ecluse

L'IA SOUVERAINE ET MAÎTRISEE

PRESENTATION

SPÉCIFICATIONS TECHNIQUES ET SÉCURITÉ

2025

www.ecluse.co

Synthèse du document

L'intégration de l'intelligence artificielle est un levier incontournable de modernisation du service public. Toutefois, face aux risques de fuite de données et aux incertitudes juridiques, les collectivités doivent impérativement se doter d'un cadre sécurisé pour protéger leurs informations sensibles et celles de leurs administrés.

Le document présente « Ecluse », une architecture hybride garantissant la **souveraineté des données** : une ingestion et une **pseudonymisation des documents** réalisées exclusivement en local ("on-premise"), couplées à l'utilisation de modèles d'IA hébergés en Europe pour les traitements génératifs. Il explicite également la **conformité** aux cadres réglementaires (**RGPD, doctrine CNIL, AI Act**) à travers des mécanismes d'audit, de journalisation et de minimisation des données. Cette approche permet à la DSI et au DPO de transformer un usage à risque en un **service managé**, offrant aux agents les **gains de productivité de l'IA** dans un cadre technique maîtrisé et auditable.

SOMMAIRE

1. Les enjeux pour les collectivités

1.1. Modernisation, pression documentaire et risques

Les collectivités font face à trois dynamiques simultanées :

Une augmentation continue du volume documentaire (délibérations, marchés, courriers, notes internes, rapports d'étude, procédures RH, etc.).

Des attentes accrues des citoyens et des élus en termes de réactivité, de pédagogie et de qualité de l'information produite.

Une adoption spontanée d'outils d'IA grand public par certains agents ("shadow AI"), en dehors de tout cadre maîtrisé, avec des risques majeurs de fuite ou de transfert de données sensibles vers des prestataires non contrôlés.

L'IA générative est donc à la fois un levier de productivité (rédaction, synthèse, aide à la décision) et une source de risque si elle est utilisée hors infrastructure maîtrisée. L'intégration de l'IA nécessite une stratégie de sécurisation documentée et conforme.

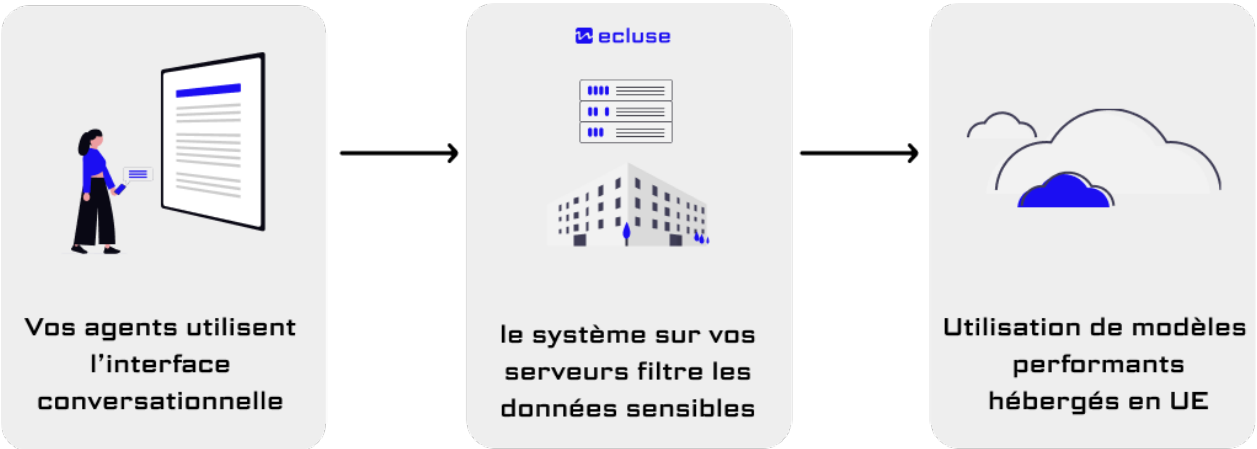
Les risques sont massifs et déjà documentés

20% des **documents** uploadés contiennent des **données sensibles**

4% des **prompts** envoyés contiennent des **données sensibles**

Source Harmony, études sur 20000 documents et 1 million de prompts

1.2. Positionnement de la solution Écluse



Ecluse est conçue comme un **"sas de sécurité et de conformité"** qui s'intercale entre les agents et les modèles d'IA externes. Elle sécurise l'intégralité des échanges à travers une interface conversationnelle unique, en traitant distinctement les deux sources de données :

- **Le flux conversationnel (Prompts) :** Les questions posées par l'agent dans le chat sont **interceptées par le moteur local**. Elles subissent un processus de détection et de masquage des données sensibles (noms, adresses, numéros) *avant* d'être envoyées au modèle.
- **Le flux documentaire (Fichiers) :** L'agent importe ses documents de travail (PDF, DOCX) dans l'interface conversationnelle. Ceux-ci sont intégralement traités en local (extraction de texte, découpage et sauvegarde dans une base de donnée vectorielle). Seulement les sections en rapport avec la demande de l'agent sont pseudonymisées et ajouté au contexte de la conversation avec le modèle externe.

