

LM01_KDD020 - Legal hold users data retention mechanism (Vault / Purview)

Status	PENDING DECISION
Owner	Damien Avril
Stakeholders	Frank Bolata, Boris Foiselle, Eric Triffaux, Hachem Osmani, James Kyndt

Decision:

Decision made by:

Date:

Online Meeting:

Issue

Migrating data from Google Vault to the Microsoft 365 Purview environment while ensuring **no data loss** and **maintaining the same functional capabilities** previously guaranteed by Google Vault (with the complement of Exterro).

This is due to the decommissioning of Google services and Google Vault in order to leverage the capabilities of Microsoft 365.

Recommendation

Option 3: Extract Google Vault data to a secure repository (GCP) + use Purview for active & new legal holds + retire Exterro

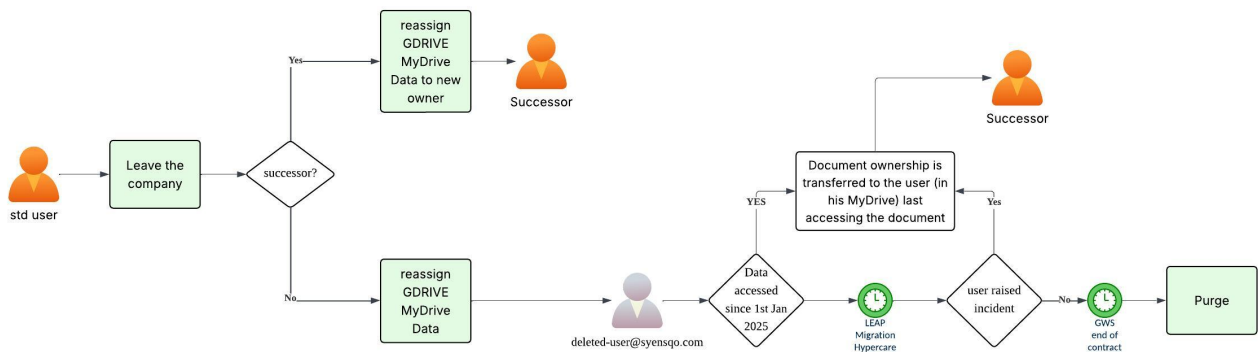
Background & Context

Historically, the legal and compliance departments have relied on a combination of Google Vault for native data retention and Exterro as the primary orchestration layer for managing legal holds and complex discovery workflows. However, as part of the broader infrastructure modernization and per the directive from the CISO Office, the Exterro platform is being phased out alongside Google services.

This shift necessitates a full transition to Microsoft Purview, which now serves as the centralized environment for information governance. The transition is not merely a technical migration of mailboxes and files, but a fundamental change in how the organization meets its statutory and litigation-related obligations.

For data move please refer to article **LM01_KDD021 - GAM to support migration (GVault / Deleted Users / Talarian)**: <https://wiki.syensqo.com/x/iDlzOg>

Current Deleted users Ownership reassignment process



Coexistence Period

During the migration from Google Workspace to Microsoft 365, "Legal Hold Subjects" will be managed in a split-governance model to ensure no data is lost during the transition.

The Strategy

- Microsoft Purview: Governs migrated Mailboxes and OneDrive (formerly MyDrive) files. Holds are applied here immediately upon migration.
 - Will hold data for new holds on active users right after the migration
 - Will hold data for new holds on users leaving the company right after the migration.
 - Will hold data for sensitive users right after the migration as per data retention policy.
- Google Vault: Governs Google Chat and any legacy data not yet decommissioned. Users must remain on hold in Vault until Google services are fully disabled.
 - Stores data for active users under litigation.
 - Stores deleted data for users under litigation.
 - Stores data for "Archived users" who left the company.

