

Mis à jour le 26/08/2026

S'inscrire

Formation Apache Gravitino

2 jours (14 heures)

Présentation

Apache Gravitino est une plateforme de metadata lake fédérée qui centralise la gouvernance, l'accès aux données et la gestion des catalogues pour les environnements data modernes.

Cette formation Apache Gravitino vous permet de comprendre le rôle d'une couche de métadonnées unifiée dans une architecture data lakehouse, multi-sources et multi-moteurs. Vous apprendrez à structurer une approche cohérente de la gouvernance, de la découverte et du contrôle d'accès.

Vous verrez comment Gravitino s'intègre à des sources relationnelles, des stockages de fichiers, des flux de messages et des environnements analytiques, tout en exposant des API et des mécanismes de gestion adaptés aux usages d'entreprise. L'objectif est de savoir organiser les métadonnées de manière lisible, robuste et exploitable par les équipes data.

La formation aborde également les notions de catalogue, de métadonnées unifiées, de sécurité centralisée et d'interopérabilité avec les moteurs de traitement. Vous gagnerez en autonomie pour préparer une base technique adaptée aux besoins de production.

À l'issue de ce parcours, vous serez capable de concevoir une organisation Gravitino claire, de piloter les opérations courantes et de poser les fondations d'une architecture data plus gouvernée et plus scalable.

Comme toutes nos formations, celle-ci vous présentera **la dernière version stable** de la technologie et ses nouveautés.

Objectifs

- Comprendre l'architecture et le positionnement d'Apache Gravitino.

- Maîtriser la logique de catalogue et de métadonnées fédérées.
- Configurer des sources et organiser une gouvernance centralisée.
- Exploiter les API et les mécanismes d'administration de base.
- Relier Gravitino aux usages data lakehouse et multi-moteurs.
- Appliquer les bonnes pratiques de sécurité, d'exploitation et de structuration.

Public visé

- Data Engineer
- Architecte Data
- Administrateur plateforme data
- Ingénieur gouvernance des données
- Data Platform Engineer

Pré-requis

- Connaissances de base en architecture data et en gouvernance.
- Pratique des environnements lakehouse, SQL ou catalogue de données.
- Compréhension des notions de stockage, de métadonnées et de contrôle d'accès.
- Expérience préalable avec un écosystème Java ou des outils data d'entreprise.

Pré-requis techniques

- Un ordinateur récent avec suffisamment de mémoire pour exécuter les outils de formation.
- Une connexion Wi-Fi ou Ethernet stable pour les manipulations en ligne.
- Java 17 installé et correctement configuré dans le PATH ou via JAVA_HOME.
- Un navigateur web moderne et, si besoin, un client Docker pour les démonstrations.
- Un éditeur de code ou IDE adapté, ainsi qu'un micro-casque pour la classe virtuelle.

Programme de notre formation Apache Gravitino

[Jour 1 - Matin]

Fondations de Gravitino

- Comprendre le rôle d'Apache Gravitino dans une architecture de données moderne.
- Identifier les enjeux de métadonnées unifiées, de gouvernance et de traçabilité.
- Découvrir les concepts de metadata lake, de catalogue et de source de vérité.
- Positionner Gravitino face aux usages lakehouse et multi-moteurs.
- Lire l'architecture fonctionnelle, les couches de service et les principaux objets métiers.
- Atelier pratique : cartographier un cas d'usage data et identifier les métadonnées à centraliser.

[Jour 1 - Après-midi]

Installation et prise en main

- Préparer l'environnement local et vérifier les dépendances nécessaires.
- Installer et démarrer les composants de base dans un contexte de démonstration.
- Explorer l'interface d'administration, la CLI et les premiers points d'accès.
- Comprendre la logique des API REST et des interactions serveur client.
- Lire la structure d'un metalake et les objets de premier niveau.
- Atelier pratique : déployer une instance de test et exécuter les premières commandes d'inventaire.

[Jour 2 - Matin]

Gestion des catalogues

- Créer et administrer des catalogues selon les types de sources supportées.
- Comprendre les différences entre catalogues relationnels, fichiers, messagerie et modèles.
- Paramétrer les propriétés principales, les emplacements et les métadonnées associées.
- Organiser la découverte, la consultation et la mise à jour des objets gérés.
- Appliquer des règles de structuration pour garder une nomenclature cohérente.
- Atelier pratique : configurer plusieurs catalogues et vérifier leur exposition dans Gravitino.

[Jour 2 - Après-midi]

Gouvernance, sécurité et usages avancés

- Mettre en place une gouvernance centralisée autour des métadonnées et des accès.
- Appréhender les mécanismes de sécurité, de contrôle et d'audit.
- Relier Gravitino aux besoins de data platform et d'intégration avec les moteurs de calcul.
- Comprendre les extensions possibles pour les policies, les services associés et les évolutions de l'écosystème.
- Préparer une démarche d'exploitation, de maintenance et de montée en charge.
- Atelier pratique : définir une stratégie de gouvernance et simuler un scénario de mise en production.

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.