

NeuralSearch: Storage-Native Search for Unstructured Data at Massive Scale

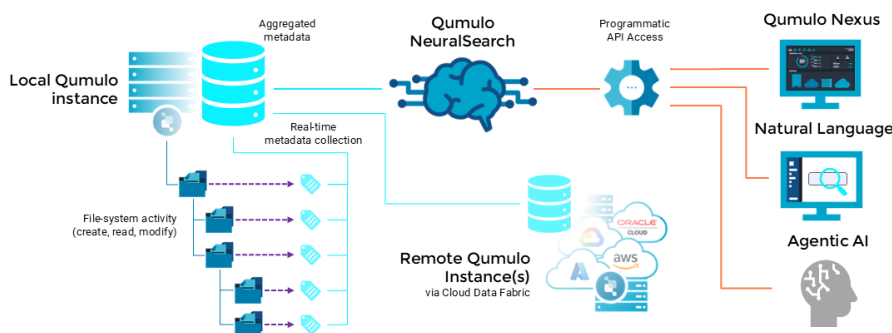
ANY DATA
ANY LOCATION
TOTAL CONTROL

EXECUTIVE SUMMARY

Even as organizations are accelerating their AI strategy, the unstructured data that powers most AI and analytics workflows is spread across file systems, object stores, edge sites, data centers, and multiple clouds. Traditional search relies on crawlers, ETL pipelines, and separate databases that quickly become stale, making discovery slow, complex, error-prone, and expensive.

NeuralSearch is Qumulo’s storage-native search feature that makes unstructured data instantly queryable by turning the file system into a real-time source of truth. Built directly into the platform, it eliminates crawlers and complex external infrastructure while delivering constant-time search capabilities across hundreds of millions of files.

With support for SQL, natural language, and semantic search, NeuralSearch helps organizations find and use data faster across core, cloud, and edge environments, powering AI/ML, analytics, compliance, and agentic AI from a single, unified metadata layer.



THE CHALLENGE

Traditional approaches to searching unstructured data create major operational and business challenges:



Slow namespace walks across large file environments



Stale indexes from crawlers and ETL pipelines



Complex infrastructure with separate databases and VMs



Manual tagging limits discovery accuracy



Poor support for AI/ML workflows and agents



Fragmented search across core, cloud, and edge

KEY BENEFITS

With Qumulo NeuralSearch, you get:

- Instant search without crawlers, indexing jobs, or manual tagging
- No separate complex search infrastructure VMs, databases, ETL pipelines, or search clusters
- Elimination of slow namespace walks and performance degradation at scale
- Search using SQL, plain English, or semantic similarity
- Unified metadata view across environments without replication or separate indexing
- Support for MCP-ready interfaces and open formats like Parquet and Iceberg, making it ideal for analytics platforms and AI/ML pipelines

As data volumes grow, these inefficiencies increase costs, slow decision-making, and reduce the value of enterprise data.

THE SOLUTION: NEURALSEARCH

NeuralSearch transforms the file system into a live query engine by using metadata the storage platform already tracks.

Instead of scanning namespaces or maintaining external search infrastructure, NeuralSearch continuously materializes metadata in real time using native columnar indexes. This enables users to instantly query data across structured metadata and unstructured content using four core capabilities:

SQL Metadata Search

Users can query system and user-defined metadata instantly, using open formats like Parquet and Iceberg through tools such as Spark, Trino, and DuckDB.

Natural Language Search

Users can ask questions in plain English, and an LLM automatically translates those requests into SQL, making search accessible to both technical and non-technical users.

Semantic Search

Vector embeddings allow users to search by meaning, similarity, and visual context rather than relying only on file names or tags.

Example: "Find desert chase scenes including a car"

Temporal Search

Users can ask questions like:

"What changed since Monday?"

This enables rapid compliance investigations, audit workflows, and operational visibility.

IDEAL USE CASES

AI/ML Data Discovery

Rapidly locate and retrieve training datasets without duplication or manual preparation.

Analytics Modernization

Replace slow namespace walks with fast SQL-based metadata queries.

Compliance and Investigations

Accelerate audit workflows and locate relevant files instantly using metadata and temporal search.

Media and Content Discovery

Find specific scenes, visual content, or creative assets using semantic search instead of manual tagging.

Agentic AI Platforms

Enable autonomous agents to discover and retrieve enterprise data using standard interfaces and real-time metadata access.

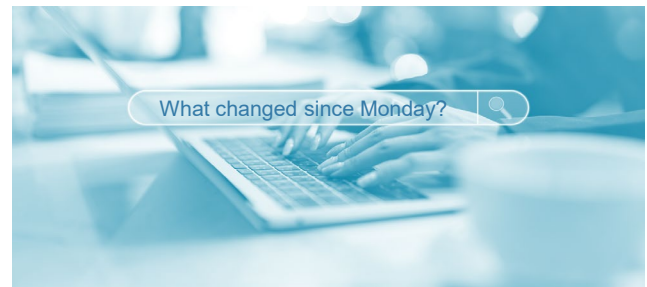
WHY QUMULO

Unlike traditional search platforms, NeuralSearch is built directly into the storage system—where the truth already exists.

Because it is integrated with Qumulo Cloud Data Fabric, organizations gain:

- A single intelligent view of data across all environments
- Strict consistency without stale indexes
- An integrated search feature that doesn't require replication or separate tools
- Real-time metadata propagation globally
- Direct retrieval using standard NFS and S3 workflows

The result is faster discovery, lower operational cost, and dramatically simpler access to unstructured data.



NeuralSearch is priced as a 25% uplift on CNQ or ANQ \$/TB pricing. For on-premises licensing, please contact sales. NeuralSearch is deployed on a separate VM.

ABOUT QUMULO

Qumulo is the leading provider of cloud file data platforms, offering unrivaled performance, scale, and data management solutions. Qumulo's platform is trusted by Fortune 500 companies and global enterprises to manage petabytes of data, unlocking its value and driving innovation. For more information, visit www.qumulo.com