International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 32 Issue 06 – AUGUST 2026.

AI-Driven Autonomous Workload Orchestration for Cost, Performance, Security, and Sustainability Optimization in Multi-Cloud Environments

Ms. Spoorthi B[#1] and Prof. Nayana K[*2]

[#1] *Department of Computer Science & Engineering at GM University, Davanagere*

[*2] *Department of Computer Science & Engineering at GM University, Davanagere*

Abstract— Generative Artificial Intelligence (GenAI) has transformed data infrastructure from a passive storage layer into a strategic component of AI development and deployment. As organizations manage massive datasets, model checkpoints, embeddings, and fine-tuning artifacts, traditional cloud object storage platforms such as Amazon Web Services (AWS) S3 face limitations in supporting AI-specific requirements, including dataset versioning, model lineage, reproducibility, metadata management, and collaborative workflows [1]. Although AWS S3 provides excellent scalability, durability, and cost efficiency, it was designed for general-purpose cloud applications rather than AI-native environments.

This paper investigates the migration from AWS S3 to Hugging Face Private Storage, an AI-native platform that integrates datasets, models, metadata, version control, and governance within a unified ecosystem. A structured AI-Native Storage Migration Framework is proposed to support secure and efficient migration while preserving data integrity and minimizing operational disruption [2]. The findings demonstrate that Hugging Face Private Storage enhances collaboration, reproducibility, governance, and lifecycle management, providing a more efficient foundation than traditional cloud storage for building scalable, secure, and intelligent Generative AI applications.

***Index Terms*— Generative AI, AI-Native Storage, AWS S3, Hugging Face Private Storage, Data Infrastructure Modernization, MLOps, Model Governance, Data Lifecycle Management, Cloud Storage Migration, Artificial Intelligence Systems.**

I. Introduction

The rapid advancement of Generative Artificial Intelligence (GenAI) has significantly transformed the role of data infrastructure, making it a critical enabler of AI development rather than merely a storage repository. Unlike traditional machine learning applications that rely on structured datasets and periodic model updates, modern GenAI systems manage massive volumes of diverse data, including text corpora, multimodal datasets, embeddings, model checkpoints, and continuously evolving foundation models [10]. As a result, storage platforms now play a central role in improving model training, collaboration, governance, reproducibility, and deployment efficiency.

For many years, Amazon Web Services (AWS) S3 has been the preferred cloud object storage solution because of its scalability, durability, and cost-effectiveness [3]. However, AWS S3 was designed primarily for general-purpose cloud applications and lacks several features required for AI-native workflows, such as integrated dataset versioning, model lineage tracking, metadata management, and seamless support for MLOps and large language model (LLM) development. These limitations become increasingly significant as organizations adopt retrieval-augmented generation (RAG), fine-tuning pipelines, and collaborative AI development environments.

To address these challenges, AI-centric platforms such as Hugging Face Private Storage have emerged, providing an integrated ecosystem for managing datasets, models, version control, access governance, and deployment [4]. Unlike conventional storage services, Hugging Face Private Storage offers native support for collaborative AI workflows, secure model repositories, reproducible experiments, and lifecycle management, simplifying the development of Generative AI applications.

This paper investigates the migration from AWS S3 to Hugging Face Private Storage as a strategic modernization of AI infrastructure[5]. It evaluates both platforms across key aspects, including storage architecture, security, governance, collaboration, operational complexity, and lifecycle management. In addition, the study proposes a migration framework to support organizations in transitioning with minimal disruption while maintaining data integrity and compliance.

The findings demonstrate that AI-native storage platforms provide significant advantages in accelerating AI development, improving reproducibility, strengthening governance, and reducing operational overhead[6]. This work highlights that the future scalability and sustainability of Generative AI depend not only on advances in model architectures but also on intelligent, AI-native data infrastructures that seamlessly integrate storage, collaboration, governance, and lifecycle management.

