500 billion pages of information. It's enough storage to hold 223,101 DVD-quality movies.

As digital pathology adopters look at the amount of storage required, it's not surprising that digital image storage becomes one of their top challenges.

Pathologists prefer to work with high-resolution whole slide images, which are between 2 and 4 gigabytes each, approximately the file size of one full-length feature film in high definition (HD). Many hospitals and labs will produce [700 to 1,000 slides per day](#), generating an average of 3,000 gigabytes, or 3 terabytes, of new data each day. This adds up as 1,000 terabytes amounts to 1 petabyte.

It's clear why the storage of digital image files leads the list of concerns with the massive amount of data. The information also needs to be stored in a way that is secure and accessible. The burden is tremendous, especially for an organization whose primary focus is healthcare.

Three patterns in storage strategy

Healthcare organizations are considering the pros and cons of three approaches to digital pathology image storage.

Option 1: An on-premises approach

For small-scale operations with limited scanning needs, scanning and storing on-premises is feasible. The organization may require a single scanner to process a moderate number of slides. For example, 50 slides per day would generate 80 gigabytes of data per day. The PC for the scanner could potentially store 22 terabytes—enough to hold a year’s worth of data. The drive could be swapped out each year to keep the operation running.

Data storage of diagnostically used whole slide image data on consumer hard drives is inadequate, and data storage on professionally maintained and secured servers is mandated.

Journal of Clinical Medicine

In this scenario, there could be difficulties with image management, such as security vulnerabilities related to protected health information and limited scalability. The organization would also need to consider how to handle redundant and offsite file storage requirements.

This approach involves considerable staffing resources, capital costs, and onsite IT expertise. Due to the listed challenges, there are better options for moderate- to large-scale operations and/or those that need to quickly establish a cost-effective digital pathology practice.

Challenges of the on-premises approach:

- Risks are higher as data storage is not redundant, secure, or in the cloud.
- Accessing the data may interfere with concurrent scanning operations.

- Dismounted drives are not accessible and can pose security vulnerabilities if not properly encrypted and managed.
- The solution is not readily scalable should scanning volumes increase.
- Data security and the protection of personal health information may not be assured.
- The scanning equipment consumes valuable local floor space and human resources.

Source: *Ten Simple Rules for Data Storage*

Any organization will need expert help in addressing each of the challenges. As the digital pathology system’s complexity increases, so do capital costs and the amount of IT expert time required to design, test, implement, and maintain the solution. There can be added difficulties if or when the operation requires scaling up, even moderately.

Option 2: A hybrid model

It seems simple enough to scan files locally and store them in the cloud. Many organizations are **embracing the cloud** for medical image storage. This hybrid solution resolves many problems an on-premises model poses, including the need for more redundant, offsite storage.

While the cost of storage continues to decrease, the volume of data to be stored impacts the choice of storage methods and locations: for large datasets, it is necessary to balance the cost of storage with the time of access and costs of regenerating the data.

PLOS Computational Biology

Cloud storage providers typically have [various data redundancy techniques](#). Cloud storage is essentially limitless, which is beneficial, but not free. The organizational costs will need to be modeled and understood for scalability.

This approach places high demands on local IT and network resources, and requires secure data connections. There are many [security concerns](#) related to public cloud storage, including data loss and unauthorized access. There could also be hefty egress fees imposed by cloud providers and complex pricing structures associated with moving data among storage tiers.

Challenges of the hybrid approach:

- Adequate data storage security may not be assured by the cloud storage provider, putting personal health information at risk and demanding [data protection techniques](#) be in place.
- [Data in motion](#) must be secured, meaning the organization must design, test, and maintain a secure data connection between the scanning location and the cloud storage service provider.
- Valuable local floor space and human resources are still consumed by scanning tasks and equipment.
- Moving large amounts of data to the cloud places high demand on local IT and network connectivity resources and will likely require infrastructure upgrades, delaying deployment time and increasing operational and capital costs.
- In some cases, [high overall storage](#) costs exceed the cost of analysis or even regenerating the data (when possible) from scratch.
- Unknown future data access volumes and timeframes make it difficult to budget for future storage costs. The problem can be compounded by changes in the cloud storage provider's fee structure.
- Many cloud storage providers charge hefty data egress fees, but this can be mitigated by carefully managing data across tiers of cloud storage.
- Financially optimized archival storage can be complex, requiring active expert management.

