# Apache Airflow — Orchestration de pipelines de données

*Lien :*
*https://innov-systems.com/formation/apache-airflow-orchestration-de-pipelines-de-donnees*

DURÉE
**5 jours (35h)**

RÉFÉRENCE
**BDI165**

CATÉGORIE
**Data Science et Data Engineering**

Innov Systems

## OBJECTIFS DE LA FORMATION

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

- Comprendre les principes de l’orchestration de workflows avec Airflow
- Concevoir et développer des DAGs complexes
- Automatiser des pipelines de données ETL/ELT
- Intégrer Airflow avec des outils data modernes
- Gérer la planification et l’exécution des tâches
- Assurer le monitoring des workflows
- Déployer Airflow en environnement scalable
- Améliorer la fiabilité des pipelines de données
- Gérer les dépendances entre systèmes distribués
- Optimiser les performances des workflows
- Appliquer les bonnes pratiques de développement Airflow
- Étudier des cas concrets d’orchestration de données

## POUR QUI ?

- Data engineers
- DevOps engineers
- ML engineers
- Data architects
- Backend developers
- SRE (Site Reliability Engineers)
- Platform engineers
- Cloud engineers
- BI developers
- Administrateurs systèmes
- Consultants data engineering
- Toute personne impliquée dans l’orchestration de pipelines de données

Innov Systems

## Programme détaillé

### 1 / Introduction à Apache Airflow

- Concept d’orchestration de workflows
- Positionnement dans les architectures data modernes
- Cas d’usage en data engineering et automation

### 2 / Architecture d’Apache Airflow

- Composants : Scheduler, Webserver, Workers
- DAG (Directed Acyclic Graph) et exécution des tâches
- Backend de métadonnées et stockage

### 3 / Création de DAGs

- Structure d’un DAG Airflow
- Définition des tâches et dépendances
- Paramétrage et scheduling

### 4 / Operators et tâches Airflow

- PythonOperator, BashOperator et opérateurs custom
- Sensors et triggers
- Gestion des tâches complexes

### 5 / Orchestration de pipelines de données

- ETL/ELT avec Airflow
- Coordination de systèmes distribués
- Intégration avec Spark, dbt et APIs

Innov Systems

### 6 / Planification et scheduling avancé

- Cron scheduling et triggers dynamiques
- Backfilling des données
- Gestion des dépendances temporelles

### 7 / Monitoring et observabilité

- Interface web Airflow
- Logs et tracking des exécutions
- Gestion des erreurs et retries

### 8 / Scalabilité et déploiement

- Mode Celery et Kubernetes Executor
- Scaling horizontal des workers
- Déploiement en cloud

### 9 / Sécurité et bonnes pratiques

- Gestion des connexions et secrets
- Isolation des environnements
- Bonnes pratiques de développement DAG

### 10 / Cas d’usage et industrialisation

- Pipelines data en production
- Orchestration de workflows ML et BI
- Études de cas en entreprise

Innov Systems

## Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

## Prochaines dates programmées

| | |
|---|---|
| 16 au 20 Nov. 2026 | Distanciel |
| 12 au 18 Jan. 2027 | Présentiel - Casablanca |
| 08 au 12 Mars 2027 | Présentiel - Casablanca |
| 03 au 07 Mai 2027 | Présentiel - Casablanca |
| 28 Juin au 02 Juil. 2027 | Présentiel - Casablanca |

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

## Réservation & Renseignements

**Téléphone :** +212 522 247 210
**Email :** contact@innov-systems.com
**Web :** https//www.innov-systems.com


Innov Systems — Tous droits réservés