

Mis à jour le 21/07/2025

S'inscrire

Formation Matillion ETL

2 jours (14 heures)

Présentation

Maîtrisez Matillion ETL dans sa globalité grâce à cette formation complète, structurée et résolument orientée pratique. De la prise en main à l'automatisation des pipelines, vous apprendrez à concevoir, orchestrer et optimiser des flux de données modernes, dans un environnement cloud-first, compatible avec Snowflake, Redshift ou BigQuery.

Vous commencerez par découvrir l'architecture de Matillion, les différences entre jobs d'orchestration et de transformation, ainsi que la configuration initiale des projets et des connexions sources. Objectif : poser des bases solides pour le développement de pipelines fiables.

Vous apprendrez ensuite à concevoir des workflows d'intégration complets, en manipulant les composants natifs, les connecteurs cloud, les API externes et les scripts personnalisés, tout en appliquant les bonnes pratiques de performance.

La formation abordera aussi la gestion avancée des variables, l'automatisation via planification ou triggers, ainsi que le monitoring détaillé des exécutions et erreurs. Vous découvrirez comment industrialiser vos jobs avec Git et orchestrateurs externes.

Comme pour toutes nos formations, celle-ci vous sera présentée avec les toutes dernières actualisations de [Matillion ETL](#).

Objectifs

- Comprendre l'architecture de Matillion ETL
- Savoir déployer, configurer et connecter Matillion ETL à des sources et cibles variées
- Maîtriser les composants de transformation de Matillion
- Être capable d'automatiser, superviser et auditer les traitements ETL

Public visé

- Data Engineers
- Data analytics

Pré-requis

- Maîtrise des fondamentaux SQL
- notions générales sur les architectures ETL / ELT

Programme de la formation Matillion ETL

Introduction à Matillion ETL

- Objectifs et cas d'usage
- Architecture et principes de fonctionnement
- Matillion ETL vs Matillion Data Loader
- Cloud supporté : AWS, GCP, Azure
- Présentation des menus
- Zones de développement

Configuration de l'environnement

- Déploiement de Matillion
- Lancement via AWS Marketplace, GCP Marketplace, ou Azure
- Prérequis réseau et sécurité
- Connexions aux bases de données
- Création de Data Source
- Authentification, clés et autorisations
- Paramètres globaux et projets
- Configuration de projet
- Variables globales vs environnementales

Création de jobs Orchestration

- Comprendre le pipeline d'orchestration
- Orchestration vs Transformation
- Usage typique d'un job d'orchestration
- Composants principaux
- Python Script, Bash Script, If Condition, Iterator
- Run Orchestration / Run Transformation
- Gestion des dépendances
- Chaînage logique des composants

- Gestion des erreurs

Création de jobs Transformation

- Concepts fondamentaux de transformation
- Types de composants : Input, Join, Filter, Calculator
- Chargement, nettoyage, enrichissement
- Construction de pipelines
- Design visuel de transformations
- Utilisation de plusieurs sources
- Optimisation des performances
- Partitionnement
- Pushdown SQL
- Réduction des étapes inutiles

Connecteurs et intégrations

- Connecteurs natifs
- API et services cloud : Salesforce, S3, REST API
- Fichiers plats : CSV, Excel, JSON
- Utilisation des composants "Query"
- Redshift Query, Snowflake Query, BigQuery Query
- Web Services et API externes
- Appels REST avec OAuth 2.0
- Traitement JSON/XML en sortie

Gestion des données et qualité

- Data Profiling
- Statistiques descriptives avec le composant "Data Quality"
- Prévisualisation de données
- Gestion des erreurs
- Filtres de données invalides
- Logs, notifications, Retry
- Bonnes pratiques de transformation
- Validation de schéma
- Documentation inline et métadonnées

Variables, paramètres et réutilisabilité

- Utilisation des variables
- Variables de projet vs d'environnement
- Variables dans les scripts et composants
- Jobs paramétrables
- Composants Paramètres
- Job Template et modularité
- Réutilisation
- Common jobs

- Référence à des transformations existantes

Planification et automatisation

- Triggers et planification
- Matillion Scheduler
- Intégration avec Airflow, cron externe ou Cloud Scheduler
- Gestion des événements
- Déclenchement conditionnel
- Notifications
- Orchestration complexe
- Boucles, splits, parallélisation

Monitoring, logs et audit

- Console de monitoring Matillion
- Logs d'exécution et d'erreur
- Historique des jobs
- Intégration avec des outils tiers
- CloudWatch (AWS), Stackdriver (GCP), Azure Monitor
- Audit et traçabilité
- Versioning des jobs
- Suivi des modifications

Bonnes pratiques et cas d'usage avancés

- Bonnes pratiques DevOps
- Git integration
- CI/CD avec Matillion et Terraform
- Sécurité
- Gestion des accès par rôle
- Cryptage et gestion des credentials
- Études de cas
- Cas Snowflake : ingestion depuis S3, transformation, export
- Cas BigQuery : nettoyage de logs applicatifs
- Cas multi-sources : fusion et enrichissement de données

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.