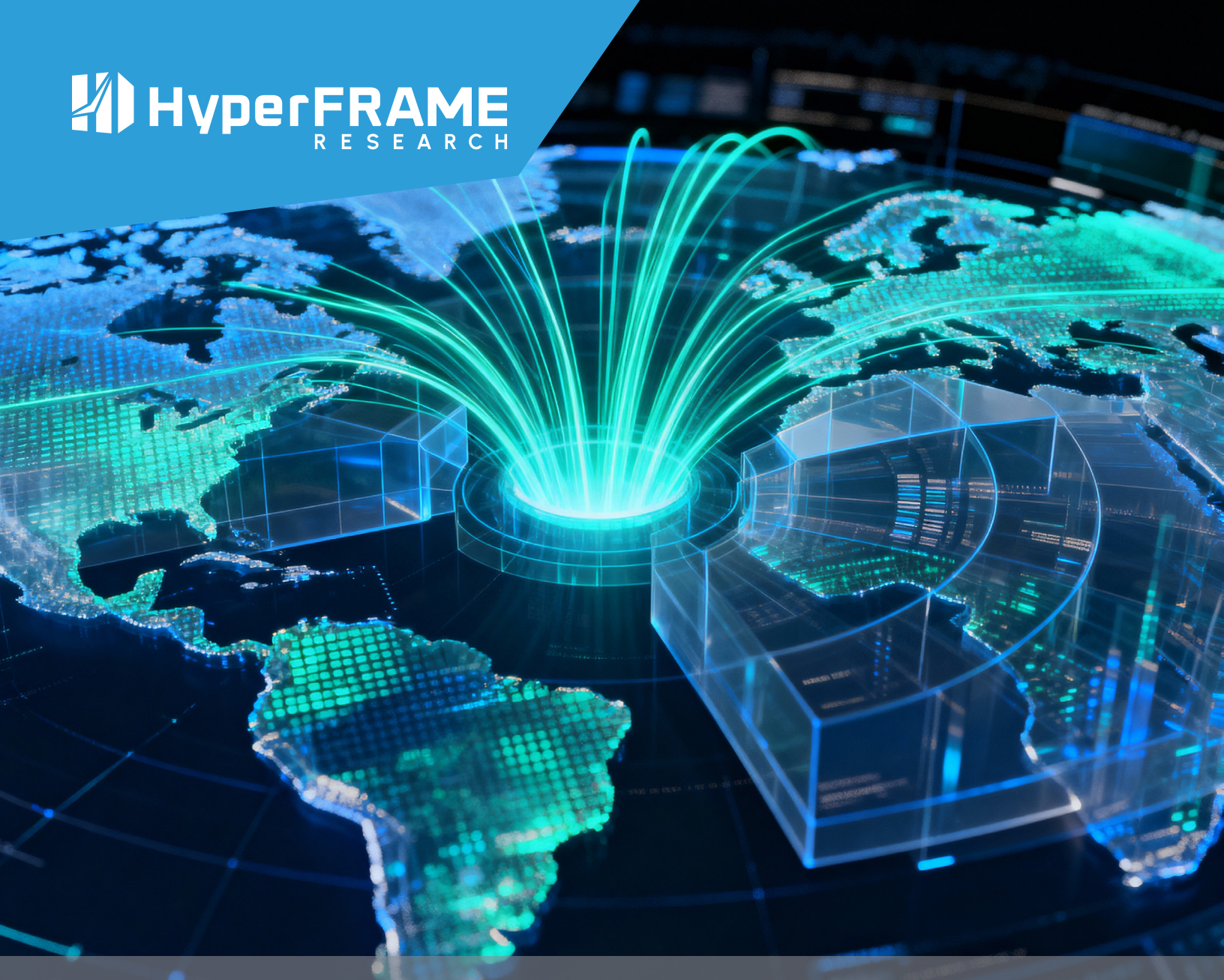




RESEARCH BRIEF

# Hammerspace: Unifying Data for Enterprise AI Success

Authors:

**Ron Westfall**  
VP and Practice Leader for  
Infrastructure and Networking

**Don Gentile**  
Analyst-in-Residence  
Storage & Data Resiliency

**Steven Dickens**  
CEO and Principal Analyst

JANUARY 2026



## Executive Summary

The enterprise data management and storage market is evolving and increasingly targeting the growing needs of AI and high-performance computing (HPC) workloads. Enterprises, as they wrestle with the increasingly dynamic landscape, are often focusing on a single, unified data environment across on-premises, edge, and cloud storage. A key trend driving this market is the “data gravity” problem, where the sheer volume of unstructured data makes it difficult and expensive to move for processing. Increasingly, the market is demanding a software-defined data layer that makes data instantly accessible to applications and GPUs regardless of their physical location.

