

# Déployer une Intelligence Artificielle (IA) open source en local

1 jour (7 heures)

Perfectionnement

Présentiel ou classe à distance

à partir de

**875 € HT** par personne (TVA en vigueur) Repas inclus en présentiel

Contenu mis à jour le 25/09/2026. Document téléchargé le 02/10/2026.

## Objectifs

A l'issue de cette formation, vous serez capable de :

- Définir les exigences fonctionnelles et non-fonctionnelles de votre modèle, ainsi que les tests
- Comparer les différents modèles open source du marché selon les principales métriques
- Explorer les différentes alternatives pour entraîner votre modèle : from scratch, transfer learning, prompting, distillation, RLHF, mixture of experts, LoRA et PEFT
- Spécifier et chiffrer les alternatives architecturales locales capables d'héberger votre modèle
- Déployer et surveiller votre modèle en production, et optimiser vos inférences.

Les compétences visées par cette formation font l'objet d'une évaluation.

## À qui s'adresse cette formation ?

Data Scientists, Data Analysts, ingénieurs Data, ingénieurs MLOps, développeurs, architectes Data, cadres dirigeants, managers, chefs de projets, managers de produits IA.

### Prérequis

Avoir suivi la formation IA-INTRO "Intelligence Artificielle (IA) pour les spécialistes de l'IT - Techniques, cas d'usage et innovations", ou avoir des compétences équivalentes. Etre familier avec le langage Python.

# Programme

## 1 Pendant votre formation

### Déployer une Intelligence Artificielle (IA) open source en local

Jour 1 – Matin

#### Introduction

- Dernières avancées dans la production de modèles d'IA
- Planification d'applications IA : les grands jalons
- IA closed source vs IA open source : comparatif des avantages et inconvénients
- Ingénierie des coûts
- Modèles open source :
  - Polyvalents (Mistral/Mixtral, Llama, Falcon)
  - Légers (Phi-3, Gemma, TinyLlama)
  - Spécialisés (CodeLlama, Meditron)
- Exigences matérielles pour chaque type de modèle (mémoire, CPU/GPU/multi-GPU/cluster/cloud privé, réseautage)
- Ingénierie d'Intelligences Artificielles : des LLM aux modèles fondateurs et écosystèmes d'IA
- Ingénierie des datasets :
  - Traitement
  - Enrichissement
  - Augmentation
  - Synthèse des données
- Les trois couches de la stack IA

#### Exemples de travaux pratiques (à titre indicatif)

- *A partir d'une étude de cas : recueil des besoins, étude de faisabilité, estimation du budget et des ressources, sélection d'une solution, implémentation de l'application IA, tests*

Jour 1 – Après-midi

#### Les modèles fondateurs

- Qu'est-ce qu'un modèle fondateur, exemples
- Processus de création d'un modèle fondateur
- Données d'apprentissage
- Modélisation : architecture du modèle, taille du modèle
- Alternatives d'entraînement :
  - From scratch
  - Transfer learning

- Prompting
- Distillation
- RLHF
- Mixture of experts
- LoRA
- PEFT Affinage (fine tuning)
- Echantillonnage (stratégies, test time compute, sorties structurées)
- Infrastructures pour les LLM :
  - Ollama
  - LM Studio
  - llama.cpp
  - vLLM
  - Unsloth
  - Hugging Face
  - Hub Space
- Optimisation des LLM :
  - Quantization
  - LoRA
  - Fully sharded data parallel–quantized LoRA
- Défis et solutions dans la mise en production des LLM et LCM

#### *Exemples de travaux pratiques (à titre indicatif)*

- Réimplémentation du modèle open source Llama
- Amélioration du modèle avec la quantization et LoRA
- Déploiement, inférence, et surveillance

## Financement

Plusieurs dispositifs de financement sont accessibles via les OPCO (Opérateurs de Compétences), organismes agréés par le ministère du Travail dont le rôle est d'accompagner, collecter et gérer les contributions des entreprises au titre du financement de la formation professionnelle.

Pour plus d'information, une équipe de gestionnaires SKOLAE Formation spécialisée vous accompagne dans le choix de vos formations et la gestion administrative.

[Nous contacter pour en savoir plus →](#)

## Modalités pédagogiques

Formation délivrée en présentiel ou distanciel\* (blended-learning, e-learning, classe virtuelle, présentiel à distance).

Le formateur alterne entre méthode\*\* démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation).

Variables suivant les formations, les moyens pédagogiques mis en oeuvre sont :

- Ordinateurs Mac ou PC (sauf pour certains cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel)
- Environnements de formation installés sur les postes de travail ou en ligne
- Supports de cours et exercices

En cas de formation intra sur site externe à M2I, le client s'assure et s'engage également à avoir toutes les ressources matérielles pédagogiques nécessaires (équipements informatiques...) au bon déroulement de l'action de formation visée conformément aux prérequis indiqués dans le programme de formation communiqué.

\* nous consulter pour la faisabilité en distanciel

\*\* ratio variable selon le cours suivi

### Accessibilité aux personnes en situation de handicap

## Détails de la session

Le 23 novembre 2026 (1 jour / 7 heures)

Lyon

Le 23 novembre 2026 Lyon

- 23 novembre 2026 ..... 9h00 – 17h00