Mixpeek Data Retention Policy

Version: 1.0 **Effective Date:** April 15, 2026 **Last Reviewed:** April 15, 2026
Owner: Mixpeek Engineering

1. Purpose

This policy defines how long Mixpeek retains customer data, operational data, and audit records, and the procedures for secure deletion when retention periods expire.

2. Scope

This policy applies to all data stored, processed, or transmitted by Mixpeek systems, including customer-uploaded content, derived data (vectors, embeddings), metadata, audit logs, and operational data.

3. Retention Schedules

3.1 Customer Data

Data Type	Storage Location	Default Retention	Deletion Trigger	Method
Uploaded files (documents, images, video)	AWS S3	Duration of subscription	Account deletion or namespace deletion	S3 object deletion
Extracted features (vectors, embeddings)	Qdrant Cloud	Duration of subscription	Namespace deletion	Qdrant collection deletion
	MongoDB			

Data Type	Storage Location	Default Retention	Deletion Trigger	Method
Document metadata		Duration of subscription	Namespace or document deletion	MongoDB document removal
Namespace configuration	MongoDB	Duration of subscription	Namespace deletion	MongoDB document removal
Collection pipeline config	MongoDB	Duration of subscription	Collection deletion	MongoDB document removal

3.2 Lifecycle Rules

Mixpeek supports declarative retention rules at the namespace level through the Retention Policy system:

Trigger Type	Description	Example
MAX_AGE	Delete documents older than N days	`max_age: 90d` — delete after 90 days
MAX_IDLE	Delete documents not accessed in N days	`max_idle: 30d` — delete if untouched for 30 days
MAX_COUNT	Delete oldest documents when count exceeds N	`max_count: 100000` — cap at 100K documents
MAX_STORAGE	Delete oldest documents when storage exceeds N bytes	`max_storage: 10GB` — cap at 10GB

Actions: `delete` (permanent removal), `archive` (metadata only), `cold` (move to cold storage).

Grace periods prevent immediate deletion of recently created documents.

3.3 Namespace TTL

Namespaces support an optional `ttl_seconds` field. When set, the namespace and all its data are automatically deleted after the specified

duration. This is used for ephemeral workloads such as E2E testing and demo environments.

3.4 Operational Data

Data Type	Storage Location	Retention Period	Notes
Audit logs	ClickHouse	365 days	Immutable, auto-purged via ClickHouse TTL
API request logs	ClickHouse	90 days	Sampled, used for analytics and billing
Canvas telemetry (errors, vitals, events)	ClickHouse	90 days	Client-side error and performance data
Canvas request logs	ClickHouse	30 days	HTTP request metadata
Batch job records	MongoDB	180 days	Processing history for debugging
Webhook delivery logs	MongoDB	30 days	Delivery receipts and retry history

3.5 Account Data

Data Type	Storage Location	Retention Period	Notes
User profiles (email, name)	MongoDB	Duration of account	Deleted on account closure
Organization records	MongoDB	Duration of account	Cascading deletion of all org resources
API key records	MongoDB	Duration of account	Hashes only; plaintext never stored
Billing records	Stripe	Per Stripe retention policy	Managed by Stripe
Auth records	PropelAuth / Clerk	Per provider retention	Managed by auth provider

4. Data Deletion Procedures

4.1 Customer-Initiated Deletion

Customers can delete their data through:

1. **Document deletion** — `DELETE /v1/namespaces/{ns}/documents/{doc_id}` removes the document from Qdrant and MongoDB.
2. **Namespace deletion** — `DELETE /v1/namespaces/{ns}` removes all documents, collections, and configurations for that namespace.
3. **Account deletion** — Triggered through support or the API. Executes `delete_organization_resources_flow()` which cascades across all stores:
4. All namespaces and their data (Qdrant + S3 + MongoDB)
5. Organization records (users, API keys, webhooks, tasks)
6. ClickHouse analytics data (`ALTER TABLE ... DELETE WHERE internal_id = {id}`)

4.2 Automated Deletion

- Retention rules are evaluated periodically by the retention enforcement service.
- Each deletion is recorded in the `RetentionAuditEntry` for compliance tracking.
- Namespace TTL enforcement runs as a background task.

4.3 Secure Deletion

- MongoDB documents are removed via standard `deleteOne` / `deleteMany` operations.
- Qdrant points are removed via the Qdrant API, which handles physical deletion during compaction.
- S3 objects are deleted via the AWS S3 API. Versioning-enabled buckets require explicit version deletion.
- ClickHouse data is deleted via `ALTER TABLE ... DELETE` which marks rows for asynchronous physical removal.

4.4 Backup Considerations

- Deleted data may persist in backups for up to 30 days (the backup retention window).
- After 30 days, all backup copies containing the deleted data will have been rotated out.
- DR backups follow the same 30-day lifecycle.

5. Data Subject Requests

For data subject access requests (DSARs) under GDPR, CCPA, or similar regulations:

1. **Access requests** — Contact info@mixpeek.com. Data export provided within 30 days.
2. **Deletion requests** — Processed via account deletion flow. Confirmed within 30 days.
3. **Correction requests** — Contact info@mixpeek.com. Corrections applied within 30 days.

6. Exceptions

Retention periods may be extended when required by: - Legal hold or litigation preservation notice - Regulatory investigation - Contractual obligation with customer

All exceptions must be documented and approved by the security owner.

7. Policy Review

This policy is reviewed annually and updated when retention requirements change due to regulatory, contractual, or operational needs.

This document is maintained by Mixpeek Engineering. Questions should be directed to info@mixpeek.com.