Though the hybrid option has flexibility, it should be carefully evaluated for fit with an organization's specific needs, budget, and security requirements. Organizations must assess potential cloud storage providers in terms of security assurances and other contractual details of the service agreement. Unless the requirements are met, this approach is not viable.

Option 3: Service provider cloud storage

The primary concern of any healthcare organization is the health and well-being of patients. Thus, many are opting to outsource the entire digital pathology image storage process. That means a trusted professional will provide redundant, secure, cloud storage and be responsible for efficiently scanning slides and managing secure data access.

In this approach, a cloud storage service provider collects pathology slides and moves them to a secure, temperature-monitored, offsite storage facility. The provider will access and scan specific slide sets at the organization's request, minimizing transport and maximizing security. The slides are returned to storage after scanning.

An organization can cut costs, simplify operations, and improve security by putting the right data in the right place for the right price.

The provider safeguards image data at one of its secure data centers – offering cost-effective storage at a reliable rate, and delivering security and data privacy requirements. Healthcare professionals can return their focus to patient care, side-stepping many of the hassles associated with digital pathology adoption.

Storage challenges and solutions

Consider how an organization can address storage challenges by employing a cloud storage service optimized for pathology slides and images:

Challenge	Solution
Redundant, secure, offsite data storage.	A pathology-optimized cloud storage provider has secure, redundant multi-location storage.
Data access interference with concurrent scanning operations	The provider manages scanning, access, and storage completely.
Dismounted drives are not accessible and can pose security vulnerabilities if not properly encrypted and managed.	There are no drives to dismount and manage.
The storage solution is not readily scalable.	Pathology-optimized cloud storage is essentially limitless. High-quality providers will offer a diminishing per-unit cost as overall storage volume increases.
Data security in motion and in storage is not assured.	A provider whose business is built around secure protection of customer assets, with decades of experience managing healthcare data, can meet security needs and regulations.
The scanning equipment consumes valuable local floor space and human resources.	Housing, operating, and maintaining offsite scanning equipment is now the responsibility of a qualified provider.
Moving large amounts of data to the cloud places high demand on local IT and network connectivity resources.	The service provider handles this concern instead of the healthcare organization's IT department.
Public cloud storage comes with security concerns.	By scanning in their secure facilities and storing information in their secure data centers, a service provider can alleviate concerns related to public cloud storage.
Many cloud storage providers charge hefty data egress fees, or offer complicated tier structures.	A pathology-optimized storage service can help eliminate unpredictable future cost problems by not charging data egress fees, simplifying the process with built-in tiered management.

Employing a service provider for cloud storage answers the challenges associated with on-premises or hybrid options. Yet, organizations are cautioned to select a pathology storage provider with a solid history of securely scanning, managing, and storing medical information and pathology data.

The bottom line is that collaborating with a cloud storage service provider is the smartest, safest, most cost-effective way for a healthcare organization to realize the many benefits of digital pathology without making substantial investments of time, money, space, and staff resources.

Understanding storage tiers

The Iron Cloud Digital Pathology Storage solution from trusted provider **Iron Mountain** combines multi-tiered storage options with on-demand pricing that includes online access, data protection, and archiving.

- ▶ Active/hot tier data is fully online and available for immediate access, held in the system for a set number of days determined by your organization.
- ▶ Warm tier data is migrated from the active/hot tier, which is still available online and stored for a length of time determined by your organization but is less likely to be needed immediately.
- ▶ Cold tier data has aged out of the Active/hot tier or warm tier after a number of days set by your organization. This data is offline and protected by an air gap, stored for up to ten years at a reduced cost, and accessible within 12 to 24 hours.

How it works: Healthcare organizations can work with Iron Mountain to configure timelines and criteria for each storage tier. Pathologists and others with access can request and retrieve data immediately from the active/hot tier, which uses the industry standard S3 protocol. Iron Mountain creates two identical copies of digital slides from the active/hot tier that are stored in the warm tier, and the cold tier.

Warm tier data can be retrieved within minutes to hours. Cold tier data is kept as archival to meet compliance and other requirements. Iron Mountain manages the data lifecycle policy, maintaining multiple copies of data stored in various locations even after the data ages out of the active/hot tier.

This high-availability model will scale as an organization's digital images expand and flex as data storage needs shift. The multi-tiered model is favorable compared to hyper-scaling, which may keep all data in active/hot storage at a higher cost.

