

Data Engineering pour l'IA

UNE IA QUI CONNAÎT VOS DONNÉES

Transformez vos données en avantage IA : construisez un assistant qui répond juste sur VOS documents (RAG). L'IA générique parle à tout le monde — la vôtre connaîtra votre entreprise.



DURÉE
7h (1 jour ou 2
demi-journées)



PUBLIC
Data engineers,
développeurs,
architectes



MODALITÉ
Présentiel /
Distanciel



FINANCEMENT
OPCO · Qualiopi

Objectifs de la formation

- ✓ Transformer les données de l'entreprise en avantage IA concret
- ✓ Maîtriser les vector databases et les pipelines d'embedding
- ✓ Concevoir un assistant qui répond juste sur VOS documents (RAG)
- ✓ Arbitrer fine-tuning vs RAG selon coûts et performance

Programme

PARTIE 1

Donner de la mémoire à l'IA (RAG)

1

- RAG par la valeur : chunking, embeddings, retrieval
- Vector databases : Pinecone, Weaviate, ChromaDB, Qdrant
- Fine-tuning vs RAG : critères de choix, coûts, performances

PARTIE 2

Pipeline & atelier

2

- Pipelines de données : ingestion, transformation, indexation
- Qualité des données : nettoyage, déduplication, enrichissement
- Atelier : pipeline RAG complet sur données d'entreprise



CE QUE VOS ÉQUIPES REPARTENT AVEC

un **pipeline RAG opérationnel** : une IA fiable qui répond juste sur les données de votre entreprise.

Fait pour vous si...

- ✓ Vous voulez une IA fiable sur vos propres données
- ✓ Vous êtes data engineer, dev ou architecte
- ✓ Vous voulez maîtriser RAG et vector DB

Pas pour vous si...

- ✗ Vous cherchez un usage IA généraliste
- ✗ Vous n'avez pas de bases en Python / données