During the coexistence:

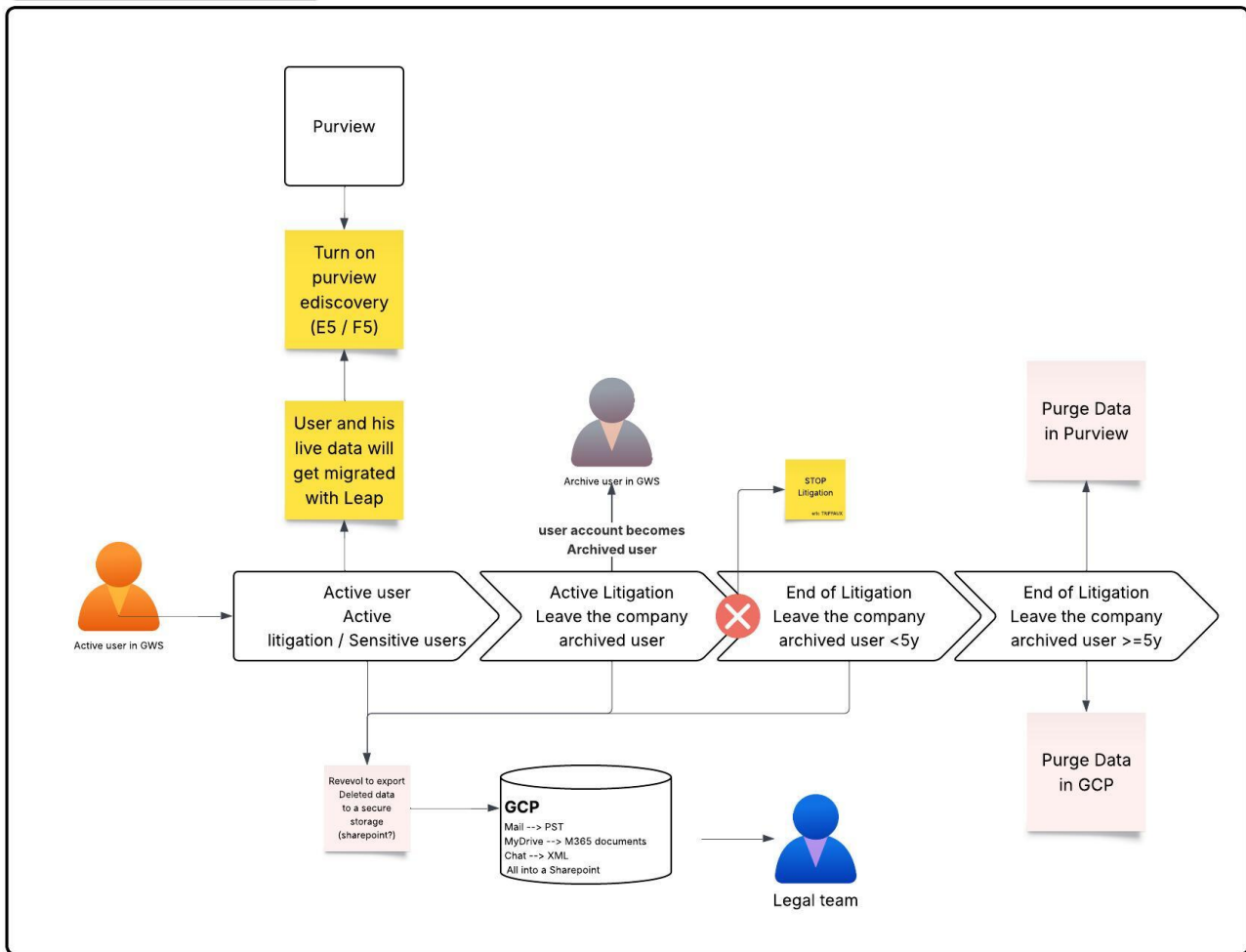
- The new cases and sensitive users being migrated to Microsoft will be covered by Purview.
- The ongoing holds will be stays in GVault for archived users, and for active users deleted data.

