Security & Compliance

Questions this doc answers

- How does Reflect Memory stay compliant with SOC 2 / HIPAA / regulated controls?
- What audit logs are available for every read/write?
- How do we ensure agents can only reach approved LLM hosts?
- Who handles telemetry and data residency permissions?

Controls & coverage

- **SOC 2 Type II alignment** is in progress for the hosted environment. The same controls (per-tenant RBAC, encrypted transport, rate limits) apply in isolated/self-host pilots.
- **HIPAA boundary** is enforced via self-host deployment, `disableModelEgress`, and private key storage. No data leaves the deployment unless you explicitly open webhooks.
- **SSO** – `RM_SSO_ENABLED=true` plus `RM_SSO_JWKS_URL`, `RM_SSO_ISSUER`, `RM_SSO_AUDIENCE` (with optional `RM_SSO_EMAIL_CLAIM`) bootstraps OIDC login. Fails fast if required values are missing.
- **Agent keys** are validated with timing-safe compares; agents cannot access non-MCP routes.

Audit trail

Every read, write, auth, and admin action writes a row in `usage_events`. The backlog feeds:

- Billing plans (`PLAN_LIMITS`, usage quotas)
- Compliance exports (JSON/CSV)
- Admin metrics (`/admin/metrics`, `dashboard page`).

Logs capture: `user_id`, `memory_id`, `operation`, `ip_address`, `timestamp`, `device`, `origin`.

Telemetry & observability

- Cloud deployments collect minimal telemetry for service health; self-hosts default telemetry to off with `RM_DISABLE_TELEMETRY`.
- Rate limits: 100 req/min per IP; admin routes tighten to 10 req/min.
- Fail-safes: WAL mode ensures SQLite availability; Postgres migration adds partitioned `usage_events`, ready for SIEM ingestion.

Guardrails

- `allowed_vendors` gating is enforced via SQL: `EXISTS (SELECT 1 FROM json_each(m.allowed_vendors) WHERE value = '*' OR value = ?)`.
- `deleted_at` ensures soft delete; no normal query sees trashed memories.
- `team / share_memory` flows add explicit references so team reads and writes are auditable separately.