A. Motivation

The rapid advancement of Generative Artificial Intelligence (GenAI) has transformed data infrastructure from a passive storage layer into a critical component of AI development and deployment. Traditional cloud storage platforms such as Amazon Web Services (AWS) S3 provide excellent scalability, durability, and cost efficiency but were designed for general-purpose cloud applications rather than AI-centric workflows. As organizations increasingly develop Large Language Models (LLMs), multimodal AI systems, and Retrieval-Augmented Generation (RAG) applications, they must manage massive datasets, embeddings, model checkpoints, and fine-tuning artifacts[7]. These workloads expose limitations in conventional storage systems, particularly in dataset versioning, model lineage, metadata management, reproducibility, and collaborative development.

Hugging Face Private Storage addresses these challenges by offering an AI-native platform that integrates datasets, model repositories, version control, security, and deployment within a unified ecosystem. This approach simplifies AI workflows, improves collaboration, enhances governance, and reduces operational complexity compared to traditional object storage solutions.

This study investigates the migration from AWS S3 to Hugging Face Private Storage by evaluating storage architecture, security, governance, collaboration, operational efficiency, and lifecycle management[8]. The findings demonstrate that AI-native storage platforms provide significant advantages in supporting modern Generative AI applications. The research concludes that future AI ecosystems require intelligent storage infrastructures that seamlessly integrate scalability, reproducibility, governance, collaboration, and lifecycle management to accelerate innovation and improve the efficiency of large-scale AI development.

B. Contributions

This research contributes to the emerging field of AI-native infrastructure by investigating how storage architectures must evolve to meet the demands of modern Generative AI systems. The key contributions of this work are as follows:

AI-Native Storage Perspective: Establishes a new perspective that positions storage infrastructure as a strategic enabler of the Generative AI lifecycle rather than merely a data repository.

Comparative Architectural Analysis: Provides a comprehensive evaluation of AWS S3 and Hugging Face Private Storage across scalability, governance, security, collaboration, and lifecycle management dimensions relevant to GenAI workloads.

Migration Framework: Proposes a structured framework for migrating datasets, model checkpoints, embeddings, and metadata from traditional cloud object storage to AI-centric storage environments.

AI Lifecycle Integration Assessment: Examines the impact of storage architecture on model development, fine-tuning, deployment, reproducibility, and continuous AI operations.

5. Governance-Centric Evaluation Model Introduces evaluation metrics focused on dataset lineage, model provenance, auditability, compliance, and collaborative research efficiency beyond conventional storage performance measures.

6.Future AI-Ready Infrastructure Model: Presents a conceptual framework for next-generation data infrastructures capable of supporting foundation models, multimodal AI applications, Retrieval-Augmented Generation (RAG), and autonomous AI systems.

C. Organizational

This paper is structured to systematically investigate the evolution of data infrastructure for Generative AI. Section II examines the state of the art in cloud-based and AI-native storage architectures. Section III defines the research framework and presents the architectural foundations of AWS S3 and Hugging Face Private Storage. Section IV details the proposed migration strategy and evaluation criteria. Section V analyses the platforms from performance, governance, reproducibility, security, and collaboration perspectives. Section VI discusses the broader implications of AI-native storage for foundation models, Retrieval-Augmented Generation (RAG), and future AI ecosystems[9]. Section VII concludes the study by summarizing the key contributions and identifying opportunities for future research in intelligent data infrastructure.

II. Evolution of Data Infrastructure for Artificial Intelligence

The history of Artificial Intelligence (AI) is closely intertwined with the evolution of data infrastructure. As AI systems have progressed from rule-based algorithms to large-scale Generative AI (GenAI) models, the underlying storage architectures have undergone significant transformation to accommodate increasing demands for scalability, accessibility, governance, and computational efficiency. This evolution reflects a broader shift in the role of data infrastructure—from passive storage repositories to intelligent platforms that actively support the complete AI lifecycle.

A. Traditional Data-Centric Infrastructure

Early AI applications primarily relied on structured and relatively small datasets stored within relational database management systems (RDBMS). These systems were designed to support transactional workloads and deterministic query processing, making them suitable for expert systems, statistical learning, and enterprise analytics. However, as data volumes increased and AI algorithms became more data-intensive, traditional databases struggled to meet the scalability and performance requirements of emerging machine learning workloads.