Against this backdrop, Hammerspace’s open approach directly benefits enterprises in their AI journeys by providing a high-performance data platform that eliminates silos and ensures fast access to data across diverse environments, from edge to cloud. This enables efficient data orchestration, crucial for scaling data-intensive AI models and maximizing GPU utilization. By leveraging existing infrastructure and eliminating manual data copying, it significantly reduces costs and accelerates time-to-insight for AI initiatives.

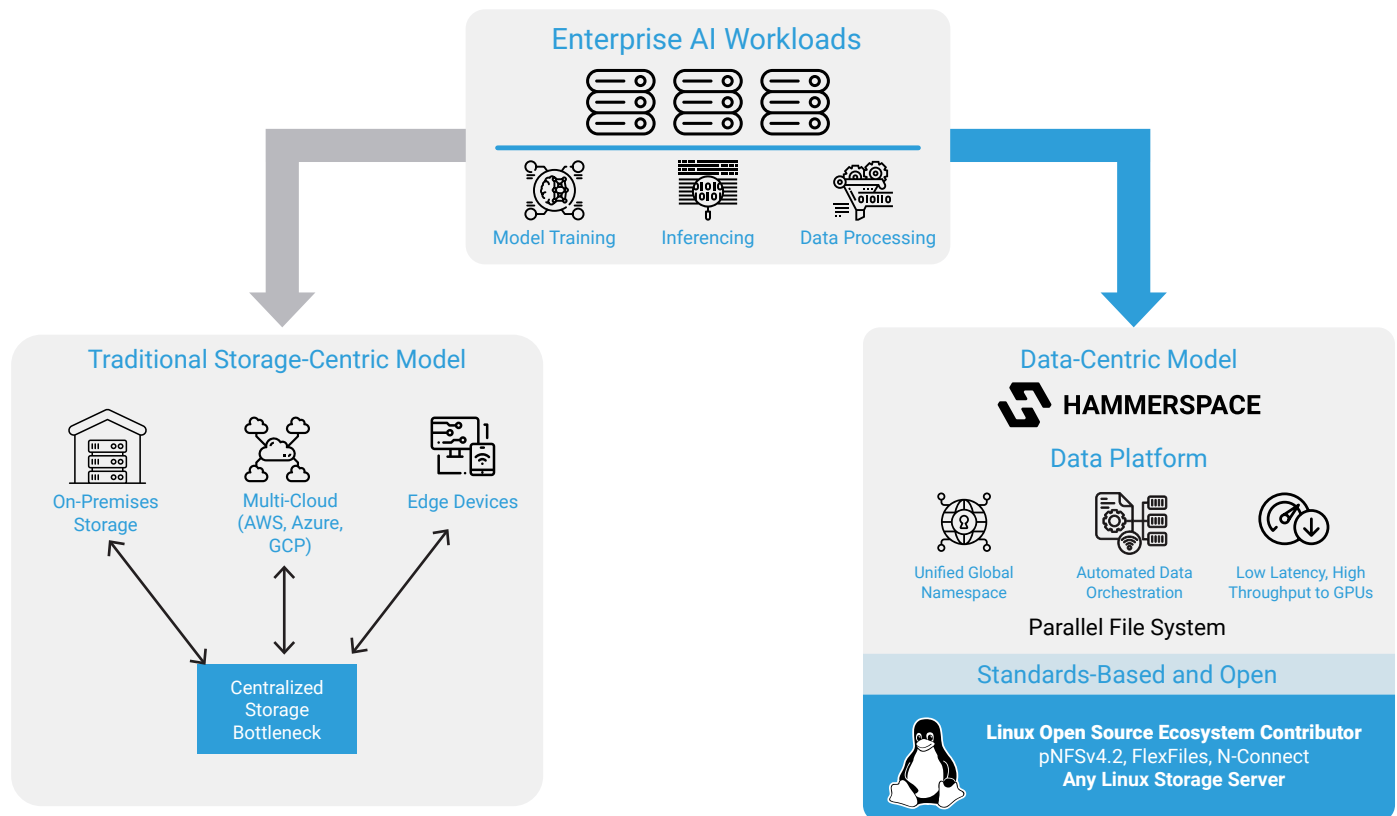
## I. Hammerspace’s Data Platform: Integral to Advancing Enterprise AI Journeys

