USCB1B - Introduction à la gestion de données à large échelle

Mis à jour le 13-03-2026

Code : USCB1B

Unité spécifique de type mixte

5 crédits

Responsabilité nationale : EPN05 - Informatique / Nicolas PIOCH

Présentation

Prérequis

Pour suivre cette UE, l'auditeur doit avoir des connaissances en bases de données (niveau L2 Bases de données).

Objectifs pédagogiques

Cette UE s'adresse aux auditeurs souhaitant obtenir le diplôme de niveau II de concepteur-architecte informatique, et/ou préparant le diplôme d'ingénieur spécialité informatique, et/ou un master spécialité informatique.

L'objectif de cette UE est d'acquérir les bases nécessaires à l'appréhension des nouvelles problématiques liées à la gestion des données à large échelle dans l'entreprise.

Les nouvelles problématiques autour de la donnée concernent d'une part le stockage et la gestion du volume et de l'hétérogénéité (BD relationnelles ou NoSQL, stockage réparti dans le cloud, entrepôts de données, lacs de données, etc.) et d'autre part les dimensions liées à la qualité des données dans ces nouvelles solutions technologiques (intégrité, sécurité des données, confidentialité des données personnelles, coût énergétique etc.).

Compétences

- Savoir concevoir et optimiser une base de données relationnelle.
- Savoir interroger une base de données relationnelle.
- Connaître les principes des bases de données fédérées.
- Connaître les architectures de systèmes décisionnels fondés sur un entrepôt de données et/ou sur un Data Lake.
- Connaître les solutions de Business Intelligence liées au Big Data et aux solutions Cloud.
- Savoir concevoir des bases de données dans les principaux types de stockages NoSQL.
- Savoir optimiser une base de données NoSQL.
- Savoir utiliser les algorithmes de base liés à l'exploitation du NoSQL.
- Connaître les règlements et les méthodologies pour une gestion responsable des données à large échelle.

Programme

Contenu

Partie 1 : Systèmes d'information et gestion des données

- Systèmes d'information et bases de données (introduction)
- Des données aux bases de données
  - Modélisation relationnelle des données
  - SQL
  - Optimisation

Partie 2 : Au-delà du relationnel

- Bases de données fédérées
  - Gérer l'hétérogénéité des données
- Bases de données multidimensionnelles
  - Modélisation multidimensionnelle
  - Architecture d'un système décisionnel (ETL, Data Warehouse, Data Mart, Reporting etc.)
- Systèmes décisionnels et Big Data vers la Business Intelligence
- Solutions Cloud

Partie 3 : Bases de données NoSQL

- Introduction au NoSQL
- Usage du NoSQL (avantage, types d'usage ..)
- Développement de BD NoSQL : modèles de données, distribution, passage à l'échelle, …
    - Les diverses BD NoSQL : Clé-valeur, orientées colonne, orientées document, orientées graphe
    - Les algorithmes (Map Reduce, Elastic Search)
- Bases de données NoSQL et performances

Partie 4 : Gestion responsable des données

- Green IT : stockage et interrogation large échelle
- Sécurité des données
- Intégrité des données
- Confidentialité des données personnelles

Modalités de validation

- Examen final

Description des modalités de validation

Examen final