B. Big Data and Distributed Storage Era

The proliferation of web applications, social media platforms, and Internet-of-Things (IoT) devices generated unprecedented amounts of unstructured and semi-structured data. This led to the emergence of distributed storage architectures such as the Hadoop Distributed File System (HDFS), distributed object stores, and NoSQL databases. These technologies enabled organizations to process petabyte-scale datasets and laid the foundation for modern machine learning ecosystems. During this period, storage infrastructures focused primarily on scalability, fault tolerance, and high-throughput data access.

C. Cloud-Native Storage Revolution

Cloud computing transformed data management through scalable object storage services such as Amazon S3, offering high availability, global accessibility, and cost efficiency. However, these platforms were designed for general-purpose storage and lack native support for AI-specific assets, including model versioning, dataset lineage, experiment tracking, and collaborative workflows, requiring additional management tools.

D. Emergence of AI-Native Infrastructure

The rise of foundation models, LLMs, multimodal AI, and RAG has increased the need for AI-native infrastructure. Unlike traditional object storage, AI-native platforms manage datasets, models, embeddings, and metadata as integrated assets, providing built-in version control, governance, collaboration, reproducibility, and lifecycle management for efficient AI development.

E. Transition from AWS S3 to AI-Native Storage Ecosystems

The migration from AWS S3 to Hugging Face Private Storage reflects a shift from cloud-native to AI-native infrastructure. While AWS S3 focuses on scalable data storage, Hugging Face integrates datasets, models, governance, collaboration, and lifecycle management, enabling efficient handling of interconnected AI assets and supporting continuous Generative AI innovation.

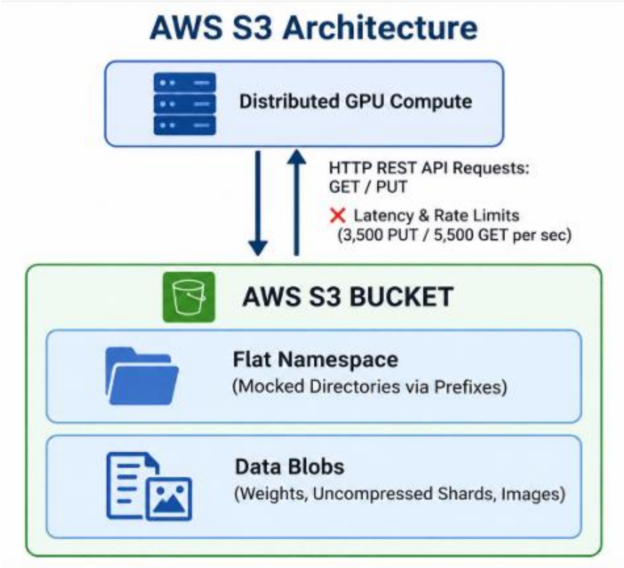
F. Toward Intelligent Data Infrastructure

Future AI infrastructures will support autonomous agents, multimodal models, federated learning, and evolving knowledge repositories through intelligent storage with automated governance, metadata management, and lifecycle integration. Data infrastructure is evolving from traditional storage to AI-native intelligent storage, making platforms like Hugging Face essential for scalable and collaborative Generative AI innovation.

III. AWS S3 Architecture and Limitations for Generative AI

A. AWS S3 Architecture: A Foundation for Modern Cloud Storage

Over the past decade, Amazon Simple Storage Service (AWS S3) has become one of the most influential technologies in cloud computing. Its success stems from a simple yet powerful design philosophy: provide highly scalable, durable, and cost-effective storage that can accommodate virtually any type of data. By organizing information as objects within buckets, AWS S3 allows organizations to store and retrieve massive volumes of data without worrying about the underlying infrastructure.



AWS S3 was designed for scalable, reliable storage of enterprise applications, analytics, and data lakes. As AI adoption grew, it became the backbone of machine learning workflows, efficiently storing datasets, models, and experiment outputs. Its scalability, security, reliability, and broad cloud ecosystem continue to make it a preferred storage solution.

B. The Changing Nature of AI Workloads

Generative AI demands more than scalable data storage. Modern AI workflows require intelligent management of datasets, models, embeddings, prompts, checkpoints, and metadata with built-in versioning, governance, reproducibility, and collaboration. The focus has shifted from simply storing data to efficiently managing interconnected AI assets throughout their entire lifecycle.

