Librarian: IT & Security One-Pager

For: IT leads, managed service providers, and technical reviewers evaluating Librarian on behalf of a law firm, insurance agency, accounting practice, or other regulated organization.

Purpose: Describe what Librarian is, where it runs, what data it handles, and what leaves your network — with enough specificity that your review can proceed without needing to contact us for basics.

> *This document is designed to be readable by both human reviewers and AI assistants. You're welcome to share it with AI tools for summary, analysis, or question-answering during your evaluation.*

At a glance

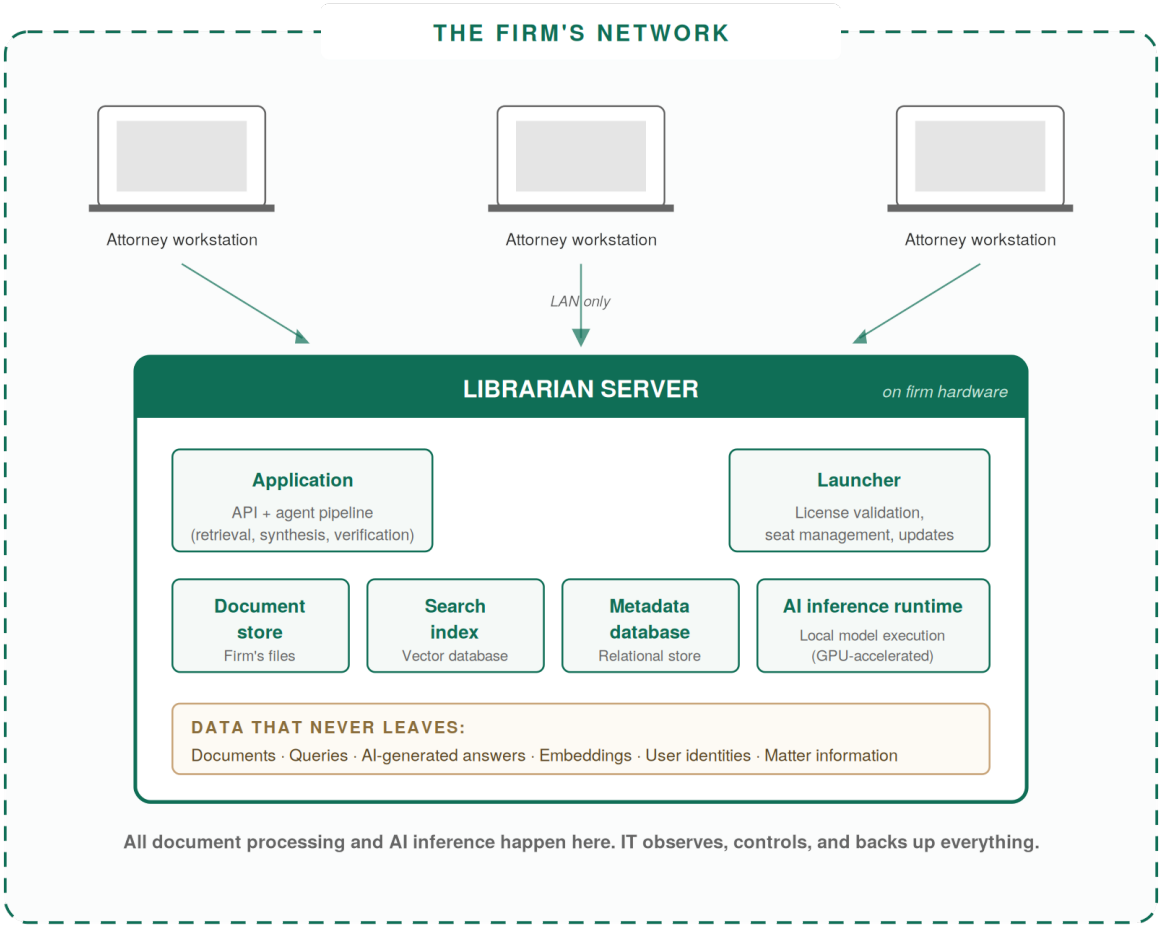
Librarian is a document intelligence and AI question-answering application that **runs entirely on your organization's own hardware**. Documents, queries, indexes, AI model weights, and generated outputs remain on machines you control. There is no vendor cloud holding your firm's data.

The application is delivered as a Windows MSI installer. Once installed, it operates as a set of local services on the user's machine or on a dedicated server within your network. A small companion application, the Launcher, handles licensing, certificate management, and update notifications.

During normal operation, Librarian makes a short list of outbound connections (license validation, update checks, optional certificate renewal, optional base model downloads) — described below. No customer documents, queries, embeddings, or AI outputs are transmitted in any of them.

Architecture at a glance

The diagram below shows every component inside your network and every outbound connection Librarian makes. This is the authoritative reference; narrative descriptions that follow correspond to this diagram.