Ainsi, l'intégralité des traitements sensibles est réalisée **(on-premise)**. Seuls des fragments de documents anonymisés et des questions épurées ("prompts assainis") sont transmis aux modèles d'IA hébergés en Europe (Zero Data Retention), garantissant qu'aucune donnée identifiante ne quitte le périmètre de la collectivité.

2. Cadre réglementaire : RGPD, CNIL, AI Act – synthèse des exigences

2.1. RGPD : minimisation, proportionnalité et responsabilité

Le RGPD impose :

- La minimisation des données : N'envoyer que ce qui est strictement nécessaire au traitement.
- La responsabilité du responsable de traitement : La collectivité reste juridiquement responsable, même en cas de recours à des prestataires.
- La documentation : Registre de traitement, analyse d'impact (DPIA).

Ecluse fournit les mécanismes techniques (anonymisation locale) pour respecter ces obligation.

2.2. CNIL : maîtrise des flux et souveraineté opérationnelle

La CNIL recommande aux organismes publics :

- D'éviter l'envoi de données sensibles vers des services d'IA non maîtrisés ou hors UE.
- De mettre en place un pré-traitement local (pseudonymisation) avant tout envoi externe.
- De privilégier des modèles exécutés localement ou des prestataires européens.

Écluse répond à ces points via un moteur local de pseudonymisation (ML PII, Regex, LLM local) et une architecture on-premise pour tous les traitements identifiants.

2.3. AI Act : transparence et usage responsable

L'AI Act introduit des obligations de transparence et de documentation des systèmes d'IA. Ecluse permet de documenter finement :

- Les flux de données (Collecte → Pseudonymisation → Requête → Restitution).
- Les modèles utilisés (Modèles locaux pour la sécurité, modèles européens pour la génération).
- Les mesures de réduction des risques.

3. Architecture technique détaillée : une IA sous contrôle

3.1. Vue d'ensemble

1. L'architecture Ecluse est déployée sur un serveur de la collectivité (ou hébergement souverain).
2. Couche d'ingestion locale : Traitement des fichiers déposés par l'utilisateur (PDF, DOCX, etc.).
3. Couche de pseudonymisation multi-niveaux : Masquage des données personnelles.
4. Couche RAG contextuel (Local) : Recherche vectorielle limitée aux documents de la conversation.
5. Couche d'orchestration : Appels sécurisés vers LLM externes (UE).
6. Couche de ré-identification : Restitution lisible pour l'agent.

3.2. Installation on-premise

Pré-requis : Serveur Linux (ou VM) dans le SI, base de données (PostgreSQL), moteur vectoriel (ex: Qdrant/pgvector), accès réseau sortant restreint aux API IA européennes.

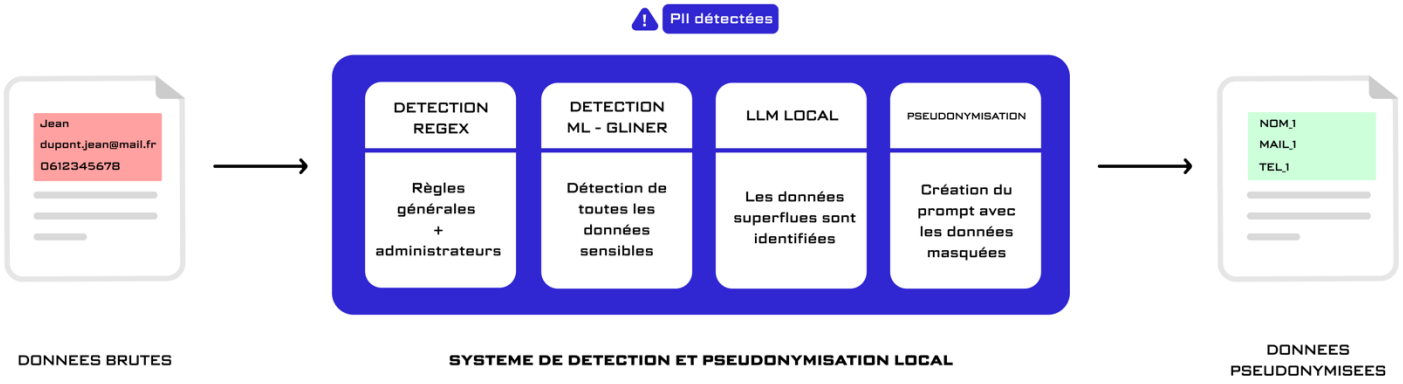
Déploiement : Installation des services backend, des modèles locaux (embeddings, détection PII, SLM (Small Language Model) spécialisé) et configuration de la passerelle externe.