C. Architectural Limitations of AWS S3 for Generative AI

Despite its strengths, AWS S3 was not originally designed with AI lifecycle management in mind. As Generative AI workloads continue to grow, several limitations become apparent.

1) Storage without AI Context: AWS S3 stores all objects uniformly, regardless of whether they are datasets, model checkpoints, embeddings, or evaluation artifacts. Although flexible and scalable, it lacks native awareness of relationships among AI assets, requiring external tools to link datasets, models, and experiments.

2) Limited Support for Reproducibility: Reproducibility is essential in AI research, requiring accurate tracking of dataset versions, model checkpoints, and training configurations. While AWS S3 supports object versioning, it does not provide integrated lifecycle management for AI assets, making experiment reproduction dependent on additional management systems.

3) Fragmented Development Ecosystem: Generative AI workflows involve multiple tools for data preparation, model training, experiment tracking, deployment, and governance. AWS S3 functions primarily as a storage service, forcing organizations to integrate several independent platforms, which increases operational complexity and reduces workflow efficiency.

4) Collaboration Challenges: Modern AI development requires collaborative features such as shared repositories, version control, change tracking, and review mechanisms. AWS S3 enables file sharing but lacks built-in collaboration capabilities, making coordinated development across distributed teams more difficult.

5) Governance and Metadata Management: Although AWS S3 offers robust security and access control, governance information such as data provenance, model lineage, compliance records, and audit history must be managed separately. This separation increases administrative overhead and complicates regulatory compliance and AI governance.

D. Why These Limitations Matter in the Generative AI Era

The emergence of Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), multimodal systems, and autonomous AI agents has significantly amplified infrastructure complexity. These systems depend on continuous interaction between datasets, embeddings, vector representations, models, and deployment pipelines.

In such environments, the effectiveness of an AI platform is determined not only by its ability to store information but also by its ability to organize, govern, and operationalize AI assets. Traditional object storage remains highly effective for data persistence; however, it provides limited support for the broader lifecycle requirements of modern AI development.

This distinction is increasingly important as organizations seek to reduce development cycles, improve collaboration, strengthen governance, and ensure reproducibility at scale.

E. Toward AI-Native Storage

The limitations of AWS S3 should not be viewed as shortcomings of the platform itself. Rather, they reflect the fact that the platform was designed for an earlier generation of computing challenges. AWS S3 solved the problem of scalable cloud storage exceptionally well. However, Generative AI introduces a new class of challenges centered on managing complex relationships among datasets, models, embeddings, and collaborative workflows.

As a result, the industry is witnessing the emergence of AI-native storage platforms that integrate storage, version control, governance, collaboration, and deployment capabilities into a unified environment. This evolution represents a natural progression in data infrastructure, where storage systems move beyond passive data management and become active participants in the AI lifecycle.

The migration from AWS S3 to Hugging Face Private Storage therefore represents more than a technological upgrade. It symbolizes a broader transition from cloud-native storage designed for data-centric applications to AI-native infrastructure designed for intelligence-centric ecosystems.

IV. Hugging Face Private Storage Architecture: An AI-Native Approach to Data Infrastructure

The rapid growth of Generative Artificial Intelligence (GenAI) has transformed how organizations manage datasets, machine learning models, embeddings, prompts, and other AI assets. Unlike traditional cloud storage systems that primarily store files, Hugging Face Private Storage provides an AI-native architecture designed to manage the complete lifecycle of AI resources. It integrates datasets, models, metadata, governance, version control, and collaboration within a unified platform.

The platform includes dedicated dataset and model repositories with built-in Git-based version control, enabling researchers to track changes, reproduce experiments, and manage model evolution efficiently. A comprehensive metadata layer records data provenance, training parameters, evaluation metrics, licensing, and compliance information, ensuring transparency, governance, and responsible AI development. Secure collaboration is supported through private repositories, role-based access control, authentication, and controlled sharing, allowing distributed teams to work together while protecting sensitive intellectual property.

