Large Swiss Bank

Turns Siloed Historical Content into a Compliant Digital Archive

DataBeagle consolidated hundreds of millions of historical objects into a searchable archive while preserving compliance requirements and enabling retirement of a legacy platform.

Overview

A large Swiss bank needed to retire a major legacy content platform while preserving access to historical information required for compliance, audit, and ongoing business use.

The estate contained approximately 200 TB of data and 400 million objects spanning SharePoint, NFS, CIFS, object storage, and file-drop locations. Moving the data alone would not solve the problem: metadata and retention requirements had to be preserved, migration integrity had to be validated, and scanned content had to become searchable.

Using DataBeagle Migration and DataBeagle Intelligence, the bank consolidated its historical content into a compliant digital archive and retained an active discovery experience after the legacy platform was retired.

KEY RESULT

SILOED → SEARCHABLE

Historical content transformed into a compliant archive

Key Stats

SOURCE ESTATE	OBJECTS
~200 TB	~400 million
OCR PROCESSED	**PROJECT DURATION**
~40 million	~7 months

SOURCES

SharePoint · NFS · CIFS · Object

USE CASE

Compliant archive · Retrieval · Retirement

Preserving Access While Retiring a Legacy Platform

THE CHALLENGE

The bank had accumulated hundreds of millions of historical files and documents that still carried regulatory, retention, and business value. A significant portion remained subject to compliance requirements, while authorized users still needed to retrieve historical information when required.

The challenge was broader than migration. Content had to be transferred consistently from multiple protocols and systems, key metadata and source information had to remain intact, retention requirements had to be preserved, and the integrity of the migration had to be verifiable.

Scanned and image-based documents also needed to become searchable. Without these capabilities, retiring the legacy environment risked simply moving inaccessible content from one archive to another.

What had to change

01 Consolidate sources

Move SharePoint, NFS, CIFS, object and drop-location content through one controlled migration framework.

02 Preserve compliance context

Carry forward metadata, source paths, timestamps and retention requirements into the target archive.

03 Make history usable

Validate every phase, OCR scanned content, then provide active search after the legacy system is retired.

WHY DATABEAGLE

Migration designed for long-term usability

1 Multi-protocol migration

DataBeagle Migration supported SharePoint, NFS, CIFS, object storage and file-drop sources within a common migration framework.

2 Metadata & retention preservation

Key metadata, timestamps, ownership, original source paths and retention requirements were preserved so information remained meaningful and compliant.

3 OCR at scale

Approximately 40 million scanned and image-based documents were OCR-processed, bringing previously non-searchable content into full-text discovery.

4 Searchable after migration

DataBeagle Intelligence became the primary discovery and retrieval interface for the archive after the legacy platform was retired.

From Data Migration to a Searchable Digital Archive

MIGRATION + ARCHIVE WORKFLOW

01 → 02 → 03 → 04

Migrate

Move content from heterogeneous sources into HCP through a controlled one-time migration.

Validate

Reconcile object counts, verify integrity and metadata, reprocess failed items, and produce migration audit reports.

OCR & enrich

Process scanned and image-based documents so their text becomes searchable alongside born-digital content.

Index & discover

HCI indexes the archive and DataBeagle Intelligence becomes the primary search and retrieval experience.

BUSINESS IMPACT

A compliant archive that remains actively usable

LEGACY RETIRED

Successful platform decommissioning

The legacy content platform was retired after migration and validation while historical information remained available through DataBeagle.

400M OBJECTS

Historical content preserved

Approximately 400 million objects were consolidated into the target archive, with roughly 40 million scanned/image documents OCR-processed.

30-40% EST.

Infrastructure savings

Retiring the legacy platform and consolidating retained information is estimated to reduce related infrastructure and platform costs.

50-70% EST.

Faster compliance discovery

With content consolidated, indexed and searchable through DataBeagle, compliance discovery and review are estimated to be significantly faster.

FROM SILOED HISTORY TO A COMPLIANT DIGITAL ARCHIVE

DataBeagle transformed historical content from a legacy dependency into an actively searchable enterprise information asset.