FOUR OUTBOUND CONNECTIONS

Every network call that crosses the firm's boundary, fully enumerated.

1 · LICENSE HEARTBEAT

Destination: api.umelle-librarian.ai
Frequency: every 5 minutes
Transmits: license key, device ID, timestamp
Purpose: seat validation, subscription renewal
Always on · Required

2 · SOFTWARE UPDATES

Destination: updates.umelle-librarian.ai
Frequency: user-triggered (check or download)
Transmits: current version number
Purpose: new version availability, installer download
User-controlled · Required for updates

3 · CERTIFICATE RENEWAL

Destination: api.umelle-librarian.ai (DNS proxy)
Frequency: ~every 60 days (auto-renew)
Transmits: license key, LAN IP, cert metadata
Purpose: HTTPS for multi-user LAN access
LAN mode only · Optional

4 · BASE MODEL DOWNLOAD

Destination: huggingface.co
Frequency: one-time per base model
Transmits: model name, optional HF token
Purpose: acquire base model before FORGE training
FORGE only · Optional

No other outbound connections exist. Document content, queries, and AI outputs are never transmitted.

Librarian architecture — components inside the firm's network and the four outbound connections

What gets installed

Component	Purpose	Runs where
Librarian application	Document ingestion, indexing, search, AI Q&A	User's machine or firm server
Librarian Launcher	License validation, update management, certificate provisioning	Same machine as Librarian, foreground only
Local relational database	Application metadata, user accounts, chat history	Same machine
Local vector database	Document embeddings	Same machine
Ollama	Third-party local AI runtime, MIT-licensed. Required for inference.	Same machine

All components run locally. None of them communicate externally by default. FORGE, the optional fine-tuning feature, is a module within the Librarian application and uses the same local infrastructure.

Application files are installed under `Program Files`. User data (documents, databases, embeddings, AI models) is stored in standard per-user Windows application-data directories, inspectable directly on the host machine.

What stays on your machines

The following data never leaves the environment where Librarian is installed:

- Documents uploaded to Librarian (original files stored on the host as regular files)
- Extracted text, structured parsing results, OCR outputs
- Document embeddings and vector database contents
- Chat history, user queries, and AI-generated responses
- FORGE training data, training pairs, adapter weights, exported model files
- AI model weights downloaded through Ollama
- TLS certificates, private keys, and license keys (encrypted at rest using OS-level key protection)

All AI inference — document enrichment, embedding generation, chat responses, OCR — executes via local loopback calls to the Ollama runtime on the host machine. These calls do not leave the host.

What leaves your machines, and why

1. License heartbeat

- **Where it goes:** UMELLE's license server
- **How often:** Periodic, while the Launcher is running
- **What it contains:** License key, random device identifier, timestamp

- **What it does NOT contain:** Documents, queries, outputs, hardware fingerprints, file paths, or any data from Librarian's databases

Offline behavior: If the heartbeat fails, existing users continue operating Librarian normally for up to 30 days. After 30 days of failed heartbeats, the application will decline to launch until connectivity is restored. See "User management" below for one important caveat about administrative operations during a disconnect.

2. Update availability check

- **Where it goes:** UMELLE's update server
- **How often:** Periodic, while the Launcher is running
- **What it contains:** Nothing (HTTPS request for a static manifest file)
- **What it does:** Checks whether an update is available

Updates are never installed automatically. The user must explicitly accept an update for anything to be downloaded or installed. All update packages are SHA-256 hash-verified before installation.

3. TLS certificate renewal (LAN Mode only)

If you enable LAN Mode for multi-user access on your firm's network, the Launcher provisions per-license HTTPS certificates via the Let's Encrypt ACME protocol.

- **Frequency:** Approximately every 60 days
- **What is transmitted:** License key, ACME account public key, certificate signing request, DNS challenge value, your LAN IP address
- **What is NOT transmitted:** Certificate private key (generated locally, stored locally, never leaves the machine)

LAN Mode disclosure: Enabling LAN Mode publishes a DNS name under UMELLE's domain that points to your private LAN IP address. This DNS name is visible in public DNS. Your LAN IP is a private (non-routable) address that cannot be reached from outside your network, but the DNS record itself is public. If zero public DNS footprint is required, use Offline Mode. LAN Mode is opt-in.

4. AI base model downloads

Base model weights are downloaded directly from their source by Ollama (for inference models) and, for the optional FORGE feature, from Hugging Face (for FORGE base models). These connections are initiated by those components, not by the Librarian application or UMELLE's servers.

- **What is transmitted:** Standard HTTP request metadata only. No user data, no documents.
- **What is received:** Model weight files, cached locally

These are one-time downloads per model. Once cached, no further network activity is required to use the model.

5. Transactional email (from UMELLE's servers, not yours)

When UMELLE sends you a license-related email (license delivery, trial expiration reminder, support ticket response), it is sent from UMELLE's license server via Microsoft Graph API. No data from your Librarian installation is involved in these emails.