Hugging Face seamlessly integrates with the entire AI lifecycle, from data preparation and model training to deployment and continuous improvement, reducing infrastructure complexity and enhancing collaboration. Its AI-centric design, native versioning, reproducibility, governance, and lifecycle awareness make it well suited for managing foundation models, Retrieval-Augmented Generation (RAG) systems, multimodal applications, and enterprise-scale AI solutions. Consequently, Hugging Face represents a shift from conventional storage to intelligent, AI-native infrastructure.

V. AI-Native Storage Migration Framework

The transition from traditional cloud storage to AI-native infrastructure is not simply a matter of moving files from one platform to another. In the context of Generative AI, migration involves rethinking how datasets, models, embeddings, metadata, and governance artifacts are organized, managed, and utilized throughout the AI lifecycle.

Traditional storage systems such as AWS S3 were designed primarily to provide scalable and reliable data persistence. While this architecture remains highly effective for general cloud workloads, modern AI ecosystems demand a more integrated approach where storage actively supports model development, collaboration, reproducibility, governance, and deployment. Consequently, migration should be viewed as an organizational transformation rather than a technical relocation exercise.

To address this challenge, this paper explains an AI-Native Storage Migration Framework that provides a structured pathway for transitioning from AWS S3 to Hugging Face Private Storage while preserving operational continuity, governance requirements, and research reproducibility.

The proposed framework is designed around five key objectives:

- Preserve the integrity of datasets and model assets.
- Ensure reproducibility across AI experiments and workflows.
- Strengthen governance, security, and compliance mechanisms.
- Reduce infrastructure fragmentation and operational complexity.
- Enable collaborative and lifecycle-aware AI development.

Rather than focusing solely on storage efficiency, the framework seeks to align infrastructure with the evolving requirements of Generative AI systems.

The rapid advancement of Generative Artificial Intelligence (GenAI) has exposed a fundamental limitation in traditional storage architectures: they were designed to manage data, not intelligence. While cloud object storage platforms such as AWS S3 provide exceptional scalability and reliability, modern AI ecosystems require infrastructures capable of managing datasets, models, embeddings, metadata, governance records, and collaborative workflows as interconnected assets.

To address this challenge, this paper proposes an AI-Native Storage Migration Framework that enables organizations to transition from conventional cloud storage environments to AI-centric infrastructure platforms such as Hugging Face Private Storage. Unlike traditional migration approaches that focus primarily on transferring files, the proposed framework focuses on preserving and enhancing the entire lifecycle of AI assets.

The framework is built on the principle that successful AI infrastructure modernization requires not only data migration but also the migration of knowledge, context, governance, and collaboration capabilities.

The proposed framework is guided by five fundamental principles:

1. Intelligence-Centric Management

The framework treats datasets, models, embeddings, prompts, and metadata as strategic intelligence assets rather than isolated files.

2. Reproducibility by Design

Every dataset version, model checkpoint, and experimental configuration must remain traceable throughout the migration process.

3. Governance Integration

Security, compliance, lineage tracking, and auditability are incorporated into the infrastructure rather than managed through external systems.

4. Collaborative Development

The framework supports distributed teams by enabling shared repositories, transparent version control, and collaborative workflows.

5. Lifecycle Awareness

Infrastructure decisions are aligned with the complete AI lifecycle, from data acquisition and model training to deployment and continuous improvement.

These principles ensure that migration enhances organizational AI capabilities rather than merely relocating storage assets.

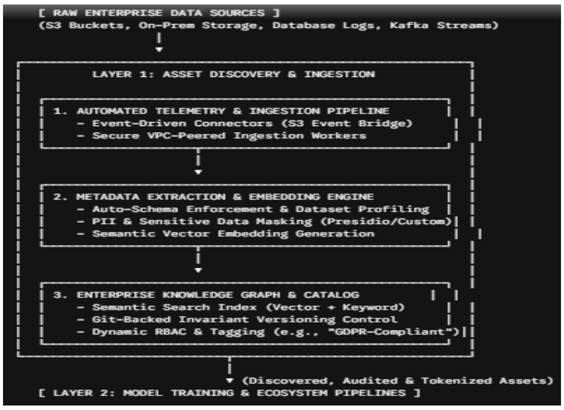
The AI-Native Storage Migration Framework consists of six interconnected layers.