Deleted User Data: GCP Cold Storage

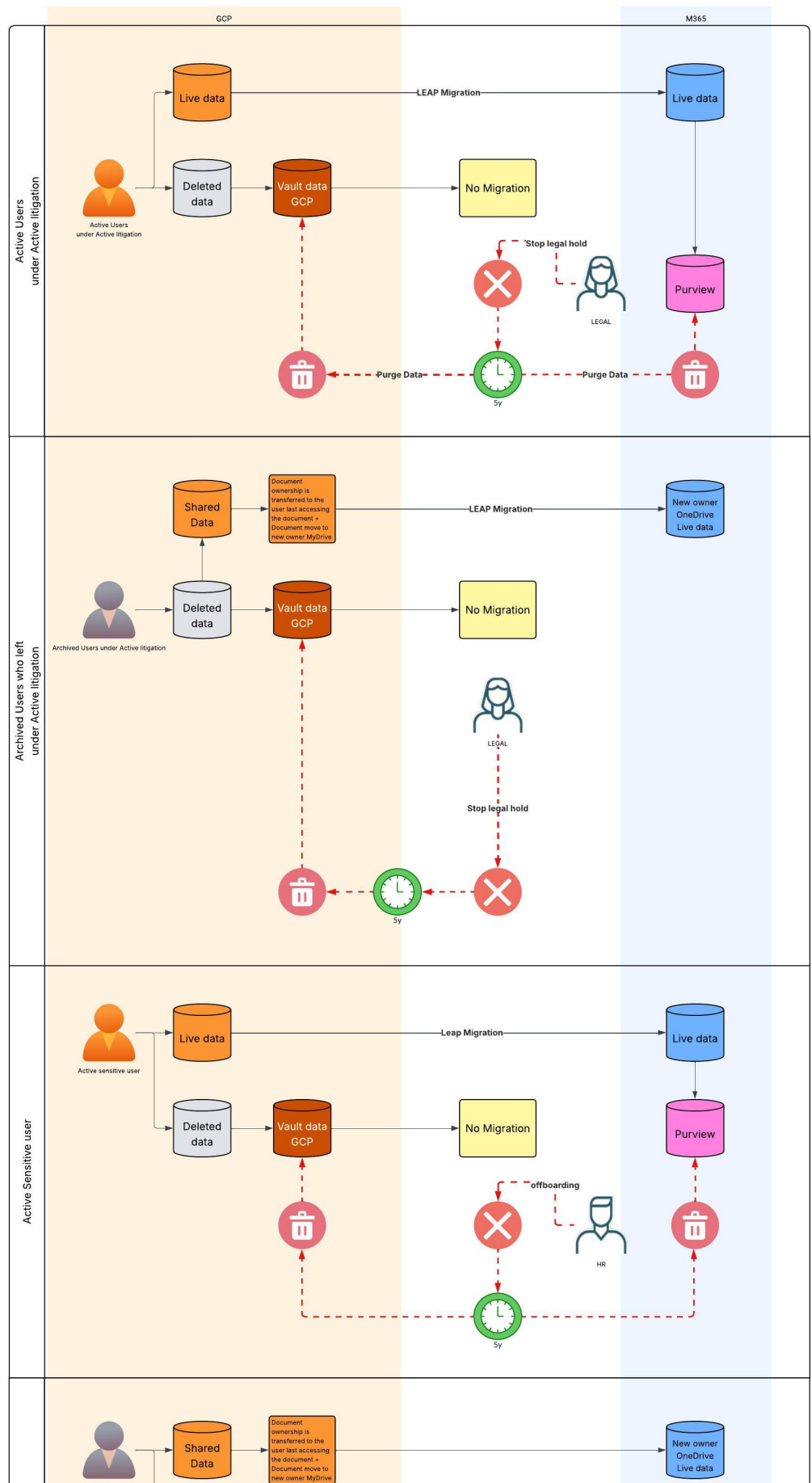
To reduce licensing costs while maintaining compliance, data for deleted users will be transferred to a Google Cloud Platform (GCP).