User management during a disconnect

One operational note that deserves explicit disclosure: **administrative operations — adding users, revoking seats, changing roles — require the application to be in contact with the license server.** This is a deliberate design choice. Seat entitlement and policy changes are signed by the license server via RS256 JWT; the application verifies these tokens rather than trusting arbitrary local configuration.

What this means in practice:

- **Existing users can continue to use Librarian normally** during a disconnect, up to the 30-day offline grace period. Document queries, AI responses, FORGE operations, and all end-user features continue to work.
- **Administrative changes to the user roster require connectivity.** If your network is down, you cannot add a new user or revoke an existing user until connectivity is restored.

Plan deployments accordingly. For the common case (existing users working through a brief connectivity issue), this has no effect. For the uncommon case (you need to terminate an employee's access during a multi-hour outage), connectivity to the license server is required.

Data at rest

Librarian does not currently implement application-level encryption of document content at rest. Data-at-rest protection relies on the disk encryption provided by the host operating system (BitLocker on Windows).

Recommendation: Deploy Librarian on machines with BitLocker enabled. For Firm and Firm+ deployments, this is treated as a required part of the reference deployment configuration.

License keys, TLS private keys, and certificate account keys are individually encrypted at rest using OS-level key protection bound to the user account, regardless of disk encryption state.

Compliance posture

Librarian's architecture supports compliance obligations under the frameworks listed below. These statements describe architectural support for your compliance work, not vendor certifications.

Framework	Support
GDPR (EU) 2016/679	Consent captured at purchase; 3-year consent record retention; right-to-erasure endpoint; Standard DPA available on request for Firm and Firm+ tiers
California ARL	Consent records retained per ARL (3 years from consent or 1 year after license termination, whichever is longer)
EU Consumer Rights Directive	Withdrawal waiver captured at checkout per Article 16(m); confirmed on durable medium per Article 8(7)
EU AI Act (Regulation 2024/1689)	Architecture supports Article 10 data governance and Article 12 record-keeping obligations for deployers. Section 16 of the Librarian EULA is a transparency statement covering the Software's AI functionality, human oversight mechanisms, and provider/deployer distinction.
HIPAA (45 CFR Part 164)	UMELLE is not a Business Associate: no Protected Health Information is transmitted to or processed by UMELLE's systems. PHI remains on hardware the Covered Entity controls. Product design supports the §164.312 technical safeguards — access control, audit controls, integrity, and transmission security. Encryption at rest (addressable specification under §164.312(a)(2)(iv)) is supported via host-level disk encryption (BitLocker recommended). A HIPAA architectural documentation deliverable can be prepared on request for qualified evaluations.

UMELLE does not make compliance determinations on your behalf. Your firm remains the data controller for documents processed within Librarian and is responsible for assessing and meeting any obligations applicable to your use.

What happens if UMELLE ceases operations

A legitimate concern for any vendor relationship. For Librarian specifically:

- Your software continues to function for the remainder of your paid subscription period, plus the 30-day offline grace period. This is enforced by the Launcher locally, not by the license server's availability.
- Your documents remain on your machines. UMELLE has no technical ability to access, modify, or delete them.
- Your FORGE-trained models are standard GGUF files that run in any compatible environment, with or without Librarian. You retain them indefinitely.
- Your database contents are exportable using standard database tools.

Unlike cloud-based legal AI vendors, there is no risk of data becoming inaccessible because a vendor service went offline. Your data's continued availability depends only on your own infrastructure.

Support model

Tier	Response SLA	Notes
Personal	Best-effort	No SLA commitment
Team	Best-effort	No SLA commitment
Firm	**48 hours**	Account and licensing issues
Firm+	**24 hours**	Account and licensing issues

Important scope note on SLAs: Our support commitment is to *respond* within the stated window, and to resolve account, licensing, and user-management issues within it.

Librarian is a desktop application. System-level issues — bugs, feature defects, or behavior changes — are resolved through software updates made available to all customers, not through customer-specific deployments. When a system-level issue is identified in your environment, our commitment is to acknowledge it within the SLA window and include the fix in a subsequent update. We do not deploy custom patches to individual customer installations.

Contact

For questions beyond the scope of this document, including formal DPA requests, deeper architectural questions, or onboarding assistance:

Email: support@umelle.com
Website: https://umelle-librarian.ai
Company: UMELLE Ltd., Manastirski Livadi, ul. Sharl Shamo 3, 1404 Sofia, Bulgaria

Firm and Firm+ tier customers receive onboarding assistance during initial deployment. Additional technical documentation is available on request for qualified evaluations.

Document Version: 1.2 · April 2026
Document type: Technical reference for IT security review. Not a contract. For contractual commitments see the Librarian End-User License Agreement available at https://umelle-librarian.ai/eula.