Layer 1: Asset Discovery Layer

The first layer focuses on identifying and cataloguing all AI-related assets currently stored within AWS S3.

These assets include:

- Training datasets
- Fine-tuning datasets
- Foundation models
- Model checkpoints
- Embeddings
- Evaluation results
- Experiment logs
- Governance artifacts

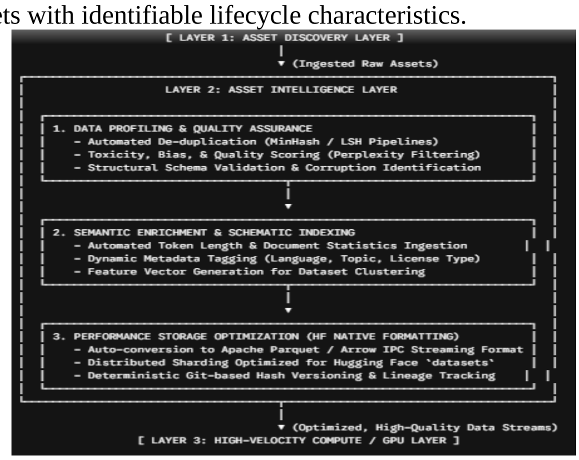


Layer 2: Asset Intelligence Layer

After discovery, assets are enriched with contextual information describing their relationships and dependencies. Examples include:

- Dataset-to-model relationships
- Model-to-experiment relationships
- Embedding dependencies
- Version histories
- Ownership information

This layer transforms storage objects into meaningful AI assets with identifiable lifecycle characteristics.



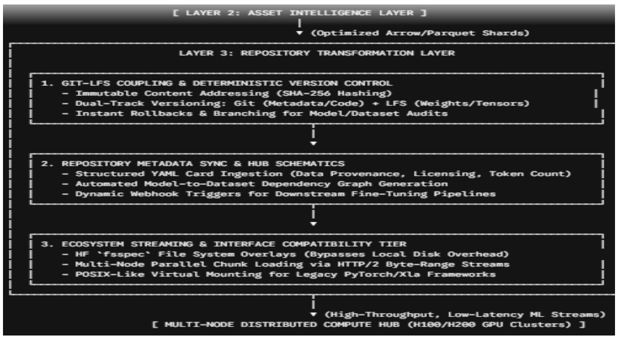
Layer 3: Repository Transformation Layer

At this stage, AI assets are reorganized into structured repositories.

The transformation includes:

- Dataset repositories
- Model repositories
- Embedding repositories
- Documentation repositories
- Governance repositories

Unlike traditional object storage, repository-based organization improves discoverability, traceability, and collaboration.



Layer 4: Governance and Compliance Layer

The fourth layer embeds governance mechanisms directly into the storage architecture.

Key capabilities include:

- Dataset lineage tracking
- Model provenance management
- Audit trails
- Access-control policies
- Compliance documentation
- Intellectual property protection

This layer ensures that governance becomes an inherent property of the infrastructure rather than an afterthought.

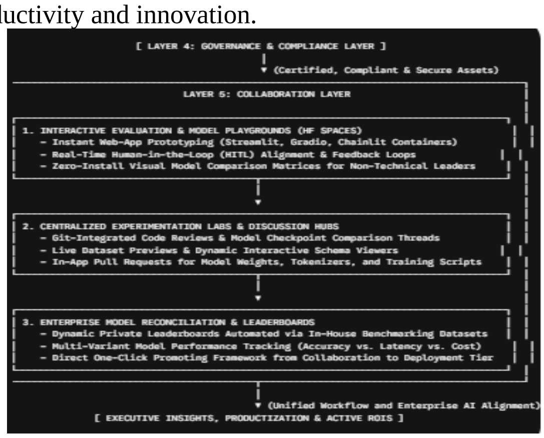
Layer 5: Collaboration Layer

Generative AI development is increasingly collaborative and interdisciplinary.

This layer introduces:

- Shared repositories
- Team-based workflows
- Version-controlled collaboration
- Knowledge sharing mechanisms
- Transparent change management

By integrating collaboration directly into the storage ecosystem, organizations can significantly improve research productivity and innovation.