Transforming the enterprise AI landscape requires a fundamental shift from traditional storage-centric models to a more agile, data-centric approach. This is where Hammerspace offers a compelling solution as a high-performance data platform that eradicates data silos across on-premises, edge, and multi-cloud infrastructures. By providing AI models with seamless, high-performance data access to GPUs - including the use of server-local flash storage as shared storage - Hammerspace enables consistent, locality-independent access to unstructured data across distributed environments. Hammerspace’s high-performance design supports data-hungry AI workflows by keeping GPUs consistently fed, regardless of where that data originates.

From our perspective, Hammerspace provides the architectural foundation to help maximize utilization of costly compute resources. This is all powered by automated data orchestration, intelligently moving and placing data based on policies, without requiring changes to applications or existing storage systems.

Furthermore, Hammerspace's strong Linux open-source ecosystem credentials, demonstrated through significant contributions to the Linux kernel and NFS protocol, underscore its commitment to open standards and high-performance, flexible data management for AI workloads without proprietary client software. Here are the key portfolio capabilities that make Hammerspace essential to fulfilling enterprise AI strategies:

- **From Storage-Centric to Data-Centric:** Hammerspace's open architecture approach is crucial for advancing enterprise AI journeys as it breaks down data silos across on-premises, edge, and multi-cloud environments by establishing a single global namespace. This means AI models can access all relevant data from anywhere, eliminating the need for costly and time-consuming data migrations.
- **High-Performance Data Access to GPUs:** AI training and inference demand extremely low latency and high throughput data access to keep expensive GPUs fully utilized. Hammerspace's parallel file system architecture provides the necessary performance to eliminate I/O bottlenecks and accelerate AI workloads.
- **Automated Data Orchestration:** Hammerspace automates the placement of data based on policies and metadata, ensuring that the right data is available at the right time and location for data processing. This intelligent data orchestration optimizes resource utilization, reduces operational complexity, and lowers costs by preventing data duplication and unnecessary data movement.
- **Linux Open Source Ecosystem Credentials:** Proven track record of significant contributions to the Linux Foundation and the open source community by actively enhancing the Linux kernel, particularly through contributions to the NFS protocol (like NFSv4.2 with FlexFiles and N-Connect), to deliver high-performance data access for enterprise and AI workloads using standard file and object protocols. Any Linux storage server can be used to build out the underlying storage architecture, giving customers unprecedented flexibility and choice.



Benefits: Eliminate data silos, accelerate AI pipelines, faster time-to-value



## II. Hammerspace's Data-Centric Approach – The Key Benefits

In the pursuit of advanced AI capabilities, enterprises frequently encounter significant hurdles related to data fragmentation, infrastructure utilization, and operational efficiency. Hammerspace's data-centric approach directly addresses these challenges by eliminating data silos through a unified global namespace, enabling AI models to access all relevant data regardless of its physical location across on-premises, edge, or cloud environments.

This innovative platform further maximizes the return on investment in expensive GPU infrastructure by intelligently orchestrating data to ensure proximity to compute resources, thereby preventing I/O bottlenecks and optimizing utilization. Hammerspace automates data placement and tiering without manual intervention or duplication, all while seamlessly leveraging existing infrastructure. The result is a dramatic acceleration of the entire AI development pipeline, from data preparation and model training to deployment, ultimately speeding up the delivery of critical business insights.

