

Intelligence Artificielle (IA) – Apache Airflow

2 jours (14 heures)

Perfectionnement

Présentiel ou classe à distance

à partir de

1 720 € HT par personne (TVA en vigueur)

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 :

- Décrire l'architecture d'Airflow et le modèle d'orchestration par DAG
- Acquérir les bonnes pratiques de conception de pipelines de données
- Configurer et étendre Airflow avec des opérateurs, hooks et connecteurs
- Appréhender les différentes stratégies d'exécution et de scalabilité
- Mettre en oeuvre le scheduling avancé et les DAG dynamiques
- Travailler en sécurité avec Airflow en production.

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

À qui s'adresse cette formation ?

Architectes, chefs de projet, développeurs, équipes data engineering et DevOps.

Prérequis

Avoir des notions d'architecture applicative et de systèmes distribués. Des connaissances en SQL et en conteneurisation (Docker / Kubernetes) sont recommandées.

Afin de valider vos prérequis avant la formation, un questionnaire est disponible ; n'hésitez pas à nous contacter.

Programme

1 Pendant votre formation

Intelligence Artificielle (IA) – Apache Airflow

Comprendre l'architecture d'Airflow et le modèle d'orchestration par DAG

- Le projet Airflow : historique, fonctionnalités, cas d'usages
- Présentation de l'architecture : Webserver, Scheduler, Executor, Metastore, Workers
- Le modèle DAG : tâches, dépendances, opérateurs, sensors

Acquérir les bonnes pratiques de conception de pipelines de données

- Ecriture de DAG en Python : structuration et bonnes pratiques

Exemples de travaux pratiques (à titre indicatif)

- Création d'un DAG ETL complet avec branchements et gestion d'erreurs
- Utilisation des XComs, Variables et Connections – templating Jinja

Savoir configurer et étendre Airflow avec des opérateurs, hooks et connecteurs

- Les opérateurs natifs : BashOperator, PythonOperator, SQLOperator, DockerOperator...
- Développement de custom operators et hooks : intégration avec des sources externes

Exemple de travaux pratiques (à titre indicatif)

- Connexion à des services tiers (S3, GCS, bases de données, API REST)

Appréhender les différentes stratégies d'exécution et de scalabilité

- Les Executors : Sequential, Local, Celery, Kubernetes
- Configuration d'un environnement multi-workers avec CeleryExecutor

Exemple de travaux pratiques (à titre indicatif)

- Déploiement d'Airflow en mode distribué avec Docker Compose

Mettre en oeuvre le scheduling avancé et les DAG dynamiques

- Scheduling : cron, catchup, backfill, SLA, time zones
- DAGs dynamiques et TaskFlow API. Deferrable Operators
- Data-Aware Scheduling : le concept de Datasets et déclenchement par données

Exemple de travaux pratiques (à titre indicatif)

- Mise en oeuvre d'un pipeline dynamique avec dépendances inter-DAG

Travailler en sécurité avec Airflow en production

- Authentification et contrôle d'accès : RBAC, LDAP / OAuth
- Chiffrement des connexions et des variables sensibles (Fernet, Secrets Backends)

- Isolation des ressources : Pools, Priorities

Exemple de travaux pratiques (à titre indicatif)

- Sécurisation d'un environnement Airflow multi-équipes

Exploitation et observabilité

- Airflow CLI et API REST
- Gestion des logs et intégration avec des systèmes centralisés (ELK, Loki)
- Monitoring avec métriques StatsD / Prometheus / Grafana – alerting
- Détection et gestion des pannes : stratégies de retry et dead letter

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

Du 1 juillet 2027 au 11 juillet 2027 (2 jours / 14 heures)

À distance

Du 1 juillet 2027 au 11 juillet 2027 **À distance**

- 1 juillet 2027 9h00 – 16h30
- 2 juillet 2027 9h00 – 16h30
- 3 juillet 2027 9h00 – 16h30
- 4 juillet 2027 9h00 – 16h30
- 5 juillet 2027 9h00 – 16h30
- 6 juillet 2027 9h00 – 16h30
- 7 juillet 2027 9h00 – 16h30
- 8 juillet 2027 9h00 – 16h30
- 9 juillet 2027 9h00 – 16h30
- 10 juillet 2027 9h00 – 16h30
- 11 juillet 2027 9h00 – 16h30