3.3. Couche 1 – Ingestion et traitement contextuel (Session utilisateur)

Lorsqu'un agent dépose un document dans le chat :

- Le fichier est traité temporairement sur le serveur local.
- Le texte est extrait et découpé en fragments
- Ces fragments sont vectorisés par un modèle d'embedding local.
- Isolation des données : L'index vectoriel créé est associé exclusivement à la session de l'utilisateur. Aucun autre agent de la collectivité n'a accès à ce contenu. Le document ne vient pas enrichir une base globale partagée.

3.4. Couche 2 – Pseudonymisation multi-couches



Avant tout usage, le contenu subit une chaîne d'assainissement locale :

1. Détection ML : Noms, adresses, dates, téléphones.
2. Règles métiers (Regex) : Numéros de dossiers, codes internes (Configurable par les administrateurs)
3. Analyse de rôle (LLM Local) : Un SLM (Small Language Model) détermine si un nom est crucial pour le contexte (ex: "Maire") ou doit être masqué.
4. Tokenisation : Remplacement par des jetons (ex: [[PERSONNE_1]]). La table de correspondance reste sur le serveur local.

3.5. Couche 3 – Moteur RAG local (Recherche sémantique ciblée)

Lorsque l'agent pose une question sur son document :

1. La question est vectorisée localement.
2. Le système recherche dans les fragments du document de la session en cours les passages pertinents.
3. Seuls ces extraits (déjà pseudonymisés) sont préparés pour l'envoi au modèle externe.

Cela garantit que l'IA ne "lit" que les paragraphes nécessaires pour répondre, et non le document entier.

3.6. Couche 4 & 5 – Appels externes et Restitution

Envoi : Les extraits pseudonymisés sont envoyés à un modèle européen (Zero Data Retention).

Réception et Ré-identification : La réponse générée contient les jetons ([[PERSONNE_1]]). Le serveur local les remplace par les vraies valeurs avant d'afficher la réponse à l'agent. L'utilisateur final voit une réponse claire, sans savoir que l'IA a traité des données masquées.

3.7. Dimensionnement de l'infrastructure et capacités de charge

L'architecture Ecluse intègre un mécanisme de **"Filtrage en entonnoir"** qui optimise drastiquement l'usage des ressources serveur. Le traitement des requêtes suit une logique de coût progressif :

1. **Niveau 1 (CPU) :** Les règles Regex et le modèle léger **GLiNER** (optimisé pour la détection d'entités nommées) filtrent 80% à 90% des flux courants ne contenant pas de données complexes.
2. **Niveau 2 (GPU) :** Le SLM local (ex: Mistral 7B) n'est sollicité que pour les cas d'ambiguïté sémantique ou de désanonymisation contextuelle complexe.

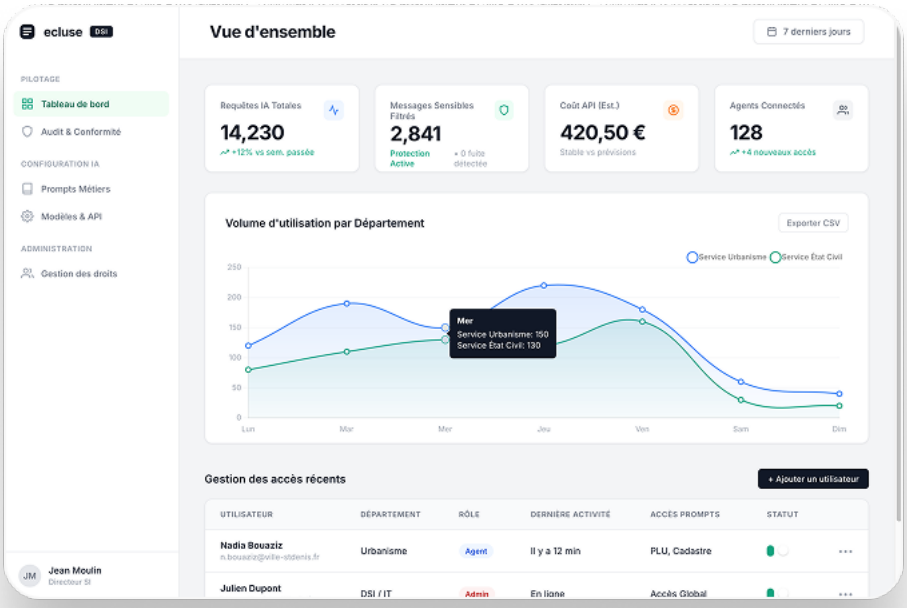
Le dimensionnement du serveur local conditionne la fluidité de l'expérience utilisateur et la capacité du système à absorber les pics de charge. Pour l'exécution d'un modèle SLM de type Mistral 7B, deux configurations sont proposées :