As such, we identify the following key benefits attained through Hammerspace's data-centric approach:

**Eliminates Data Silos:** Unlike storage-centric models that trap data within specific, often incompatible, storage systems, Hammerspace creates a unified global namespace that spans sites, clouds, and storage systems. This allows AI models to access all data, regardless of its physical location (on-premises, edge, or cloud).

**Maximizes GPU Utilization:** AI workloads are incredibly data-intensive and expensive, with GPUs being a major investment. Hammerspace's ability to orchestrate and move data intelligently prevents bottlenecks and maximizes the

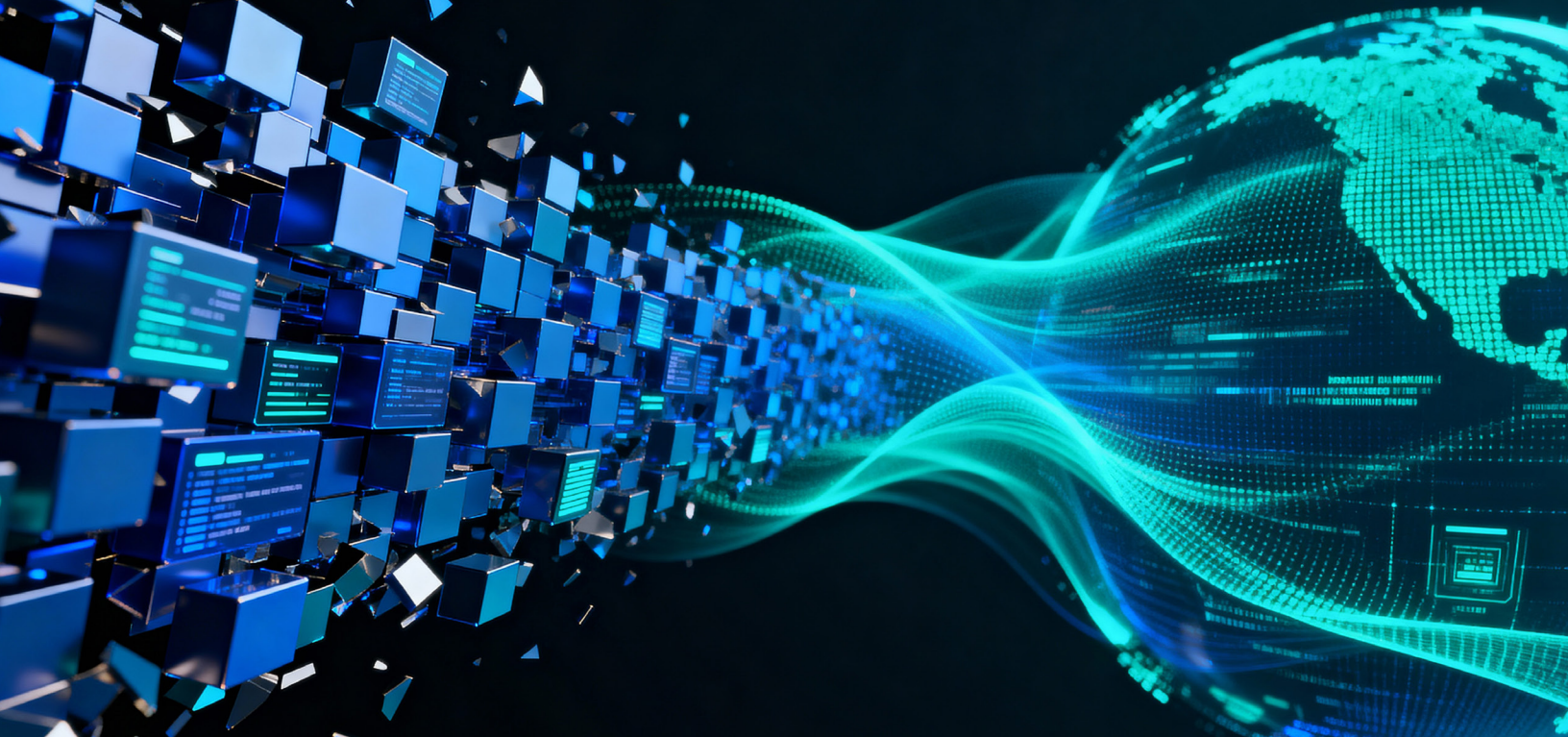
efficiency of valuable GPU assets, leading to better ROI. And Hammerspace unlocks the full potential of the GPU server-local storage - activating that storage as a "Tier 0" of high-performance shared storage. A Tier 0 architecture enables automatic pre-placement of files onto GPU server-local NVMe for maximum throughput, dramatically reducing latency by bypassing traditional network I/O paths.

