

Talend Open Studio, mettre en œuvre l'intégration de données

Cours Pratique de 3 jours - 21h

Réf : TOT - Prix 2024 : 2 230€ HT

Talend Open Studio (TOS) est une application libre de type ETL permettant d'effectuer des synchronisations massives d'informations d'une base de données vers une autre. Cette formation vous apprendra à concevoir, modéliser et développer des "jobs" TOS afin de faire face aux différentes problématiques.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Concevoir et développer des jobs dans l'application ETL Talend

Optimiser les jobs développés par l'utilisation de contextes et jeux de données

Réaliser des transformations plus complexes en utilisant variables, expressions et jointures

Exécuter et déboguer un job, tracer les statistiques d'exécution

MÉTHODES PÉDAGOGIQUES

Apports théoriques illustrés par des exemples de mise en œuvre puis consolidés par la réalisation de mini projets Talend.

TRAVAUX PRATIQUES

Suite de mini projets permettant de concevoir des jobs Talend de plus en plus élaborés.

LE PROGRAMME

dernière mise à jour : 04/2023

1) Présentation

- L'intégration de données. Les solutions ETL.
- TOS : installation, préférences utilisateur. Documentation de référence.
- Philosophie du produit. Conception d'un job.

Travaux pratiques : Prise en main de l'outil. Présentation du cahier des charges du projet fil rouge.

2) Modéliser le besoin, concevoir un premier job

- Business Modeler. Job Designer.
- Principales connexions. Composants CSV et XML.
- Composants de transformations simples.
- Visualiser du code généré, exécuter un job.

Exercice : Développement d'un job assurant le tri d'une source CSV, le filtrage de données et le stockage du résultat dans un fichier XML.

3) Optimiser le développement des jobs

- Configurer des connexions réutilisables à l'aide des métadonnées.
- Mettre à jour les métadonnées et les propager sur les jobs, importer/exporter des métadonnées.
- Paramétrer les jobs par des contextes.
- Externaliser les variables de contexte dans des fichiers ".properties" et ".ini".
- Créer et gérer ses propres variables.
- Générer des jeux de données pour des tests.

Exercice : Refactorer un job en utilisant métadonnées et contextes. Générer un jeu de données de test pour ce job.

PARTICIPANTS

Développeurs, chefs de projet, consultants en informatique décisionnelle, administrateurs de base de données.

PRÉREQUIS

Bonne connaissance des SGBDR et du SQL. La connaissance du langage Java est un plus.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

4) Travailler avec des bases de données

- Bases de données supportées et principaux composants.
- Paramétrages des opérations sur les tables.
- Métadonnées et contexte de connexion à un schéma de base de données.
- Partage de connexions et gestion des transactions.
- Créer des requêtes à l'aide de SQLBuilder.

Exercice : Lecture et mise à jour d'un référentiel de données hébergé sur un serveur MySQL.

5) Travailler sur des données multisources

- Présentation du composant tMap.
- Configuration des flux d'entrée, créations de jointures.
- Réaliser des transformations en utilisant variables, expressions et jointures.
- Qualifier les données à l'aide de filtres.
- Générer des sorties multiples, gérer les rejets.

Exercice : Consolidation de données multisources et génération d'un entrepôt.

6) Compléments

- Décomposer un job en sous-jobs, utilisation du tRunJob. Lancer les jobs en ligne de commande. Exécution périodique.
- Débogage d'un job, tracer les statistiques d'exécution.
- Reporting tJasperOutput.

Exercice : Génération d'un rapport Jasper Report à partir d'un entrepôt.

LES DATES

CLASSE À DISTANCE

2024 : 19 juin, 23 sept., 13 nov.

PARIS

2024 : 12 juin, 16 sept., 06 nov.

LYON

2024 : 23 sept., 13 nov.

AIX-EN-PROVENCE

2024 : 23 sept., 13 nov.

BORDEAUX

2024 : 23 sept., 13 nov.

LILLE

2024 : 23 sept., 13 nov.

NANTES

2024 : 23 sept., 13 nov.

SOPHIA-ANTIPOLIS

2024 : 23 sept., 13 nov.

STRASBOURG

2024 : 23 sept., 13 nov.

TOULOUSE

2024 : 23 sept., 13 nov.