Layer 6: Lifecycle Integration Layer

The final layer connects storage with the broader AI development lifecycle.

Integration points include:

- Data preparation pipelines
- Model training workflows
- Fine-tuning processes
- Evaluation frameworks
- Deployment pipelines
- Continuous learning systems

As a result, storage evolves from a passive repository into an active component of AI operations.

The framework is expected to provide several organizational and technical advantages:

Improved Reproducibility

Researchers can accurately reproduce experiments through integrated version control and asset lineage tracking.

Enhanced Governance

Compliance requirements become easier to satisfy through centralized governance mechanisms.

Reduced Infrastructure Complexity

Multiple disconnected tools can be consolidated into a unified AI-native ecosystem.

Stronger Collaboration

Teams can work more effectively through repository-based workflows and transparent change tracking.

Faster Innovation Cycles

Integrated lifecycle management accelerates experimentation, model development, and deployment.

Future Readiness

Organizations become better positioned to support foundation models, multimodal AI systems, Retrieval-Augmented Generation (RAG), and autonomous AI agents.

VI. Comparative Analysis and Evaluation Metrics:

A. *Comparative Analysis:*

1. Scalability:

Scalability has long been the defining strength of cloud object storage platforms. AWS S3 offers virtually unlimited storage capacity and has proven its effectiveness across enterprise-scale environments. Hugging Face Private Storage also provides scalable repository management; however, its primary focus extends beyond storage capacity to include lifecycle-aware management of AI assets.

From a pure storage perspective, AWS S3 demonstrates exceptional scalability. However, when scalability is considered in terms of managing growing collections of datasets, models, and collaborative AI projects, Hugging Face offers additional advantages through its repository-centric architecture.

Observation:
AWS S3 excels in data scalability, whereas Hugging Face Private Storage enhances scalability at the level of AI asset management.

2. Reproducibility
Reproducibility is a cornerstone of trustworthy AI research and development. Organizations must be able to recreate models, experiments, and results with confidence.
AWS S3 provides object versioning but does not inherently manage relationships between datasets, models, and experiments. In contrast, Hugging Face integrates version control directly into repositories, allowing teams to track asset evolution more systematically.
Observation:
Hugging Face provides stronger support for reproducible AI workflows by maintaining contextual relationships among datasets and models.

3. Governance and Compliance
As AI systems become increasingly regulated, governance has emerged as a critical infrastructure requirement.
AWS S3 offers robust security controls and auditing capabilities but often requires additional tools to establish comprehensive model governance and lineage tracking. Hugging Face integrates governance elements directly into repository structures, making provenance and documentation more accessible.
Observation:
AWS S3 provides strong enterprise governance foundations, while Hugging Face simplifies governance through tighter integration with AI assets.

4. Collaboration

Modern AI innovation is rarely an individual effort. Large-scale models are typically developed through collaboration among researchers, engineers, and domain experts.

AWS S3 facilitates data sharing but was not designed as a collaborative AI work. Hugging Face adopts repository-based workflows that encourage knowledge sharing, version tracking, and collaborative model development.

Observation:
Hugging Face significantly improves collaborative AI development through integrated repository management and transparent version control.

5. Lifecycle Integration
One of the most important differences between traditional and AI-native storage lies in lifecycle awareness.
AWS S3 primarily functions as a storage layer that must be connected to external services for training, deployment, and governance. Hugging Face is designed around the complete AI lifecycle, supporting workflows from dataset management to model deployment.

Observation:
Hugging Face demonstrates stronger alignment with end-to-end AI lifecycle requirements.

6. Operational Complexity

As AI ecosystems expand, operational complexity becomes a major concern.

Organizations using AWS S3 frequently integrate multiple services for version control, governance, experiment tracking, and deployment management. While highly flexible, this approach can increase infrastructure complexity.

Hugging Face consolidates many of these functions within a unified environment, reducing the need for multiple disconnected tools.

Observation:
AI-native storage architectures can significantly reduce operational overhead by consolidating lifecycle functions.

7. Security and Access Control
Security remains a fundamental requirement regardless of infrastructure type.
AWS S3 provides mature enterprise-grade security capabilities, including encryption, access management, and compliance support. Hugging Face Private Storage also provides robust access controls and repository-level permissions designed for collaborative AI environments.
Observation:
Both platforms offer strong security capabilities, although their implementation approaches reflect different architectural priorities.