**Enables Global Data Orchestration and Mobility:** A data-centric approach enables policy-driven automation of data placement. This means hot data can be automatically moved closer to active AI training, while less frequently accessed data can be tiered to more cost-effective storage, all without manual intervention or data duplication.

**Enables Hybrid-Cloud and Multi-Cloud Agility:** Hammerspace also automates data movement between on-prem and cloud environments, and between cloud regions, so you can migrate data into different cloud providers, and bring the data to available cloud GPU resources for burst computing.

**Leverages Existing Infrastructure:** Instead of forcing enterprises to rip and replace their existing storage infrastructure, Hammerspace acts as an intelligent overlay while still providing high-throughput, low-latency data access. This allows organizations to reuse existing storage systems while gaining the benefits of a modern, data-centric architecture for AI.

**Accelerates Time-to-Insight:** By providing agile, high-performance access to all relevant data, Hammerspace significantly reduces the time it takes to prepare data for AI models, train them, and deploy them for inference. This starts with making millions of files on existing file storage systems visible and accessible in minutes, instead of waiting weeks or months for data migrations. This accelerates the entire AI development pipeline and speeds up the delivery of business value.



### III. Turning Raw Data into AI-Ready Data via Hammerspace AI Data Platform with NVIDIA

The fusion of the Hammerspace AI Data Platform with NVIDIA's accelerated computing is essential for converting the massive quantities of raw, unstructured data – from the edge, core data center, and multi-cloud environments – into a readily usable format for AI and machine learning models. Traditional data silos and low-performance storage often bottleneck the entire AI pipeline, particularly the crucial data collection, ingest and preparation stages before even reaching the performance heavy computing phase. In alignment with NVIDIA's AI Data Platform blueprint, the Hammerspace AI Data Platform enables a complete file-to-embedding workflow using NVIDIA AI Enterprise software and microservices.

This includes document parsing and chunking through NVIDIA NIM microservices, semantic vector generation with NeMo Retriever, GPU-accelerated index creation via NVIDIA cuVS, and storage of embeddings in a vector database of record. Hammerspace exposes this AI-ready context via APIs to external tools and even natural language queries in addition to making it immediately available to RAG pipelines, agentic AI workflows, and downstream applications across the AI Factory.

This integrated solution fundamentally accelerates the AI journey by optimizing the data path to the GPU. Hammerspace leverages technologies such as Tier 0 for GPU-local NVMe staging to provide the lowest latency access possible, which is critical for demanding NVIDIA AI workflows. By ensuring that NVIDIA's powerful compute resources are constantly fed with data at maximum speed, the platform dramatically reduces idle GPU cycles and improves resource utilization. This efficiency cuts down the time and cost of developing and deploying AI models while scaling easily to handle exascale data growth and increasingly sophisticated AI projects, effectively turning data chaos into AI-ready intelligence.

We see that organizations are rapidly shifting their focus from foundational AI training to maximizing value through AI inference and fine-tuning – critical workloads for realizing production ROI. This paradigm shift makes the virtual data layer provided by the Hammerspace Data Platform strategically vital.

