SCRS Security Whitepaper

The architecture behind Other Me’s governed AI platform — Scoped Cryptographic Retrieval Shield.

By Pop Hasta Labs Ltd · www.pophastalabs.com/security/

Version 2 · 2026

This whitepaper summarises the technical architecture published in full on the public Trust Center. It is written for a CISO, compliance officer, DPO, or regulator doing a deliberate evaluation — not for a web visitor skim-reading in five minutes.

Live interactive demo: **securitylayerdemo.pophastalabs.com**

Specific questions: security@pophastalabs.com (2-business-day SLA on compliance questionnaires).

SECTION 01

Position

SECURITY IS ARCHITECTURE, NOT A FEATURE

Most AI governance products work by letting data through and scanning logs afterwards to tell you what leaked. That approach is structurally incapable of *preventing* a leak — at best it reports the cost later.

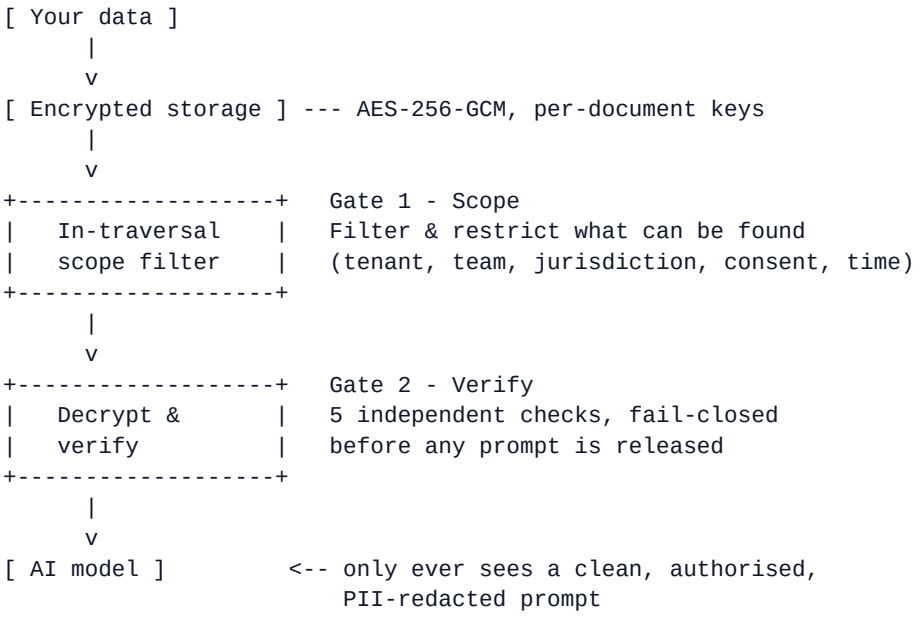
Other Me is architecturally different. Every query from a user to a model passes through two cryptographic gates that enforce scope and entitlement *before* the model ever sees data. If a gate fails, the query is refused; there is nothing to scan afterwards because no data was released.

We call this the **Scoped Cryptographic Retrieval Shield** (SCRS). It is the subject of UK Patent Application 2602911.6, filed January 2026 (pending).

SECTION 02

The two-gate flow

Blocked data never reaches the AI model. There is no downstream “review what was sent” workflow because nothing unauthorised was sent.



SECTION 03

Gate 1 — scope-constrained retrieval

WHAT MOST RAG SYSTEMS DO

Run a vector similarity search across all documents; return the top 100 nearest neighbours; *then* check access controls and discard anything unauthorised. This leaks information through embedding adjacency — even documents you throw away after the fact have already influenced which authorised documents were selected.

WHAT SCRS DOES

Authorisation is pushed into the graph-traversal step itself. Unauthorised documents are never candidates at any depth of the similarity search.

```
# Fail-closed: build_scope_filter raises -> empty queryset
include_q, exclude_q = build_scope_filter(user)
chunks = (FileChunk.objects
    .filter(collection__enabled=True)
    .filter(include_q)        # tenant, team, parent, time, jurisdiction
    .exclude(exclude_q)
    .annotate(distance=CosineDistance('embedding', query_vec))
    .order_by('distance')[:k])
# pgvector evaluates WHERE during the HNSW walk - unauthorised
# chunks are never candidates for the top-k at any depth.
```

Scope predicates: tenant (organisation), team, parent (for child accounts), time (content-aging rules), jurisdiction (e.g. EU-resident staff can only retrieve EU-resident documents). All evaluated during the pgvector HNSW walk, not afterwards.

