EPIC 3 — Stack IA locale RGPD (POC contrôlé)

1. Objectif de l'EPIC

Déployer une **stack IA locale** en mode **POC contrôlé**, permettant : - de valider techniquement l'exécution locale de modèles IA, - de maîtriser totalement les flux de données sensibles, - de préparer l'industrialisation ultérieure (EPIC 6).

Cet EPIC est volontairement **limité, non industrialisé**, et sert de **laboratoire sécurisé**.

---

2. Positionnement dans la roadmap

- EPIC 1 : socle applicatif sécurisé (IAM, Gateway LLM)
- EPIC 2 : durcissement serveur & réseau
- **EPIC 3 : validation technique IA locale (POC contrôlé – ce document)**
- EPIC 4 : stockage IA & données utilisateur RGPD
- EPIC 5 : pipeline RGPD (droits des personnes)
- EPIC 6 : stack IA dockerisée RGPD-ready (industrialisation)

⚠️**EPIC 3 ne constitue ni une base de production, ni une implémentation RGPD complète.** Il s'agit d'une **phase de validation technique préalable**, destinée à identifier les contraintes réelles (latence, consommation, limites modèles) qui **alimenteront la conception des EPIC 4, 5 et 6**.

---

3. Périmètre couvert

Inclus

- Exécution locale de modèles IA
- Flux IA via Gateway LLM (définie EPIC 1)
- Tests fonctionnels et de performance
- Mesures de sécurité minimales

Exclus

- Orchestration avancée
- Scalabilité
- Haute disponibilité
- Multi-tenant complet
- CI/CD

---

4. Principes directeurs

1. **Aucune donnée réelle de production**
2. **Isolation maximale même en POC**
3. **Pas de réseau sortant non contrôlé**

4. **Reproductibilité minimale documentée**

---

5. Architecture cible (POC)

- Serveur local durci (issu EPIC 2)
- Modèle IA local (ex : Ollama ou équivalent)
- Appel via Gateway LLM
- Stockage temporaire chiffré

---

6. User Stories & Exigences

US-1 — Exécution d’un modèle IA local

Je veux exécuter un modèle IA sans dépendance externe

Exigences - Modèle local uniquement - Pas d’appel réseau externe - Journalisation minimale

---

US-2 — Intégration avec la Gateway LLM

Je veux que tous les appels passent par la Gateway

Exigences - Aucun appel direct au runtime IA - Application des règles de redaction - Journalisation safe

---

US-3 — Stockage temporaire sécurisé

Je veux stocker temporairement les entrées/sorties

Exigences - Données chiffrées - Durée de vie courte - Suppression automatique

---

US-4 — Tests de performance contrôlés

Je veux mesurer la faisabilité technique

Exigences - Mesures CPU/RAM - Latence observée (non contractuelle) - Rapport de faisabilité

US-5 — Documentation POC

Je veux documenter les enseignements

Exigences - Limites connues - Risques identifiés - Recommandations pour EPIC 6

---

7. Dépendances

- Dépend de : EPIC 1, EPIC 2
- Prérequis pour : EPIC 6

---

8. Livrables

- Modèle IA local fonctionnel
- Gateway connectée
- Rapport POC
- Décision Go / No-Go industrialisation

---

9. Risques & points de vigilance

- Confusion POC / production
- Dérive fonctionnelle
- Mauvaise interprétation des performances

---

10. Definition of Done (DoD)

- Aucun flux IA hors Gateway
- Aucune donnée réelle utilisée
- Documentation rédigée
- Décision prise pour EPIC 6

Fin EPIC 3