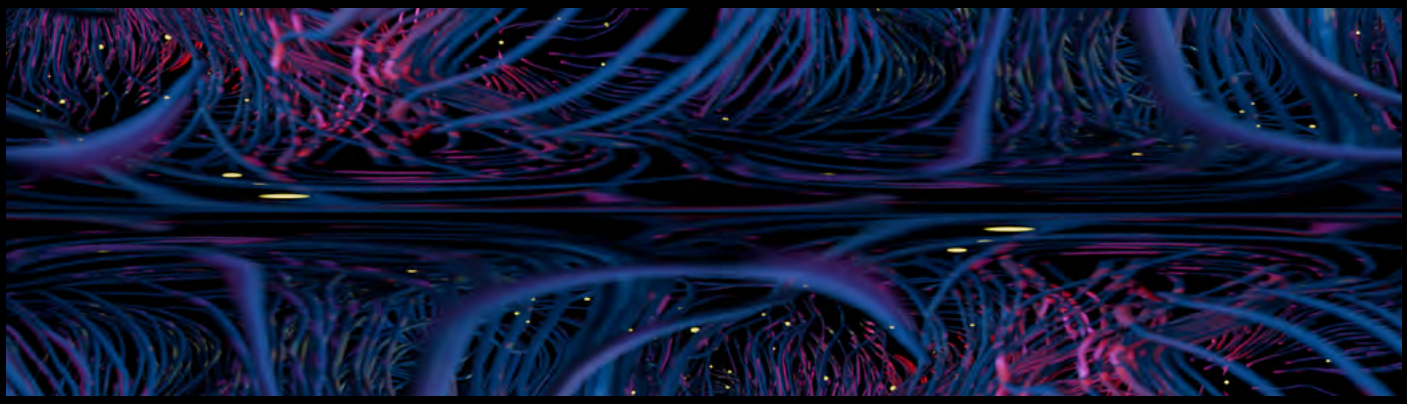
ARCITECTA + GRAU DATA + COMBACK

Scalable Data Lifecycle Management for AI, HPC, and Research Environments



Modern research, HPC, and AI environments are generating massive volumes of unstructured data across performance and archive storage tiers.

Arcitecta, GRAU DATA, and COMBACK deliver an integrated architecture that combines intelligent data orchestration, metadata-driven archive, and enterprise tape infrastructure into a unified data lifecycle platform.

Together, the solution enables organisations to:

- Manage large-scale datasets across active and archive environments
- Maintain visibility and accessibility across the data lifecycle
- Automate policy-driven data movement and retention
- Reduce long-term storage cost and energy consumption
- Scale sustainable archive infrastructure for petabyte- and exabyte-scale environments

The platform connects high-performance workflows with durable long-term archive while preserving operational visibility, governance, and control.

www.arcitecta.com | talk@arcitecta.com

The Challenge

As data volumes expand, organisations face several key infrastructure challenges:

Fragmented Data Visibility

Data is often distributed across high-performance storage, object storage, cloud platforms, and archive systems with limited visibility between tiers.

Escalating Infrastructure Costs

Retaining large inactive datasets on performance storage significantly increases operational and capital expenditure.

Long-Term Archive Complexity

Organisations require durable archive infrastructure capable of preserving data over decades while maintaining accessibility and discoverability.

Sustainability Pressures

Data centres are under growing pressure to reduce energy consumption, infrastructure footprint, and long-term environmental impact.

The Solution

Arcitecta, GRAU DATA, and COMBACK enable organisations to move seamlessly between high-performance workflows and long-term archive without losing operational visibility or control.

MEDIAFLUX
MADE BY ARCITECTA

Intelligent Data Orchestration and Global Data Management

At the core of the solution is Arcitecta's Mediaflux® platform. Mediaflux provides:

- Global namespace management
- Hierarchical Storage Management (HSM)
- Workflow orchestration and automation
- Metadata management
- Storage virtualisation across heterogeneous infrastructure
- Policy-driven data movement
- Secure API-driven integration

Mediaflux creates a unified data fabric spanning performance storage, object storage, cloud platforms, and tape archive. This enables organisations to manage data consistently across multiple storage environments while preserving accessibility and governance.

COMBACK

Enterprise Tape Infrastructure for Sustainable Scale

COMBACK's ORION tape libraries provide the physical archive infrastructure layer for long-term data retention.

ORION systems are designed for:

- Large-scale archive environments
- High-density tape storage
- Modular scalability
- Operational reliability
- Cost-efficient long-term retention

The platform supports scalable enterprise archive deployments for organisations managing petabyte- and exabyte-scale datasets.

GRAU DATA
Your data \ Your control _

Metadata Intelligence and Scalable S3-to-Tape Archive

GRAU DATA extends the architecture with deep metadata intelligence and scalable archive integration.

MetadataHub

MetadataHub extracts and indexes metadata from hundreds of complex file formats, enabling:

- Advanced search and discovery
- Rich metadata visibility
- Faster data analysis and classification
- Improved governance and compliance
- Greater operational insight into unstructured datasets

Integrated with Mediaflux, MetadataHub enables organisations to maintain visibility and discoverability across both active and archived data.

XtreemStore

XtreemStore provides scalable S3-to-tape archive integration for long-term retention of large inactive datasets.

Key capabilities include:

- S3 Glacier-compatible archive workflows
- Automated data tiering to tape
- Scalable archive management
- Long-term data preservation
- Integration with modern object storage workflows

The platform enables organisations to reduce dependence on high-cost performance storage while maintaining access to archived content.

Tape Infrastructure for the Exabyte Era

Modern research and HPC environments require an archive infrastructure that can scale sustainably over decades, not just years.

Tape continues to provide a foundational technology for large-scale long-term data retention due to its:

- Low energy consumption
- Long media lifespan
- High durability
- Strong cyber resilience
- Exceptional storage economics

By integrating modern metadata management and S3 archive workflows with enterprise tape infrastructure, organisations can create sustainable archive architectures without sacrificing accessibility or operational efficiency.

Use Cases

Research and Scientific Computing

Manage large experimental and simulation datasets while maintaining long-term accessibility and archive sustainability.

AI and Machine Learning Workflows

Support the retention, governance, and lifecycle management of large-scale training and research datasets.

HPC Environments

Optimise the movement of data between performance storage and scalable archive infrastructure.

Long-Term Research Retention

Preserve valuable institutional and research data across decades while maintaining discoverability and governance.

Proven in Research Environments

The integrated Arcitecta and GRAU DATA solution was selected by the Technical University of Dresden to optimise large-scale tape archive while maintaining visibility and access to archived research data.

By combining Mediaflux and XtreemStore, the University benefits from scalable archive infrastructure, metadata-rich discoverability, and efficient long-term retention workflows.