

## Programme de formation Pentaho Data Integration

### ● Objectifs

Cette formation Pentaho Data Integration vous apportera les outils et connaissances nécessaires pour des intégrations de données réussies, exploiter le référentiel Pentaho mais aussi mettre en place, suivre et piloter vos intégrations.

### ● Pré requis

Connaissance de l'environnement des ETL recommandée.

### ● Durée

3 jours

### ● Public

Utilisateurs

### ● Plan de formation

#### Introduction à l'outil Pentaho Data Integration

Présentation et objectif de la formation  
Vue d'ensemble des possibilités offertes par PDI  
Lancer et configurer PDI  
Les transformations en pratique  
Transformation d'un Hello World  
Gestion des erreurs

#### Le référentiel

Créer un référentiel  
Alimenter un référentiel

#### Lecture et écriture de fichier

Données d'entrée et de sortie  
Nettoyage de données  
Paramétrer le chemin d'accès au fichier dans les propriétés de Kettle  
Générer des données de sortie multiples grâce aux paramètres  
Créer un fichier sérialisable à partir de plusieurs fichiers  
Désérialiser un fichier

#### Travailler avec une base de données

Se connecter et explorer une base de données source  
Lecture et écriture de tables  
Lire les données d'une table et les écrire dans un fichier texte  
Epuración de données

Variables et paramètres  
Insertion de données à l'aide des fonctions Insert et Update  
Tables input et paramètres timestamp  
Utiliser l'assistant copie de table (Copy Table Wizard)

#### Flux de données et lookups

Copier et distribuer les données  
Les lookups  
Fusion et jointures de données  
Traitement parallèle  
Utilisation de l'étape « Stream Lookup »

#### Les calculs

Group By  
Utiliser la calculatrice  
Les expressions régulières  
Expression en Java définie par l'utilisateur  
Le JavaScript dans Pentaho  
Exercices de calcul

#### Les jobs

Introduction  
Envoi d'alertes  
Utiliser le kitchen pour exécuter des jobs via la ligne de commande  
Vérifier l'état de la JVM  
Les loops  
Gestion des erreurs et envoi de mails de notification



## Planification

Paramétrer le planificateur

Planifier un job

Pilotage des jobs planifiés