LIVRE BLANC OFFICIEL DE CONFORMITÉ · ÉDITION 2026

Intelligence Artificielle Souveraine : Guide de Conformité CNIL, RGPD, DORA & EU AI Act

Spécifications juridiques et d'architecture technique pour le déploiement de moteurs RAG sans exfiltration de données.

Livre Blanc : Conformité RGPD, Secret Professionnel & RAG Souverain (Yevi)

- **Auteur** : Odoun AI & Équipe Yevi Core
- **Public** : DPO, RSSI, Directeurs Juridiques, Notaires & Avocats
- **Version** : 1.0 (Août 2026)

---

1. Le Défi de la Protection des Données dans les RAG Entreprise

L'adoption de l'intelligence artificielle générative dans les secteurs régulés (juridique, santé, finance, secteur public) heurte deux exigences légales majeures :

1. **Le Règlement Général sur la Protection des Données (RGPD)** et les directives de la CNIL concernant la transmission non consentie de données à caractère personnel (PII) à des processeurs tiers.

2. **Le Secret Professionnel et Notarial** (ex: Article 66-5 de la loi du 31 décembre 1971 pour les avocats, secret notarial), interdisant toute divulgation de documents de dossiers clients à des tiers non habilités.

L'utilisation de solutions SaaS distantes basées sur des API externes cloud expose l'organisation à des risques de sous-traitance non conforme, de fuite de données par ré-entraînement de modèles et de sanctions administratives.

---

2. L'Architecture de Sécurité Yevi : Le Triangle de Confiance

Yevi résout ces exigences réglementaires grâce à une architecture à triple barrière sécuritaire :

graph TD

A["Document Confidentiel / Requete"] --> B["1. Anonymisation PII & Coffre AES-256-GCM"]

B --> C["2. Ingestion & Search Hybride Local"]

C --> D["3. LLM Local-First / Ollama / On-Premise"]

D --> E["4. Verification des Citations & Restitution"]

E --> F["5. Journal d'Audit Append-Only JSONL"]

Barrière 1 : Masquage PII Réversible avec Coffre-Fort Chiffré AES-256-GCM

- Avant toute soumission au moteur vectoriel ou au modèle de langage, le texte subit une détection hybride PII par expressions régulières (emails, IBAN, numéros de Sécurité Sociale, téléphones) et reconnaissance d'entités nommées spaCy (`PER`, `ORG`, `LOC`).
- Chaque entité est remplacée par un jeton anonyme (ex: `[[PER_1]]`, `[[ORG_3]]`).
- **Coffre-fort AES-256-GCM** : Le dictionnaire de correspondance est immédiatement chiffré en mémoire RAM volatile à l'aide d'une clé symétrique AES-256 de session. Il n'est **jamais stocké en clair sur disque**.

Barrière 2 : Traitement 100 % Local-First (Zero Cloud Leak)

- Les embeddings (`SentenceTransformers`) et la base vectorielle (`ChromaDB`) s'exécutent entièrement sur l'infrastructure locale ou sur le serveur d'étudeOn-Premise.
- Le LLM s'exécute localement via Ollama ou un runtime On-Premise dédié. Aucune donnée ne franchit le périmètre du réseau local de l'organisation.

Barrière 3 : Traçabilité & Citations Anti-Hallucination

- Le prompt système contraint le modèle à fournir l'identifiant exact de la source `[doc_id#passage_idx]` pour chaque affirmation.
- Si le fonds documentaire ne contient pas la preuve source, la réponse est marquée comme non-ancrée ("Is Grounded: False").

---

3. Matrice de Conformité aux Exigences CNIL & RGPD

| Exigence RGPD / CNIL | Mécanisme de Réponse Yevi | Statut |

| --- | --- | --- |

| **Art. 5.1(c) - Minimisation des données** | Masquage strict PII pré-indexation et pré-génération | **Conforme** |

| **Art. 32 - Sécurité du traitement** | Chiffrement AES-256-GCM en RAM volatile du dictionnaire PII | **Conforme** |

| **Art. 28 - Souveraineté & Sous-traitants** | Exécution 100% Local-First / On-Premise (Zero sous-traitant Cloud) | **Conforme** |

| **Art. 30 - Registre des activités de traitement** | Journal d'audit append-only JSONL traçant requêtes et passages cités | **Conforme** |

| **Secret Professionnel (Art. 66-5 / Secret Notarial)** | Aucune fuite réseau sortante, isolation totale dans l'étude | **Conforme** |

---

4. Recommandations de Déploiement pour DSI & DPO

1. **Option On-Premise Réseau Isolé** : Déployer Yevi via `docker-compose.yml` sur un serveur local ou un NAS d'étude sans accès Internet sortant.

2. **Relecture "Zero-Trust"** : Pour les dossiers juridiques ultra-sensibles, activer l'affichage des jetons anonymisés (`[[PER_1]]`) sur la console Web Yevi avant l'exportation finale.

3. **Conservation des Journaux d'Audit** : Sauvegarder les fichiers JSONL générés dans `core/audit` pour répondre aux contrôles périodiques du DPO.