Hammerspace transforms raw, siloed data into AI-ready representations, delivering the high-performance global data environment required for modern AI/ML pipelines, including:

- **Vector Database population**
- **Retrieval-Augmented Generation (RAG)**
- **Data conditioning and feature engineering**

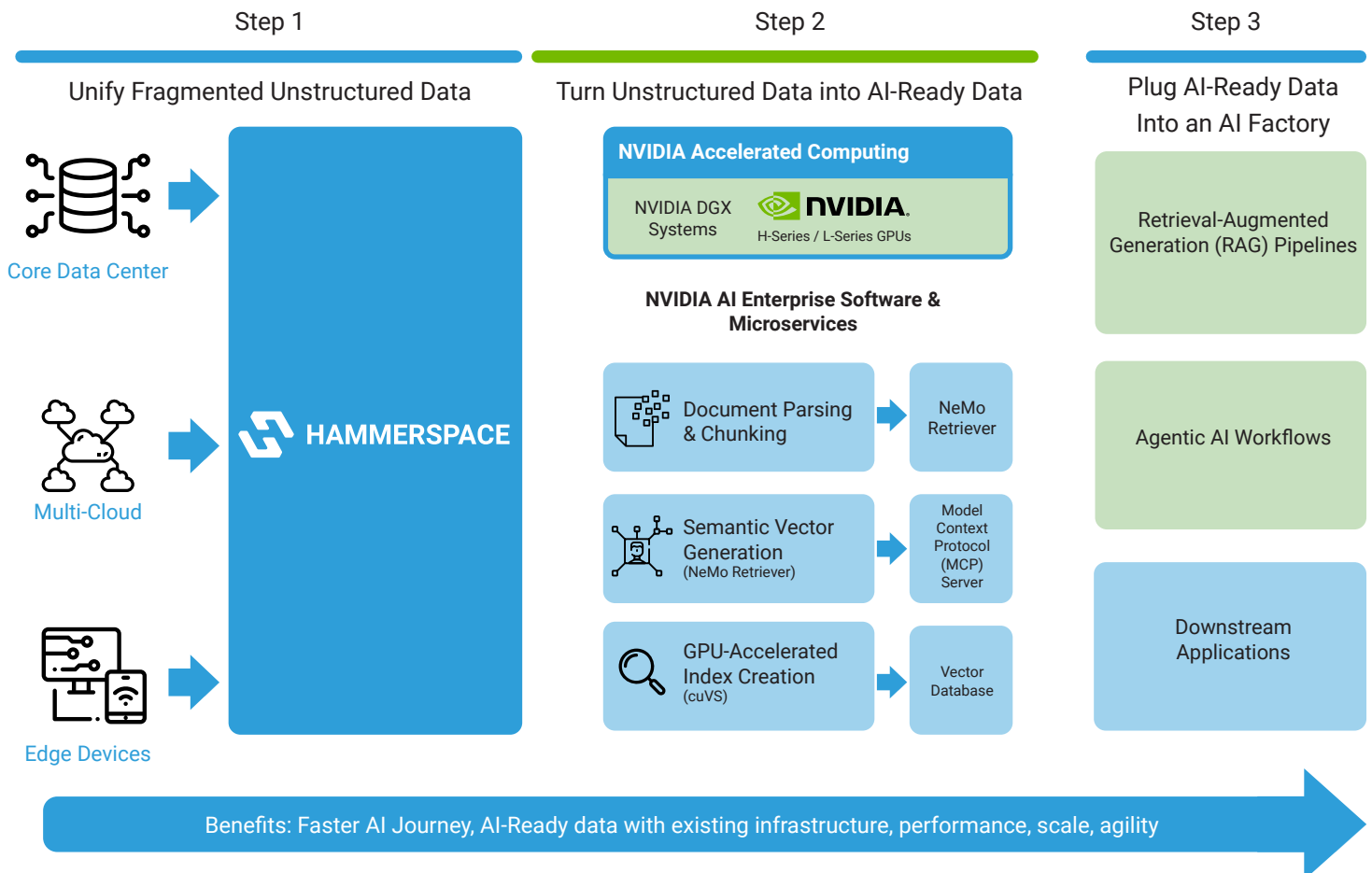
These capabilities complement the high-performance AI Factories architected with the NVIDIA portfolio (e.g., NVIDIA DGX systems, NVIDIA H-series and L-series GPUs, and NVIDIA AI Enterprise software.) By delivering global data mobility and locality-independent access, Hammerspace ensures organizations move away from the system-bound approach of rival implementations that trap data within proprietary silos, instead supplying an organized flow of data directly into the AI Factory.

For instance, Hammerspace dynamically orchestrates the movement and availability of data, enabling pre-placement of raw data directly onto NVMe storage local to NVIDIA GPUs. This ensures ultra-low latency access for high-speed preprocessing and the transformation of raw data into embeddings essential for vector databases.