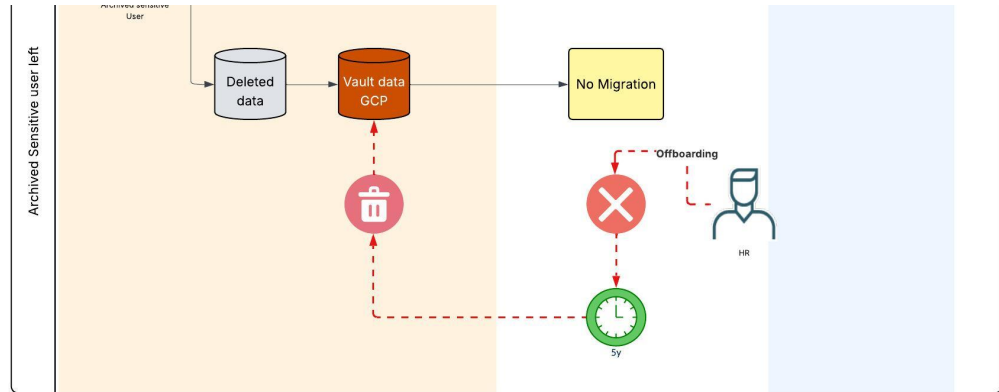
1. For archived users, before a Google account is **deleted**, OR for Active users on hold (litigation or with sensitive role) a full export of Vault data is performed.
 - a. This data is moved to a GCP Archive bucket.
 - i. Retention:
 1. All exported data are on hold during the active litigation case, until it ends.
 2. All exported data after sensitive role departure, or litigation case is over, are subject to a strict **5-year retention policy** to meet regulatory requirements.
 3. After 5 years, the GCP lifecycle policy **will automatically purge the data unless a manual legal extension** is applied.

Leap Project data migration



GVault to Purview use cases





Operational Impact:

Storing deleted users' data in GCP while retaining new information in Microsoft 365 creates a fragmented discovery environment, complicating workflows for the Legal team.

Assumptions

- **Secure Staging:** The GCP storage will serve as a secure, high-capacity staging area where data integrity (hashing) is verified before ingestion into M365.
- **Chain of Custody:** The migration process will preserve all essential metadata (creation dates, authors, recipients) to ensure the data remains legally defensible.
- **Tooling & Connectivity:** The VM will be equipped with necessary migration tools and sufficient network bandwidth to handle large-scale data transfers between Google and Microsoft.
- **Formatting Parity:** Google-native files (Docs, Sheets) will be successfully converted to Microsoft-compatible formats during the VM staging phase to allow for full indexing by Purview.
- **Capacity:** The destination Microsoft 365 environment has been provisioned with adequate storage to accommodate the total volume of legacy data from Google Vault.

Constraints

- Migration from GVault to Microsoft Purview is not supported, while the metadata are lost.
- Migrating Data from GVault into Microsoft purview without metadata are not exploitable.
- Migration of data from **Gvault** will be a process that may take up to **4–5 weeks**.
- The migration can start only after the project completes the **“Coexistence”** phase. Users under legal hold must no longer be able to continue using **GWS**, because for the migration a set of **closed/static data is required**.
- During coexistence, data keeps changing/being overwritten, which prevents the migration.
- Migration timeline should be set to **early October 2026**.

Impacts

- **Lower long-term risk through structured migration:** Moving Google Vault data to a secured repository and then closing Vault can simplify the legacy footprint—provided holds/retention and chain-of-custody documentation are recreated—making compliance posture clearer over time.
- Exported Vault data may be **harder to search/use in Purview** (format/metadata/indexing gaps)
- **Manageable change with short transition phase:** Expect a temporary ramp-up (training, updated procedures, initial validation), but day-to-day operations should stabilize quickly once Purview processes and access to legacy data are validated.

Retention Policy

Services	Active std user	Active + litigation	Active vip / sensitive
Teams chat (personnal)	10 y	Retain forever	10 y
Teams channel message (within Teams sharepoint)	10 y	Retain forever	10 y
Mail	10 y	Retain forever	10 y
Mail trash Bin	30d + (admin 14d < 30d)	Retain forever	30d + (admin 14d < 30d)