SECTION 04

Three-store architecture

Customer data is deliberately split across three independent stores, each with a different role and a different threat model. Compromise of any one store alone does not yield usable data.

Store	Contents	Compromise yields
Encrypted blob store	Original document bytes, AES-256-GCM encrypted with per-document DEK	Opaque ciphertext — garbage without the DEK.
Redacted vector index	Chunk embeddings + scope metadata. PII already replaced with opaque tokens before embedding.	Vectors and tokens like [[PERSON_1]]
PII vault + audit log	Mapping from opaque tokens back to real PII values, kept in a hardened, append-only store	Locked to the original authenticated caller; any attemp

A meaningful retrieval requires the caller to have authorised access to all three stores simultaneously. An attacker who exfiltrates “the database” receives one piece of a three-piece puzzle.

SECTION 05

Gate 2 — verify before reveal

Once Gate 1 returns candidate chunks, Gate 2 performs five independent checks. Any single failure aborts the query.

#	Check	What it does
1	Re-verify caller authorisation	Re-evaluate tenant, role, team, parental scope, consent at retrieval time — not at login time. Rev
2	Decrypt only authorised chunks	Per-chunk DEKs unwrapped via UK KMS, bound to caller session and original scope predicates.
3	Integrity & tamper check	HMAC-SHA256 over chunk ciphertext + DEK id + scope metadata. A mismatched tag (corruption
4	PII pseudonymisation (defence in depth)	Even though indexed vectors were PII-redacted at ingest, Gate 2 re-scans decrypted plaintext wi
5	Tamper-evident audit write	Hash-chained audit entry records caller, scope predicates, chunk ids, outcome. Previous entry's

Any failure = fail closed. The LLM never sees partial data. The user is told the request could not be authorised, and the audit entry records why.

SECTION 06

PII vaulting and rehydration

Personally identifiable information is never sent to an LLM provider in plaintext. The tokens are opaque pointers into the PII vault (store 3 above). The LLM provider — whether OpenAI, Anthropic, Google, or xAI — never sees real values.

BEFORE THE QUERY REACHES THE LLM

```
User prompt:
   "Find me the email for Sarah Johnson who works at
    Meridian Capital on the Q3 audit engagement."

What actually leaves the UK boundary:
   "Find me the email for [[PERSON_1]] who works at
    [[ORGANISATION_1]] on the [[ENGAGEMENT_1]]."
```

AFTER THE RESPONSE COMES BACK

```
LLM response:
   "[[PERSON_1]]'s email is sarah.j@[[ORGANISATION_1]]-domain.com"

Rehydrated on the authenticated caller's device:
   "Sarah Johnson's email is sarah.j@meridiancapital.com"
```

Rehydration happens on the caller’s device, only for the authenticated caller. On disk, in memory, and in transit, real PII stays inside the tenant boundary.

SECTION 07

Encryption at rest, in transit, in use

Layer	What we use
At rest	AES-256-GCM with per-document Data Encryption Keys (DEKs). DEKs are wrapped by Key Encryption Ke
In transit	TLS 1.3 with approved cipher suites. HSTS with 1-year max-age and preload. No fallback to TLS 1.2 or olde
In use (Enterprise)	Bring-Your-Own-Keys (BYOK). Enterprise clients can supply their own KMS key; we never hold the decrypt

KEY HIERARCHY

```
Master KEK (UK KMS, outside the app perimeter)
  |
  +-- Tenant KEK (per organisation)
        |
        +-- Document KEK (per file)
              |
              +-- Chunk DEK (per chunk, wrapped in storage)
```

A key revocation at the tenant level invalidates all downstream keys instantly. No re-encryption required — the data becomes cryptographically unreadable on revocation.

SECTION 08

Data residency

All customer data is stored in the United Kingdom. This includes:

- Conversations (every message, every thread)
- File uploads (original ciphertext + derived chunks + embeddings)
- Encryption keys (UK KMS)
- Audit logs (append-only, hash-chained)
- Backups

What leaves the UK: LLM prompts — but only after PII pseudonymisation at Gate 2. What the US-based (or EU-based, or globally-replicated) LLM provider actually receives is a pseudonymised prompt with opaque tokens where names used to be. Real customer data does not cross the UK boundary in plaintext.

SECTION 09

Key revocation — the cryptographic kill switch

When an employee leaves, is offboarded, or has access suspended:

- The admin triggers key revocation for that user.
- The user's session is invalidated.
- All DEKs associated with that user's scope become unusable.
- Any ciphertext they previously had access to is now cryptographically unreadable — for them, and for anyone impersonating them.