Watch the video to learn more about multi-tier Iron Cloud storage from Iron Mountain.

Consider Iron Cloud digital pathology storage

The healthcare industry already faces many challenges. Digital pathology image storage doesn't need to be one of them. A healthcare organization can simplify its approach by partnering with a trusted provider of both scanning and storage services.

For over 70 years, Iron Mountain has been a strategic partner to more than 225,000 organizations in 63 countries, including 95% of the Fortune 1000 companies. Iron Mountain's holistic, as-a-service operational model allows an organization to scale up or down over time with a predictable price point. Pathologists, lab managers, IT teams, and hospital executives can transition to digital pathology without losing focus on their core business functions.

Iron Mountain empowers healthcare organizations to store and archive with confidence. A global leader in information and asset management, Iron Mountain offers a comprehensive and scalable solution for digital pathology scanning and storage.

Key takeaway

Three criteria are essential when choosing a service provider: reputation, experience, and capabilities. For these three reasons, Iron Mountain is trusted by **2,000 healthcare customers, 94 imaging centers,** and the **Top 10 global pharmaceutical companies.** Iron Mountain maintains **850 million+ patient records, 1 billion+ medical images,** and **1 billion+ pathology slides** in its secure storage.

Find out how Iron Mountain can meet your organization's digital pathology needs. Visit ironmountain.com/digital-pathology.

About Iron Mountain

For over 70 years, Iron Mountain Incorporated (NYSE: IRM) has been your strategic partner to care for your information and assets. A global leader in storage and information management services and trusted by more than 225,000 organizations around the world, including 95% of the Fortune 1000, we protect, unlock, and extend the value of your work—whatever it is, wherever it is, however it's stored.

We create the framework necessary to bridge the gaps between paper, digital, media, and physical data and extract value along its lifecycle, helping to build your organizational resilience. And all this with a commitment to sustainability at our core.

Our relationship is a true collaboration where you trust us not only to preserve institutional knowledge and enhance efficiency, security, and access but to make your work mean more. Because in that work is the power to not only accelerate your business but elevate it.

Trusted by more than 225,000 organizations around the world, and with a real estate network of more than 85 million square feet across more than 1,400 facilities in over 50 countries, Iron Mountain stores and protects billions of valued assets, including critical business information, highly sensitive data, and cultural and historical artifacts. Providing solutions that include information management, digital transformation, secure storage, secure destruction, as well as data centers, cloud services, and art storage and logistics, we help customers lower cost and risk, comply with regulations, recover from disaster, and enable a more digital way of working—and all this with a commitment to sustainability.



800.899.IRON | [ironmountain.com](https://www.ironmountain.com)

About Iron Mountain

Iron Mountain Incorporated (NYSE: IRM), founded in 1951, is the global leader for storage and information management services. Trusted by more than 220,000 organizations around the world, and with a real estate network of more than 85 million square feet across more than 1,400 facilities in over 50 countries, Iron Mountain stores and protects billions of information assets, including critical business information, highly sensitive data, and cultural and historical artifacts. Providing solutions that include secure storage, information management, digital transformation, secure destruction, as well as data centers, art storage and logistics, and cloud services, Iron Mountain helps organizations to lower cost and risk, comply with regulations, recover from disaster, and enable a more digital way of working. Visit www.ironmountain.com for more information.

© 2023 Iron Mountain, Incorporated and/or its affiliates ("Iron Mountain"). All rights reserved. Information herein is proprietary and confidential to Iron Mountain and/or its licensors and may not be used for competitive analysis or building a competitive product or otherwise reproduced without Iron Mountain's written permission. Iron Mountain believes the information is accurate as of its publication date, but such information is subject to change without notice, is provided AS-IS without representation or warranty of any kind including regulatory status or recommended use, does not provide a commitment to regional or future availability, and does not represent an affiliation with or endorsement by any other party. Recipients acknowledge Iron Mountain's products and services are not intended for primary interpretation or diagnostics or in connection with any other clinical decision-making, and recipients retain sole responsibility for ensuring their use is compliant with relevant laws and regulations, including those set by the FDA or similar regulatory bodies. "Iron Mountain", the Iron Mountain logo, and combinations thereof, and other marks marked by © or TM are trademarks of Iron Mountain Incorporated. All other trademarks may be trademarks of their respective owners.