8. Future AI Readiness
Future AI systems will increasingly rely on foundation models, multimodal learning, Retrieval-Augmented Generation (RAG), and autonomous AI agents.
Traditional storage architectures are highly effective at storing data but may require extensive integration efforts to support emerging AI workflows. AI-native platforms are specifically designed to accommodate evolving model-centric ecosystems.

Observation:
Hugging Face demonstrates greater alignment with future AI infrastructure requirements due to its lifecycle-oriented design philosophy.

Evaluation Metric	AWS S3	Hugging Face Private Storage
Storage Scalability	Excellent	Excellent
Dataset Versioning	Moderate	Excellent
Model Versioning	Limited	Excellent
Reproducibility	Moderate	Excellent
Governance Integration	Good	Excellent
Collaboration Support	Moderate	Excellent
Lifecycle Integration	Limited	Excellent
Operational Simplicity	Moderate	High

Security & Access Control	Excellent	Excellent
AI Readiness	Moderate	High

The future competitiveness of AI systems will depend less on the volume of data they store and more on the effectiveness with which they govern, reproduce, and operationalize intelligence assets. This shift makes AI-native evaluation metrics as important as traditional storage performance metrics.

VII. CONCLUSION:

The rapid growth of Generative Artificial Intelligence (GenAI) has transformed storage infrastructure into a strategic component of AI development. This study examines the migration from Amazon Web Services (AWS) S3 to Hugging Face Private Storage as a shift from traditional cloud storage to AI-native infrastructure. While AWS S3 offers excellent scalability, reliability, and durability, it lacks native support for AI asset management. Hugging Face Private Storage integrates datasets, models, version control, governance, and collaboration within a unified platform, improving reproducibility and lifecycle management. The proposed AI-Native Storage Migration Framework enables organizations to build scalable, secure, collaborative, and trustworthy Generative AI ecosystems.

REFERENCES

[1] J. Santos, C. Wang, A. Tantawi, O. Tardieu, P. Maniotis, T. Wauters, and F. De Turck, "Toward Efficient Orchestration of GenAI Workloads in Modern Container Clouds," *IEEE Communications Magazine*, vol. 64, no. 2, 2026. DOI: 10.1109/MCOM.001.2500262.

[2] J. Perez-Valero, G. Garcia-Aviles, A. Skarmeta, and T. Chen, "AI-Powered Orchestration-as-a-Service for 6G Networks: The 6G-CLOUD View," *IEEE Communications Standards Magazine*, vol. 10, no. 2, pp. 33–39, 2026. DOI: 10.1109/MCOMSTD.2026.3656636.

[3] Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "A Survey of AI-Enabled Security for Cloud Computing," *IEEE Network*, 2021

[4] X. Xu et al., "AI-Driven Resource Management for Cloud Computing: A Survey," *IEEE Access*, vol. 9, 2021

[5] M. Mao, J. Li, and M. Humphrey, "Deep Reinforcement Learning for Intelligent Resource Management in Cloud Computing: A Survey," *IEEE Transactions on Cloud Computing*, 2022

[6] Z. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Deep Reinforcement Learning for Resource Management in Network Slicing and Multi-Cloud Computing," *IEEE Wireless Communications*, 2020

[7] T. Chen et al., "AI-Driven Network and Service Orchestration in Cloud-Native 6G Systems," *IEEE Communications Magazine*, 2025

[8] Performance-Aware Zero-Trust Architecture for Secure and Scalable Multi-Cloud Computing, *IEEE International Conference on Intelligent Systems in Engineering, Secured Systems and Cybersecurity*, 2026

[9] J. Santos et al., "Toward Efficient Orchestration of GenAI Workloads in Modern Container Clouds," *IEEE Communications Magazine*, 2026

[10] Cloud Supply Chain Optimization using AI-driven Forecasting, *Proceedings of the 2026 International Conference on Smart Futuristic Technology (ICSFT)*, IEEE, 2026. DOI: 10.1109/ICSFT66733.2026.11507034.