This is structurally different from "we remove their login." Login removal just means they can't authenticate. **Key revocation** means that even if they held a copy of ciphertext in their laptop's memory, it is now unreadable because the unwrapping key is gone.

WHERE THIS MATTERS MOST

- Regulated professional-services firms where departing staff could otherwise take client information with them.
- NHS and education where Caldicott and safeguarding rules require immediate access termination.
- Any firm with acrimonious-departure risk — a disgruntled ex-employee who might attempt exfiltration.

SECTION 10

Bring-Your-Own-Keys (Enterprise)

Enterprise clients can bring their own key material. In BYOK mode:

- Master KEK is held in the client's KMS (AWS KMS / Azure Key Vault / GCP KMS), never in ours.
- Every decryption operation requires a signed request to the client's KMS.
- If the client revokes our access to their KMS, we instantly become unable to decrypt anything they have stored with us.
- The client can cryptographically "wipe" their data from our platform by revoking KMS access — without waiting for us to do anything.

This is the strongest privacy guarantee we can offer. It is available on Enterprise contracts (from £297/month).

SECTION 11

Tamper-evident audit trail

Every access, every retrieval, every model inference writes an entry to an append-only hash-chained audit log. Each entry includes:

- Caller identity
- Timestamp (nanosecond precision, NTP-synced)
- Scope predicates active at the time of the call
- Chunks accessed (IDs only, never content)
- Model inference outcome (authorised / refused / error)
- Hash of the previous entry

Because each entry embeds the previous hash, any attempt to delete or modify a historical entry invalidates every entry written afterwards. An auditor can verify the chain from any point back to the genesis entry. Truncation is detectable. Silent deletion is detectable.

For regulated clients (FCA, SRA, Caldicott, ICAEW), this provides the forensic foundation their regulators expect for incident investigations and inspections.

SECTION 12

Compliance posture

Framework	Status
UK GDPR	Full compliance — lawful bases documented, DPIA conducted, subject-rights requests fulfille
Data (Use and Access) Act 2025	Full compliance — in force since February 2026; tracked directly in the architecture, not a retr
Age Appropriate Design Code	Full compliance — relevant to the Family tier with child accounts.
EU AI Act	Risk-assessed; we do not operate systems classified as high-risk under the Act, but the SCR
ICAEW / SRA / FCA / Caldicott	Aligned — sector-specific documentation available to regulated-sector clients on request.
SOC 2 Type II	Roadmap — targeting 2026 H2.
ISO 27001	Roadmap — targeting 2026 H2.

For specific compliance questions from a prospective client — especially regulated-sector firms — email security@pophastalabs.com. We respond to compliance questionnaires within 2 business days.

SECTION 13

What this means for your clients

Three sentences a client can say to their own auditors, regulators, and customers after adopting Other Me:

1. *Our customer data never leaves the UK in plaintext.*

True, because PII is pseudonymised at Gate 2 before any prompt leaves the UK boundary.

2. *Access control is enforced at the retrieval layer, not just at the interface.*

True, because Gate 1 filters during the graph traversal itself — unauthorised documents are never candidates, at any depth.

3. *If we revoke an employee's access, their access to historical data is cryptographically terminated, not just logged out.*

True, via key revocation at the tenant/user level — ciphertext becomes unreadable, not just invisible.

These three sentences are not available to a client using ChatGPT Plus, Google Gemini, Microsoft Copilot for M365, or any of the enterprise AI platforms outside the £10,000/month price band. They are the specific commercial value of the SCRS architecture, and they are why the product justifies its price.

SECTION 14

Further reading

- Public Trust Center — **www.pophastalabs.com/security/**
- Live interactive demo — **securitylayerdemo.pophastalabs.com**
- Data Processing Agreement — **www.pophastalabs.com/data-processing-agreement/**
- Privacy Policy + DPIA summary — **www.pophastalabs.com/privacy-policy/**
- Responsible AI statement — **www.pophastalabs.com/responsible-ai/**
- Acceptable Use policy — **www.pophastalabs.com/acceptable-use/**
- Subprocessor list — on request to security@pophastalabs.com.
- Penetration test summary — on request under NDA.

This whitepaper is distilled from the live Trust Center. If anything here has changed since the version date on the cover, the Trust Center is the authoritative source — email us if something in this PDF is inconsistent with what is on the site, and we will reconcile.

Questions?

security@pophastalabs.com — 2-business-day SLA on compliance questionnaires. For anything that needs a call, include preferred times and your team’s jurisdiction.

Pop Hasta Labs Ltd · Registered in England and Wales · Public Trust Center: www.pophastalabs.com/security/ · App: otherme.pophastalabs.com