

# PENTHAO

Durée

3 jours

Référence Formation

4-OSHIFT

## Objectifs

Implémenter une solution ETL avec PDI.

Découvrir le traitement et l'alimentation des données, la sécurisation et l'optimisation des flux.

## Participants

Administrateurs systèmes/BDD, Chefs de projets, Développeurs

## Pré-requis

Connaissance du SQL.

## Moyens pédagogiques

Accueil des stagiaires dans une salle dédiée à la formation équipée d'un vidéo projecteur, tableau blanc et paperboard ainsi qu'un ordinateur par participant pour les formations informatiques.

Positionnement préalable oral ou écrit sous forme de tests d'évaluation, feuille de présence signée en demi-journée, évaluation des acquis tout au long de la formation.

En fin de stage : QCM, exercices pratiques ou mises en situation professionnelle, questionnaire de satisfaction, attestation de stage, support de cours remis à chaque participant.

Formateur expert dans son domaine d'intervention

Apports théoriques et exercices pratiques du formateur

Utilisation de cas concrets issus de l'expérience professionnelle des participants

Réflexion de groupe et travail d'échanges avec les participants

Pour les formations à distance : Classe virtuelle organisée principalement avec l'outil ZOOM.

Assistance technique et pédagogique : envoi des coordonnées du formateur par mail avant le début de la formation pour accompagner le bénéficiaire dans le déroulement de son parcours à distance.

## PROGRAMME

### - Introduction et découverte de l'environnement

Pourquoi Pentaho Data Integration (PDI) ?

Historique et présentation

Architecture, vue d'ensemble des composants.

Installation et Configuration

### - Découverte des flux

Les transformations

Les jobs

Text file Output et CSV Input

Exécuter un job.

Exécuter une transformation.

### - Traitement des données

Qualifier les données à traiter : Sort, Unique et Filter Rows...

## CAP ÉLAN FORMATION

www.capelanformation.fr - Tél : 04.86.01.20.50

Mail : contact@capelanformation.fr

Organisme enregistré sous le N° 76 34 0908834

[version 2023]

Gérer les chaînes de caractères : String cut  
Effectuer des calculs : Calculator  
Joindre des données : Join Rows, Merge Join et Rows...).

#### - Les métadonnées

Métadonnées : vue d'ensemble et création.

#### - Travailler avec les bases de données

Bases de données supportées.  
Récupérer une table d'une métadonnée.  
Interaction avec la base de données.

#### - Développement de transformations avancées

Propriétés d'un job et transformation.  
Utiliser des variables : Set / Get Variables.  
Copy / Get Rows to Result.  
Alimenter des données via des services Web  
Monitoring (" Write to log " et mails).

#### - Gestion des logs et des erreurs

Les types d'erreurs.  
Gérer les erreurs (dans un job, une transformation).  
Paramètres de gestion d'erreurs.  
Le composant " Data Validation ".  
Les niveaux de logs (via la console, sauvegardés en base...).

#### - Fonctionnalités avancées, automatisation et documentation

Débogage (prévisualisation et ajout de points d'arrêt).  
Parallélisation des traitements.  
Export des travaux.  
Kitchen : l'exécuteur de jobs (fichiers, repository, tâches planifiées).  
Documenter ses jobs.