

## Formation Talend Open Studio Data Integration : Initiation

■ <b>Durée :</b>	3 jours (21 heures)
■ <b>Tarifs inter-entreprise :</b>	1 895,00 CHF HT (standard) 1 516,00 CHF HT (remisé)
■ <b>Public :</b>	Administrateurs systèmes/BDD, Chefs de projets, Développeurs
■ <b>Pré-requis :</b>	Notions de Java et fondamentaux SQL
■ <b>Objectifs :</b>	Comprendre le processus d'intégration de données - Maîtriser l'outil Talend Open Studio et la bibliothèque de composants
■ <b>Modalités pédagogiques, techniques et d'encadrement :</b>	<ul style="list-style-type: none"><li>• Formation synchrone en présentiel et distanciel.</li><li>• Méthodologie basée sur l'Active Learning : 75 % de pratique minimum.</li><li>• Un PC par participant en présentiel, possibilité de mettre à disposition en bureau à distance un PC et l'environnement adéquat.</li><li>• Un formateur expert.</li></ul>
■ <b>Modalités d'évaluation :</b>	<ul style="list-style-type: none"><li>• Définition des besoins et attentes des apprenants en amont de la formation.</li><li>• Auto-positionnement à l'entrée et la sortie de la formation.</li><li>• Suivi continu par les formateurs durant les ateliers pratiques.</li><li>• Évaluation à chaud de l'adéquation au besoin professionnel des apprenants le dernier jour de formation.</li></ul>
■ <b>Sanction :</b>	Attestation de fin de formation mentionnant le résultat des acquis
■ <b>Référence :</b>	BUS875-F
■ <b>Note de satisfaction des participants:</b>	4,59 / 5

■ <b>Contacts :</b>	commercial@dawan.fr - 09 72 37 73 73
■ <b>Modalités d'accès :</b>	Possibilité de faire un devis en ligne (www.dawan.fr, moncompteformation.gouv.fr, maformation.fr, etc.) ou en appelant au standard.
■ <b>Délais d'accès :</b>	Variable selon le type de financement.
■ <b>Accessibilité :</b>	Si vous êtes en situation de handicap, nous sommes en mesure de vous accueillir, n'hésitez pas à nous contacter à referenthandicap@dawan.fr, nous étudierons ensemble vos besoins

## 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

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 métadonnées, 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.**