

# Talend Open Studio Data Integration : Initiation + Approfondissement

## INFORMATIONS GÉNÉRALES

**Type de formation :** Formation continue

**Éligible au CPF :** Non

**Domaine :** Décisionnel, collaboratif

**Action collective :** Non

**Filière :** Stratégie décisionnelle

**Rubrique :** Modélisation & conception

**Code de formation :** RDBN409

## € Tarifs

**Prix public :** 2995 €

### Tarif & financement :

Nous vous accompagnons pour trouver la meilleure solution de financement parmi les suivantes :

**Le plan de développement des compétences de votre entreprise :** rapprochez-vous de votre service RH.

**Le dispositif FNE-Formation.**

**L'OPCO** (opérateurs de compétences) de votre entreprise.

**France Travail:** sous réserve de l'acceptation de votre dossier par votre conseiller Pôle Emploi.

**CPF -MonCompteFormation**

Contactez nous pour plus d'information : [contact@aston-institut.com](mailto:contact@aston-institut.com)

## PRÉSENTATION

### Objectifs & compétences

A l'issue de la formation, le stagiaire sera capable :

- Comprendre le processus d'intégration de données
- Maîtriser l'outil Talend Open Studio et la bibliothèque de composants

### Public visé

Administrateurs systèmes/BDD, Chefs de projets, Développeurs

### Pré-requis

Connaissance de Java

## 📍 Lieux & Horaires

**Durée :** 35 heures

**Délai d'accès :** Jusqu'à 8 jours avant le début de la formation, sous condition d'un dossier d'inscription complet

## PROGRAMME

### Introduction

Problématique d'intégration de données

Intégration Opérationnelle et Décisionnelle

Outils ETL (Extract Transform Load) disponibles

Talend Open Studio : présentation, licences, lien avec Eclipse

Support disponible : documentation, communauté, ...

**Atelier :** Installation et présentation de l'environnement Talend Open Studio

### Découvrir l'interface

Les perspectives et les vues Talend

Utilisation du Business Modeler

Les jobs : grille, palette de composants

Les flux principaux (main)

La notion de schéma : types des colonnes

Premiers composants : entrée, sortie, génération de flux

**Atelier pratique :** création et exécution d'un job simple, génération de contenu dans un fichier

### Les principaux composants de Talend

Composants d'entrée (tFileInput, tMySQLInput, ...)

Génération de contenu (tRowGenerator, tFixedFlowInput, ...)

Sorties et affichage (tLogRow, tFileOutput, tMySQLOutput, ...)

Logique et chaînage des composants

Suivi pas à pas des traces de l'exécution du job

Gestion des rejets

Transformations simples (filtre, tri, remplacement, jointure)

**Atelier pratique :** génération de contenu dans une base de données, extraction et traitement des données pour affichage en console

### Garantir la pérennité des jobs

Versions et statuts des jobs

## 📅 Prochaines sessions

Consultez-nous pour les prochaines sessions.

Les métadonnées : propriétés et schémas.

La découverte de schéma pour les fichiers CSV

Récupération des schémas de bases de données

Mises à jour des jobs depuis les métadonnées

**Atelier pratique** : insérer les sources et cibles de données dans les métadonnées du référentiel

### Variables et contextes

Utiliser des variables dans un job (tSetGlobalVar)

Mise en place d'un contexte local au job

Groupes de contexte dans le référentiel

Les métadonnées comme contextes

**Atelier pratique** : utiliser une variable dans un job, contextualiser les paramètres d'une métadonnée, créer un contexte de test, et de production

### Le mappage de données : tMap

Le composant ETL par excellence

Gérer les entrées : flux principal, secondaires, nommage des flux

Gérer les sorties : l'éditeur intégré de la tMap

Filtrer les données : en entrée, sortie

Sorties conjointes, capture des rejets de sortie, de jointure interne

L'éditeur d'expression

Les performances

**Atelier pratique** : extraction multi-sources, multi-destinations avec une tMap, sur un exemple complet

### Orchestration des jobs, gestion d'erreurs

Exécutions parallèles de sous-jobs

Déclencheurs Talend (triggers)

Enchaînement de sous-jobs et de composants (OnSubjobOK, OnComponentOK)

Gestion des erreurs de composants, sous-job (OnComponentError, OnSubjobError)

Composant de départ, et de fin (tPreJob, tPostJob)

Attraper et lever des erreurs : tWarn, tDie, tLogCatcher

**Atelier pratique** : assembler plusieurs jobs en un seul, factoriser la gestion d'erreurs pour générer un fichier de compte-rendu

### Les flux de programmation : liens iterate

Différencier les flux principaux des flux iterate

Boucles et programmation via tLoop, tForeach

Passer d'un flux principal à un flux iterate, et l'inverse (tFlowToIterate, tIterateToFlow)

Synchronisation de jobs avec les tWaitForFile, tWaitForSocket, tWaitForSQ

**Atelier pratique** : paralléliser des sous-jobs, synchroniser les jobs avec un événement externe

### Utilisation avancée

Utilisation de composants évolués : Java ou expressions régulières

La spécificité des flux XML dans Talend

Réutilisation du code : les routines

**Atelier** : Utilisation des composants tJavaRow, tJavaFlex, utilisation de l'éditeur Java embarqué, déploiement d'une routine personnalisée

### Débogage dans Talend

Rappels sur le débogage de traces

Utilisation du mode Debug Java : perspective Debug, mode pas à pas

Visualisation et modification à chaud des variables

**Atelier** : Debuggage basique/avancé de tâches

### Déploiement et industrialisation

Rappel sur les contextes

Export des jobs pour exécution sur un serveur

Passage de contexte entre jobs, à l'exécution

Utilisation d'un fichier pour charger le contexte (tContextLoad)

**Atelier** : Déploiement de tâches et chargement dynamique du contexte

### Utilisation avancée

Utilisation de composants évolués : Java ou expressions régulières

Création de logs spécifiques

Gestion des flux XML

Réutilisation du code

**Atelier** : Utilisation des composants tJavaRow, tJavaFlex - création de logs avec tCatch, tDie ... - utilisation de l'éditeur Java embarqué

### Les composants Talend

Concepts essentiels pour étendre Talend Open Studio

Création de la maquette du composant avec tJavaFlex / interface graphique

Découvrir la forge Talend et ses composants communautaires

Intégration dans Talend Open Studio des plugins créés ou téléchargés

**Atelier :** Création d'un composant simple via tJavaFlex, intégration de composants de la forge Talend

## MODALITÉS

### Modalités

**Modalités :** en présentiel, distanciel ou mixte . Toutes les formations sont en présentiel par défaut mais les salles sont équipées pour faire de l'hybride. – Horaires de 9H à 12H30 et de 14H à 17H30 soit 7H – Intra et Inter entreprise.

**Pédagogie :** essentiellement participative et ludique, centrée sur l'expérience, l'immersion et la mise en pratique. Alternance d'apports théoriques et d'outils pratiques.

**Ressources techniques et pédagogiques :** Support de formation au format PDF ou PPT Ordinateur, vidéoprojecteur, Tableau blanc, Visioconférence : Cisco Webex / Teams / Zoom.

**Pendant la formation :** mises en situation, autodiagnostic, travail individuel ou en sous-groupe sur des cas réels.

### Méthode

**Fin de formation :** entretien individuel.

**Satisfaction des participants :** questionnaire de satisfaction réalisé en fin de formation.

**Assiduité :** certificat de réalisation.

**Validations des acquis :** grille d'évaluation des acquis établie par le formateur en fin de formation.