Mail junk	30d + (admin 14d < 30d)	Retain forever	30d + (admin 14d < 30d)
OneDrive	Retain forever	Retain forever	Retain forever
OneDrive trash Bin	30d + (admin 93d)	Retain forever	30d + (admin 93d)
Sharepoint	Retain forever	N/A	Retain forever
Sharepoint trash Bin	30d + (admin 93d)	N/A	30d + (admin 93d)
License when active user	E5/F3	E5/F3	E5/F3

Options considered

Option 1: Maintain Google Vault for legacy data + Purview for new cases + Exterro (notifications only)

- keep Google Vault for legacy Google data (need to license all existing vaulted users active and inactive ~2200 x7\$/user = 15400 yearly)
- Purview eDiscovery for new litigation on M365 : covered by E5 / F5 add-on
- reduce Exterro scope for litigation notification only (missing in eDiscovery) : cost to be defined

Option 2: Extract Google Vault data to secure repository (GCP) + Purview for new cases + Exterro (notifications only)

- Extract GVault data into a secured place (to be defined) today using GCP project
- New Microsoft Purview (eDiscovery) for new litigation: covered by E5 / F5
- Reduce Exterro scope for litigation notification only (missing in eDiscovery): cost to be defined

Option 3: Extract Google Vault data to secure repository (GCP) + Purview for new cases + retire Exterro

- Extract GVault data into a secured place (to be defined) using GCP project
- New Microsoft Purview (eDiscovery) for ongoing active users & new litigation: covered by E5 / F5
- **Stop Exterro** and utilize (eDiscovery) functionality

Evaluation

Feature	Legacy: Google Vault + Exterro	New: Microsoft Purview (eDiscovery)
Architecture	Multi-vendor: Hybrid integration between Google Cloud and a third-party legal platform.	Native: Unified, "in-place" solution built directly into the Microsoft 365 stack.
Data Collection	API-dependent: Data is copied/transferred from Google Vault into Exterro for processing.	In-Place: Search and review happen directly where the data lives, eliminating transfer time.
Legal Hold Management	Centralized Orchestration: Exterro manages notifications and triggers holds via API in Vault.	Policy-Driven: Holds are managed via Purview cases; notification tracking is a manual or Power Automate task.
User Interface	Legal-Centric: Designed specifically for lawyers and forensic investigators.	Admin-Centric: Integrated into the broader Compliance/IT Admin center.
Advanced Analytics	Exterro Smart Lab: Proprietary AI for threading, deduplication, and predictive coding.	Security Copilot: Native generative AI for summarizing threads and identifying key evidence.
Platform Scope	Cross-Platform: Can simultaneously search Slack, Zoom, Box, and other non-Google sources.	Ecosystem-Specific: Optimized for M365; requires specific connectors for non-Microsoft data.
Metadata Integrity	Risk of Loss: Metadata can occasionally be altered or lost during the export/import process.	Highest Integrity: No data movement is required for search, ensuring 100% metadata preservation.
Licensing Cost	Dual-Cost: Requires both Google Workspace licenses and a separate Exterro subscription.	Consolidated: Included in the existing Microsoft 365 E5 / Compliance subscription.

The shift to **Purview** reduces architectural complexity and licensing costs, but moves the organization from a specialized "Legal-First" tool (Exterro) to a "Platform-First" tool.

The primary trade-off is the loss of Exterro's robust **Legal Hold Notification** automation, which now requires a defined internal process within the Microsoft environment.

See also

The following section describes relevant documentation:

Description	Repository
-------------	------------

Workshop presentation	1.3 Teams & Collaboration Design
Workshop meeting minutes	Workshop meeting minutes
HLD: 6.Collaboration, page 99, 6.6Onedrive Sync page 103	1. SYENSQO-Microsoft365 High-Level Design v1.docx

Version	Published	Changed By	Comment
CURRENT (v. 13)	May 05, 2026 14:58	TODESCHINI-ext, Gautier	
v. 12	May 05, 2026 11:54	TRIFFAUX, Eric	
v. 11	May 05, 2026 11:50	TRIFFAUX, Eric	
v. 10	May 05, 2026 09:05	TODESCHINI-ext, Gautier	
v. 9	May 05, 2026 08:46	TODESCHINI-ext, Gautier	

[Go to Page History](#)