CAEVES

Whitepaper

Unlocking Engineering Insights with Deep Storage

Executive Summary

A multinational engineering company with global operations across North America, Europe, and Asia approached CAEVES to transform its storage archives. With over 50 years of design, testing, and compliance experience, the organization was handling several petabytes of data, distributed across different data centers, storage platforms and data management tools. Retrieving anything from this layered archive was nearly impossible, especially as engineering firms often repurpose CAD/BIM models from historical projects, which slowed down project delivery and impacted customer satisfaction.

This project, which helped the organization to define a more future-proof data strategy, examined the client's environment, previous systems, and the impact of their current infrastructure on productivity and business performance. Noted issues included slow retrieval of archival data, high costs associated with maintaining data storage, and challenges in utilizing past data to implement new solutions.

CAEVES migrated their data to Microsoft Azure to leveraging its AI-powered Deep Storage™ platform, enabling them to utilize critical functions such as AI-indexed search and Agentic AI, using Microsoft 365 Copilot. Their archive can now be searched at will, setting them up for further AI tool development and easier access to critical archive data by employees.

"...as this AEC firms tried to embrace digital transformation, they were held back by over 50 years of design, testing, and compliance history stored across legacy systems."

Jaap van Duijvenbode, Head of Product & Customer Experience

Significant Challenges for Modern AEC Businesses

For Architecture, Engineering, and Construction (AEC) businesses, data is their lifeblood. From old drawings to building layouts and CAD files, each byte of data represents the company's history and identity. How they approach projects, old workflows, and vital customer information is stored in these legacy archives. Unfortunately, while the sector continues to grow in scale and sophistication, AEC businesses are confronting a range of data management challenges that demand both strategic agility and operational transformation.

The digital transformation of the AEC sector has accelerated, yet integration remains inconsistent, particularly in terms of data storage. Many firms rely on patchwork legacy systems, disconnected design platforms, and siloed data environments. This makes it near impossible to track file versions or keep legacy information accessible, not to mention the limits it places on collaboration and real-time decision-making. Despite the rise of Building Information Modeling (BIM), cloud collaboration tools, and AI-powered tools, legacy data archives remain largely unchanged, effectively serving as a foundation for these businesses.

The Problem with Legacy Archives in AEC

In an industry that thrives on precision, documentation, and accountability, the value of legacy archives in AEC cannot be overstated. Decades of drawings, blueprints, project records, and compliance documents form the backbone of institutional knowledge. These documents represent the history of the company and its evolving identity. Being able to access and use these files can prompt strong internal growth, feed ideas, or protect the business in compliance use cases.

Yet, as this AEC firms tried to embrace digital transformation, they were held back by over 50 years of design, testing, and compliance history stored across legacy systems. The company holds petabytes of archived project files, test results, and maintenance logs on old infrastructure and tape libraries. Their legacy archives have become both an invaluable resource and a significant operational burden.

- **Slow retrieval:** Weeks to access files from tapes or cold tiers.
- **Limited search:** Filename and metadata only, no contextual intelligence.
- **High costs:** Retrieval fees, tape management, and siloed infrastructure.
- **Compliance pain:** Manual, error-prone audit preparation.
- **Innovation Bottlenecks:** Limited ability to use past data for trend analysis or AI modeling.

Traditional archive systems don't work with today's data-rich workflows. Indexed files use filename or basic metadata only, and this approach offers little to no insight into the actual content or relationships between assets. Without intelligent search or contextual recognition, valuable information remains buried, forcing teams to dig through repositories manually. This not only wastes time but also risks overlooking critical project details that could inform current design or engineering decisions.

While these pain points highlight inefficiencies, they also underscore opportunity. Legacy archives represent a rich historical dataset; one that, if digitized and intelligently indexed, can drive future innovation.

The CAEVES AI-powered Deep Storage Solution

To overcome the inefficiencies of legacy archive systems, AEC businesses are turning to modern data infrastructure solutions that merge scalability, intelligence, and compliance readiness. Our approach utilizes CAEVES' cloud-native Archive Data Lake, which uses Microsoft Azure. This solution transforms decades of unstructured archives into searchable, secure, and AI-ready digital assets, effectively turning static repositories into living, value-generating data ecosystems.

t its core, the CAEVES platform solves the most persistent archival challenges faced by AEC firms—eliminating slow retrievals, fragmented data management, and costly tape dependencies. By leveraging AI-Indexed Search, Agentic AI analytics, and low-cost retention, AEC organizations gain immediate, actionable access to their historical records without compromising compliance or cost efficiency.

Legacy Archive Challenge	Deep Storage Solution (CAEVES on Azure)
Slow Retrieval—Weeks to access files from tapes or cold tiers.	AI-Indexed Search enables near-instant retrieval of any file using semantic and metadata-based intelligence.
Limited Search—Filename-only indexing with no contextual insights.	Contextual AI Search scans and understands file content, allowing users to query by concept, project, or keyword.
High Costs—Tape management, retrieval fees, and siloed infrastructure.	Low-Cost Retention eliminates retrieval fees and reduces long-term storage costs through cloud-native efficiency.
Compliance Pain—Manual, error-prone audit preparation.	Audit-Ready Access provides instant retrieval of compliance records and automated data lineage tracking.

CAEVES offers a genuinely unique approach to this company's problems. Since traditional keyword-based searches cannot meet the complexity of AEC documentation, AI-indexed search is the most obvious choice. While employees referencing a single drawing across dozens of projects and revisions might take weeks, an AI-indexed search can systematically sort through every document, image, and CAD file, automatically analyze it, and then create relevant tags. In this way, AI-indexed search enables project teams to locate information by meaning, rather than just by filename, dramatically accelerating retrieval times and reducing the dependency on manual indexing.

Unlike conventional systems that require archived data to be restored before analysis, CAEVES' Agentic AI runs analytics directly within the archive environment. This eliminates costly rehydration processes and allows firms to derive insights—such as material trends, cost histories, or maintenance patterns—from decades of archived data in real time.

The Results

Through Microsoft Azure's secure cloud foundation, CAEVES delivers a resilient, future-ready data protection environment built on encryption at rest and in transit, role-based access control, and automated compliance monitoring. This combination ensures that sensitive design and regulatory data remain protected throughout its lifecycle—without adding administrative complexity.

The platform's low-cost retention model keeps archives both accessible and affordable, eliminating the burden of physical media management while maintaining seamless connectivity to critical business systems. Its audit-ready governance framework further simplifies regulatory reporting, version control, and verification, giving stakeholders confidence in the integrity of every archived record.

With CAEVES, they quickly achieved:

- **90% Faster Retrieval – Cut search time down to minutes instead of weeks.**
- **40% Cost Savings – Eliminated tapes, on-premise archive storage, and retrieval fees. .**
- **Improved Insight – Gained new tools to utilize predictive design and reduce recalls.**
- **Compliance Confidence – Achieved peace of mind with instant audit responses.**

With the Archive Data Lake in place, teams can instantly locate drawings, specifications, and compliance documents that once took weeks to retrieve. The shift from tape-based storage to an AI-indexed, cloud-native architecture has not only eliminated retrieval delays and costs but also opened new pathways for data-driven decision-making.

Rather than serving as a cost center, the company's archives now act as a strategic asset—powering more innovative projects, faster compliance, and sustained innovation.