MINIO

AIStor

The Data and Memory Foundation for Enterprise AI

AIStor is one unified data store for memory, tables, and objects. Agents, analytics engines, and AI pipelines work from the same governed data, at the performance and scale enterprise AI requires. It holds what agents remember, what analytics query, and the models, checkpoints, and embeddings in between. AIStor deploys on the infrastructure you choose across core, edge, and cloud, and scales from a first AI project to an exabyte estate.

Overview

AI applications utilize persistent memory, structured tables, and billions of objects. AIStor brings them together in one enterprise AI data foundation built for performance, governance, and scale. Unified means every workload acts on the same data at the same moment: what an agent writes to memory, an engine can query and a pipeline can act on, with no copy step and no drift between systems. Governance is configured once and holds everywhere, so the controls that protect a table also protect what an agent remembers about it. Enterprise data compounds instead of fragmenting.

Unified Architecture

Memory, tables, objects, and files on one foundation.

Enterprise AI programs spend more of their engineering investment moving data between systems than building on it. AIStor gives them one namespace, one identity model, and one binary, so an agent, a query engine, and a training job address the same data without a pipeline in between. The integration work disappears, and the time goes back into the application.

Memory. Agents are becoming pervasive across the enterprise, alone or alongside people, reviewing documents, building workflows, and making consequential decisions. What they do and learn must persist. Sessions end and sandboxes recycle, so without agentic memory an agent's judgment disappears with the run. AIStor Memory keeps it. Long-term memory fills automatically through Agent Biography, a time-stamped, searchable record of every run, and deliberately through memory tools. Teams can read, audit, and correct any memory. Skills carry what one agent masters to the agents that follow. Workspace carries work in progress across runs and handoffs, and Vault holds the credentials an approved agent needs, kept separate from both. Together they form organizational memory: decisions, context, and reasoning from completed work, owned by the enterprise rather than the session. Every new agent starts from what the organization knows, not from zero.

Tables. Structured data belongs inside the data store, not beside it. AIStor Tables implements Apache Iceberg natively with a built-in Iceberg REST catalog, so lakehouse architectures, feature stores, and AI pipelines query governed tables with no separate table platform to license, deploy, or secure. Tables persist as objects on the same foundation, which is why there is no second system to keep in sync. Existing Iceberg catalogs migrate with a single command, and the data files stay where they are.

AT A GLANCE

Designed for the way enterprise AI actually runs

- **One data store:** Memory, tables, objects, and files on a unified data foundation.
- **Memory that lasts:** Agents keep, share, and build on what they learn.
- **Software-defined:** One static binary, industry-standard hardware, edge to exabyte.
- **Iceberg inside:** Native table support, with no separate platform to run.
- **S3 without changes:** Existing applications and AI frameworks just work.
- **Proven at exabyte scale:** Microsecond latency from edge to the largest clusters.
- **Standards-based, no lock-in:** Open formats and protocols at every layer.
- **Your keys, your environment:** Nothing your agents learn leaves your infrastructure.
- **Predictable economics:** Capacity-based subscription, no egress fees.

Objects and Files. Everything else in the AI lifecycle lands here: models, checkpoints, training datasets, embeddings, documents, media, and logs. AI is built on object storage. Frameworks, training pipelines, and query engines speak S3 first, and a flat namespace scales where hierarchical filesystems stall. The object layer is S3-native and engineered for both sides of the problem, the throughput GPU infrastructure requires and the scale AI accumulates. Billions of objects live under a single namespace, and clusters grow to exabytes without re-platforming. Capacity and throughput scale together, which allows memory and tables to run on the same foundation without a performance penalty. AIStor Files presents the same objects over POSIX for applications that expect a filesystem, an access path rather than a second system.

One platform beneath all three. Every layer inherits the same platform. Identity and access management, encryption and key management, object immutability and anti-ransomware protection, replication, data resilience, observability, traffic management, multi-tenancy, and proactive support ship as part of AIStor. Workloads connect over standards-based protocols including S3, S3 Express, Iceberg Catalog, OpenSharing, MCP, and SFTP, keeping every layer portable and free of lock-in.

The AIStor Difference

The data foundation for AI is not the foundation of the last decade. Blocks and files were built for applications that read a file and wrote it back, not for GPU clusters consuming petabytes, agents accumulating knowledge, and query engines working the same estate. Enterprise AI needs one foundation for every data type it produces, fast enough to keep accelerators working and large enough to hold all of it.

Architected for Modern AI & Analytics Workloads

Memory, tables, and objects run on one software-defined architecture with one identity model and one governance model, from a single edge node to an exabyte cluster. A common foundation means fewer moving parts and fewer places for data to diverge: one copy of the truth, one set of controls to audit, one system to operate and staff. Native Apache Iceberg support removes the proprietary table layer from the stack entirely.

Unmatched Performance & Scale

AIStor keeps GPU infrastructure working instead of waiting on data. Throughput scales with the cluster rather than bottlenecking at a centralized metadata tier, and the same architecture holds from a single node to billions of objects across exabytes of capacity. Growth requires no re-platforming, so the architecture that runs the first AI project is the architecture that runs the estate.

Operational Simplicity & Lower TCO

AIStor is software. It ships as a single static binary with no external metadata database and no background services, and it runs on the industry-standard hardware you choose, at the edge, in the data center, or in the cloud. A small team operates it at scales that traditionally require a storage organization. Subscriptions are capacity-based and all-inclusive, with no per-operation or egress charges.

Built for Every Enterprise AI Workload

AI Agents

Fleets share a common memory, so knowledge compounds across sessions and agents instead of resetting with each run.

RAG and Inference

Models receive exactly the context they need, improving answer quality while cutting token spend and latency.

AI Training

Data arrives as fast as GPU clusters consume it, and checkpoints complete without stalling the run.

Analytics and Lakehouse

Every engine runs against one governed copy of enterprise data, ending the silo sprawl that fragments most estates.

By the Numbers

23.5 TiB/s	Millisecond	40%	Exabyte	77%
Demonstrated peak throughput	Access latency for inference	Lower TCO vs proprietary AI storage	Scale demonstrated in benchmarking	Of the Fortune 500 use MinIO

ABOUT MINIO

MinIO is the data and memory foundation for enterprise AI. AIStor and MemKV unify every layer of the data stack, from agentic AI and inference context memory to tables and objects across core, edge, and cloud. Each layer is built for the speed, scale, and economics that AI and analytics demand. Trusted by 77% of the Fortune 500, MinIO is redefining how AI factories, intelligent applications, and autonomous agents secure, persist, and unlock the full value of their data.