

Gestion de données massives



Composante
École Nationale
Supérieure des
Ingénieurs en
Arts Chimiques



Volume horaire
12h

En bref

- **Code:** LP1A5230
- **Ouvert aux étudiants en échange:** Oui

Présentation

Objectifs

Théorie et concepts

- Données massives (« Big data » ou mégadonnées)
- Modèle non relationnel

Compétences

- Comprendre les limites des bases de données relationnelles
- Comprendre la différence ACID <> BASE
- Découvrir les différents types de bases de données •Connaître leurs avantages et inconvénients
- Savoir choisir le type de base adapté
- Manipuler des bases de données non relationnelles

Outil

- XML, XSD, Neo4J, Cypher

Compétences (détails)

- #Identifier le "Big Data" et ses 5 caractéristiques et les limites des bases de données relationnelles
- #Comprendre la différence ACID <> BASE
- #Découvrir les différents types de bases de données "NoSQL" et identifier leurs avantages et inconvénients
- #Manipuler des bases de données non relationnelles de type orientées agrégats et orientées graphes
- #Maîtriser les technologies XML, XSD, Neo4J et CYPHER

#Savoir exploiter AuraDB graph database

Pré-requis obligatoires

Sur la base des acquis de 1A et 2A "savoir exploiter une base de données relationnelle" (Méthode Merise, Normalisation, Langage SQL, SGBD-R)

Syllabus

- I - Généralités, vers le « big data »
- II - Base de données non relationnelle, « NoSQL »
- III - Bases de données documents (XML)
- IV - Bases de données graphes (Neo4J)

Informations complémentaires

2 CM puis séances en salle informatique pour l'apprentissage des technologies en mode "LearningByDoing" via une série de problèmes à résoudre de difficulté progressive.

Une première série pour l'apprentissage de la modélisation et technologie XML et XSD.

Une deuxième série pour l'apprentissage de la modélisation et technologie Neo4J/AuraDB. Par exemple, organisation des données relatives à la coupe du monde de rugby 2023, en utilisant une base de données orientée graphe.

Exploitation d'IA génératives (ChatGPT, Copilot, ...) pour la génération de code et analyse comparative-critique.

Evaluation sous forme d'Epreuve écrite individuelle

Bibliographie

#Bases de données : Concepts, utilisation et développement Ed. 4, JL Hainaut, Dunod, 2018

#Management des données de l'entreprise : Master Data Management et modélisation sémantique, P. Bonnet, Hermès-Lavoisier, 2009 (A7Media)

#Enjeux et usages du big data, Technologies, méthodes et mise en œuvre, C. Brasseur, Hermès-Lavoisier, 2013

#Les bases de données NoSQL et le BigData, Rudi Bruchez, Eyrolles, 2015

#Bases de données orientées graphes avec Neo4j, François-Xavier Bois, Eyrolles, 2016

#Neo4j, A graph protect story, Sylvain Roussy, D-Booker, 2019