- **Option A : Modeste**
  - **Configuration :** GPU 16 Go VRAM (ex: NVIDIA RTX A4000 / 4060 Ti 16GB), 32 Go RAM, CPU 8 Cœurs.
  - **Capacité réelle** *:* Avec le filtrage actif, ce serveur peut supporter une file active de **30 à 50 agents travaillant simultanément**.

- **Option B : Haute disponibilité**
  - **Configuration :** GPU 24 Go VRAM Datacenter (ex: NVIDIA L4 ou A10), 64 Go RAM, CPU 16 Cœurs.
  - **Capacité réelle** *:* Cette configuration peut gérer un parc de **150 à 200 agents connectés en simultané**.

4. Administration, Contrôle des données et Audit

ecluse ADMIN
L'IA SOUVERAINE ET MAÎTRISEE



Cette couche transversale est essentielle pour le pilotage par la DSI et la conformité supervisée par le DPO.

4.1. Gouvernance des accès et sécurité

Gestion des rôles (RBAC) : Distinction stricte entre les profils "Administrateur" (accès aux logs, configs, quotas) et "Agent Utilisateur" (accès uniquement à l'interface de chat).

Sécurité des accès : Politique de mots de passe forts imposée et procédure de réinitialisation sécurisée

Révocation immédiate : Fonctionnalité de "Kill Switch" permettant de suspendre ou supprimer instantanément l'accès d'un agent (départ, mobilité) pour empêcher toute fuite de données post-contrat.

4.2. Outils de supervision de la pseudonymisation

Listes d'exclusion/inclusion : Possibilité d'ajouter manuellement des termes sensibles spécifiques à la collectivité (noms de projets secrets, acronymes) qui doivent être systématiquement masqués.

4.3. Audit, Journalisation et Exports (Compliance)

Ecluse permet d'exporter les preuves nécessaires aux contrôles internes et externes :

a) Journaux d'activité (Logs fonctionnels)

Suivi des usages : Qui utilise la solution, à quelle fréquence, sur quels types de fichiers (ex: PDF, Excel), sans révéler le contenu des fichiers.

Volume de tokens consommés.

b) Traçabilité des traitements (Audit Trails pour DPO)

Export des journaux techniques permettant de prouver la minimisation.

Possibilité (sur activation temporaire pour audit) de journaliser les paires [Prompt envoyé pseudonymisé / Réponse reçue]. Cela permet au DPO de vérifier a posteriori qu'aucune donnée identifiante (PII) n'a quitté l'enceinte de la collectivité.

c) Cycle de vie des données

Purge automatique : Configuration de la suppression automatique des documents et des vecteurs temporaires à définir par l'administrateur.

5. Conclusion : Une architecture prête pour la production

Pour un DSI ou un responsable applicatif, la solution Ecluse offre un équilibre pragmatique :

Sécurité : Les documents restent sous contrôle jamais indexés dans une base globale externe.

Souveraineté : Les traitements critiques (anonymisation) sont "on-premise".

Auditabilité : Le module d'administration fournit les métriques et les logs nécessaires au DPO.

Performance : L'usage du RAG permet d'analyser des documents longs et complexes (marchés publics, rapports) que les agents traitent quotidiennement.

La solution transforme l'usage non sécurisé de l'IA (Shadow AI) en un service managé, tracé et sécurisé, adapté aux exigences du secteur public.

6. Contact

Pour toutes informations complémentaires ou questions :

Laurent LEFORT

Fondateur et Directeur Technique (CTO) - Ecluse

Mail : laurent.lefort@ecluse.co

Téléphone : 0612602182

Site Internet : www.ecluse.co