As a result, Hammerspace uniquely streamlines the process of converting global data into actionable embeddings and contextual inputs.

- **Step 1:** Unify Fragmented Unstructured Data with Hammerspace: Leverage the software-defined, high-performance Hammerspace Global File System to provide ubiquitous, high-speed access to data across any silo, cloud, or geography, ensuring data locality independence.
- **Step 2:** Accelerated Embedding Operations (AI Compute Locality): Use Hammerspace data orchestration policies to pre-place data close to the powerful NVIDIA GPU compute resources for high-speed chunking of source materials and embedding generation.
- **Step 3:** Interoperability with AI Agents & RAG (AI-Ready Data): The newly conditioned, embedded data is immediately available to downstream AI Agents and RAG processes across the entire AI Factory, ensuring seamless and scalable AI interworking for production inferencing.

## Turning Raw Data into AI-Ready Data via the Hammerspace AI Data Platform with NVIDIA





## IV. Conclusion: Why Hammerspace Is Essential for Organization-wide AI Strategy Success

We believe Hammerspace offers a comprehensive, open, and data-centric solution that can fundamentally shift enterprise AI infrastructure from storage-centric to data-centric, directly addressing the core challenges of data silos, high costs, and operational complexity across distributed environments. By establishing a unified global namespace, the platform is positioned to minimize as well as eliminate data silos, ensuring high-performance access to all unstructured data across edge, on-premises, and multi-cloud infrastructure, which is critical for maximizing expensive GPU utilization and accelerating the entire AI development pipeline.

This innovative approach integrates with NVIDIA accelerated computing - leveraging technologies such as local NVMe storage for an optimized data path - to ensure compute resources are constantly fed with AI-ready context, transforming raw data and significantly speeding up workloads like fine-tuning, embedding generation, and RAG. Ultimately, Hammerspace empowers organizations to unlock their AI potential with greater efficiency, flexibility, and clarity, while leveraging existing infrastructure and ensuring consistent governance and data integrity across the global data estate.



#### ABOUT HYPERFRAME RESEARCH:

HyperFRAME Research delivers in-depth research and insights across the global technology landscape, spanning everything from hyperscale public cloud to the mainframe and everything in between. We offer strategic advisory services, custom research reports, tailored consulting engagements, digital events, go to market planning, message testing, and lead generation programs.

Our industry analysts specialize in rigorous qualitative and quantitative assessments of technology solutions, business challenges, market forces, and end user demands across industry sectors. HyperFRAME Research collaborates closely with your Analyst Relations, Product, and Marketing teams to build and amplify your thought leadership, positioning your expertise to enhance brand and product recognition. Through content that engages readers, viewers, and listeners alike, we ensure your voice resonates across channels.

#### CONTACT HYPERFRAME RESEARCH:

**Steven Dickens**

CEO & Principal Analyst | HyperFRAME Research

**Email Address:**

[steven.dickens@hyperframeresearch.com](mailto:steven.dickens@hyperframeresearch.com)

**Telephone Number:**

+1 845 505 1678

**X: @StevenDickens3**

**LinkedIn: Steven Dickens**

**BlueSky: Steven Dickens**

#### CONTRIBUTORS

**Ron Westfall**

VP and Practice Leader for  
Infrastructure and Networking

**Don Gentile**

Analyst-in-Residence  
Storage & Data Resiliency

**Steven Dickens**

CEO and Principal Analyst

#### INQUIRIES

Contact us if you would like to discuss this report and HyperFRAME Research will respond promptly.

#### CITATIONS

This paper can be cited by accredited press and analysts, but must be cited in-context, displaying author's name, author's title, and "HyperFRAME Research." Non-press and non-analysts must receive prior written permission by HyperFRAME Research for any citations.

#### LICENSING

This document, including any supporting materials, is owned by HyperFRAME Research. This publication may not be reproduced, distributed, or shared in any form without the prior written permission of HyperFRAME Research.

#### DISCLOSURES

HyperFRAME Research provides research, analysis, advising, and consulting to many high-tech companies, including those mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

