



## ÍNDICE GENERAL DEL CURSO – Apache Airflow

### M0 · Orientación y Objetivos

- **M0.1:** ¿Qué resuelve Airflow?
  - **M0.2:** Arquitectura de alto nivel
  - **M0.3:** Flujo de ejecución y ciclo de vida de un DAG
- 

### M1 · Instalación y Primer Entorno

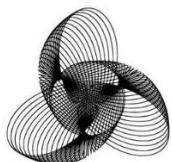
- **M1.1:** Requisitos y versiones
  - **M1.2:** Instalación local
  - **M1.3:** Quickstart con Docker Compose
  - **M1.4:** Variables de entorno y *airflow.cfg*
  - **M1.5:** Arranque y verificación básica
- 

### M2 · Conceptos Nucleares

- **M2.1:** DAGs, Tasks y dependencias
  - **M2.2:** Operadores, Sensores, Hooks y Providers
  - **M2.3:** Programación vs. orquestación
  - **M2.4:** Paralelismo, Pools, Prioridades y Concurrencia
  - **M2.5:** Timezones y fechas lógicas
- 

### M3 · Web UI y Operación Diaria

- **M3.1:** Vistas principales
- **M3.2:** Pausar/activar DAGs, correr manual, limpiar, reintentos



# AEDIA

Asociación Española para la Difusión de la Inteligencia Artificial

- **M3.3:** Logs, XCom viewer y Dataset details
  - **M3.4:** RBAC básico
  - **M3.5:** Documentación embebida en DAGs
- 

## M4 · Programación de DAGs

- **M4.1:** *schedule* y Timetables
  - **M4.2:** *catchup*, *backfills* y *data intervals*
  - **M4.3:** Data-aware scheduling con *Datasets*
  - **M4.4:** SLA y alertas por retrasos
- 

## M5 · TaskFlow API Moderna

- **M5.1:** Decoradores @dag y @task
  - **M5.2:** expand() y Dynamic Task Mapping
  - **M5.3:** TaskGroup vs. SubDAG
  - **M5.4:** Parámetros y validación
- 

## M6 · Plantillas y Ninja

- **M6.1:** Contexto de plantillas: macros
  - **M6.2:** Campos *templated* en Operadores
  - **M6.3:** Buenas prácticas de templating
- 

## M7 · XComs y Paso de Datos

- **M7.1:** Lectura/escritura y TaskFlow
  - **M7.2:** Backends de XCom
  - **M7.3:** Límites y patrones seguros
-

## M8 · Operadores, Sensores y Hooks

- **M8.1:** Operadores Bash, Python, Branch
  - **M8.2:** File/HTTP/S3/GCS/FTP
  - **M8.3:** BaseHook y conexiones
  - **M8.4:** Providers: Amazon, Google, Snowflake...
- 

## M9 · Deferrable Operators y Triggerer

- **M9.1:** Arquitectura Triggerer
  - **M9.2:** Migración desde sensores *poke*
  - **M9.3:** Patrones de espera y ahorro
- 

## M10 · Argumentos y Configuración Avanzada

- **M10.1:** *default\_args* vs. *@task*
  - **M10.2:** *depends\_on\_past* y *max\_active\_runs*
  - **M10.3:** Pools, Queues y Executors
  - **M10.4:** Prioridad y *weight\_rule*
- 

## M11 · Conexiones, Variables y Secret Backends

- **M11.1:** Gestión desde UI/CLI
  - **M11.2:** Backends: Vault, AWS, GCP
  - **M11.3:** Versionado y seguridad
- 

## M12 · API, CLI y Automatización Operativa

- **M12.1:** REST API

- **M12.2:** CLI esencial
  - **M12.3:** Automatización con CLI/API
- 

## **M13 · Testing y Calidad**

- **M13.1:** Linting / pre-commit
  - **M13.2:** Unit tests de DAGs
  - **M13.3:** airflow dags test
  - **M13.4:** Validaciones y rendimiento
- 

## **M14 · Observabilidad, Alertas y Rendimiento**

- **M14.1:** Logs remotos
  - **M14.2:** Métricas y Grafana
  - **M14.3:** Alerting y notificaciones
  - **M14.4:** Optimización del parser
- 

## **M15 · Despliegue y Producción**

- **M15.1:** Docker Compose *production-like*
  - **M15.2:** Kubernetes + Helm
  - **M15.3:** Escalado de workers
  - **M15.4:** Backends de metadatos
  - **M15.5:** Remote logging y artefactos
- 

## **M16 · Integraciones de Datos**

- **M16.1:** Ingesta batch
  - **M16.2:** Spark / EMR / Dataproc
-

- **M16.3:** Databricks
  - **M16.4:** dbt + Airflow
- 

## **M17 · Seguridad y Gobierno**

- **M17.1:** RBAC avanzado
  - **M17.2:** Auditoría y trazabilidad
  - **M17.3:** Multi-tenant
- 

## **M18 · Migraciones y Upgrades**

- **M18.1:** Buenas prácticas de upgrade
  - **M18.2:** Cambios frecuentes
  - **M18.3:** Estrategia de rollback
- 

## **M19 · Patrones y Anti-Patrones**

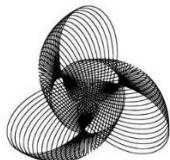
- **M19.1:** Fan-out / fan-in
  - **M19.2:** Idempotencia y *retries*
  - **M19.3:** Antipatrones comunes
- 

## **M20 · CI/CD para DAGs**

- **M20.1:** Repos y estructura
  - **M20.2:** GitHub / GitLab CI
  - **M20.3:** Promoción y Canary DAGs
- 

## **M21 · Operación Avanzada en Kubernetes**

- **M21.1:** KEDA y autoescalado
-



# AEDIA

Asociación Española para la Difusión de la Inteligencia Artificial

- **M21.2:** Pod templates
  - **M21.3:** Montaje de secrets
- 

## M22 · Plataformas Gestionadas

- **M22.1:** Astronomer
  - **M22.2:** AWS MWAA y GCP Composer
  - **M22.3:** Costes y gobernanza
- 

## M23 · Troubleshooting de Producción

- **M23.1:** Cuellos de botella
  - **M23.2:** Deadlocks y *stuck tasks*
  - **M23.3:** Estrategias de mitigación
- 

## M24 · Preparación Certificación Astronomer

- **M24.1:** Blueprint de objetivos
  - **M24.2:** Preguntas tipo
  - **M24.3:** Checklist y simulacros
- 

## M25 · Proyecto Integrador

- **M25.1:** Pipeline E2E
  - **M25.2:** Despliegue en Kubernetes/Helm
  - **M25.3:** Documentación y presentación
-