

# DataOps – Concevoir et automatiser des pipelines résilients et scalables

5 jours (35 heures)

Perfectionnement

Présentiel ou classe à distance

à partir de

**3 925 € HT** par personne (TVA en vigueur) **Repas inclus en présentiel**

Contenu mis à jour le 25/09/2026. Document téléchargé le 07/10/2026.

## Objectifs

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

- Définir une stratégie Data ambitieuse et réaliste
- Evaluer les priorités Data de l'organisation en fonction de l'état des lieux et de la vision
- Identifier et pratiquer les meilleures technologies disponibles et leur articulation / complémentarité (MLflow, Airflow, DVC, Delta Lake, Evidently AI, Optuna...)
- Construire des plateformes DataOps pour minimiser le Time To Value.

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

## À qui s'adresse cette formation ?

Ingénieurs DevOps, ingénieurs Data, Data Scientists, Data Analysts, développeurs, architectes Data, chefs de projets, managers, cadres dirigeants, consultants.

### Prérequis

Avoir des connaissances en bases de données et/ou analytiques.

# Programme

## 1 Pendant votre formation

### DataOps – Concevoir et automatiser des pipelines résilients et scalables

Jour 1

#### Fondamentaux

- Qu'est-ce qu'une donnée ?
- Evolution des analytiques : des Data Warehouses aux Lakehouses et Data Mesh
- Stratégie des données, valeur des données, sémantique des données, BPM (Business Process Management)
- La stratégie des données, et différence avec les stratégies IT et de gouvernance de données
- Définir la vision
- Déterminer et aligner les objectifs des pipelines avec la vision
- Cartographie de la chaîne de valeur (Value Stream Mapping)
- Apports des analytiques et du Cloud
- Qu'est-ce qu'une organisation Data Driven ?
- Cycle de vie de la Data, gouvernance et qualité des données
- Lean, Agilité et DevOps appliqués aux données
- Apports du DataOps
- DevOps vs DataOps vs MLOps
- Analytiques et opérations : briser le mur de la confusion
- Constituer et former les Data Teams
- Introduction à Microsoft Azure
- Etudes de cas DataOps

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

- Contrôle de version des données avec DVC et Git

Jour 2

#### Scoping et définition des éléments des pipelines DataOps

- Processus de développement Agile DataOps : idéation, "inception", R&D, transition / production
- Conception de pipelines
- Les trois variantes d'implémentation de pipelines
- MLflow: la plateforme indispensable pour le tracking des expériences et la gestion des modèles
- Exemples d'implémentations avec MLflow Projects et MLflow Recipes
- Importance du versioning modèle-dataset pour la traçabilité

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

- Création d'un pipeline MLflow intégrant Git et DVC

- Déploiement du meilleur modèle suite à Grid Search / Randomized Search / Optimisation bayésienne
- Détection du data drift et du feature drift avec MLflow et Evidently

Jour 3

## Les technologies DataOps

- Ecosystème des technologies DataOps
- Build vs Buy
- Choisir les bons outils : cartes de Wardley, radar technologique
- Intégration, préparation, traitement et management des données
- Reproductibilité, déploiement, orchestration, et surveillance
- Infrastructure de calcul et moteur d'exécution de requêtes
- Stockage, plateformes et outils d'analytics
- Faire évoluer la pile technologique
- Concevoir des pipelines robustes avec Airflow
- Planifier des workflows complexes avec Airflow

*Exemples de travaux pratiques (à titre indicatif)*

- Amélioration du pipeline précédent avec Airflow
- Planification de l'exécution du pipeline en fonction d'évènements

Jour 4

## Produire la valeur

- Planifier les itérations
- Extraire la valeur des données
- Surveiller la qualité avec la MSP (Maîtrise Statistique des Procédés)
- Mesurer les bénéfices
- Animer les rétrospectives et assurer l'amélioration continue

*Exemples de travaux pratiques (à titre indicatif)*

- Optimiser l'empreinte des modèles (CPU, RAM, stockage) avec MLflow et Optuna
- Serving de plusieurs modèles via un seul point de terminaison avec MLflow et PyFunc

Jour 5

## Le DataOps Factory

- Importance de la gouvernance des données
- Culture DataOps et conduite du changement
- Centraliser les plateformes : exemple de Delta Lake
- Tout automatiser
- L'organisation Self-Service

*Exemples de travaux pratiques (à titre indicatif)*

- *IA responsable : suivre l'interprétabilité et l'explicabilité des modèles avec MLflow, SHAP, LIME et interpretML*
- *Mise en place d'une plateforme de données avec Spark et Delta Lake*

## 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 26 octobre 2026 au 30 octobre 2026 (5 jours / 35 heures)

Tours

Du 26 octobre 2026 au 30 octobre 2026

Tours

- 26 octobre 2026 ..... 9h00 – 15h00
- 27 octobre 2026 ..... 9h00 – 15h00
- 28 octobre 2026 ..... 9h00 – 15h00
- 29 octobre 2026 ..... 9h00 – 15h00
- 30 octobre 2026 ..... 9h00 – 15h00