

Défis informatiques du Big Data

Semestre 3
Code Apogée : LE201MXC

3 ECTS
16h de CM et 12h de TD

Master Santé Publique
Parcours M2 B3H
Biostatistics, Bioinformatique, Biomathématiques for Health
Responsables de l'UE :
Stéphane DERRODE

DOMAINE SCIENCES TECHNOLOGIE SANTE
Composante de gestion de l'UE :
Médecine Lyon Est
Service des Etudes Interdisciplinaires en Santé

PROGRAMME DE L'ENSEIGNEMENT

Objectifs de la formation

Dans de nombreux domaines scientifiques, tels que la biologie ou les sciences de l'environnement, l'évolution rapide des instruments scientifiques ainsi que le recours intensif à la simulation informatique ont conduit, ces dernières années, à une production importante de données. Les applications scientifiques sont confrontées à de nouveaux problèmes liés essentiellement au stockage et à l'exploitation de ces données. L'enseignement proposé permettra de découvrir les grandes problématiques soulevées par l'émergence de ces flots de données (stockage, interrogation, analyse et visualisation), et d'aborder certaines solutions technologiques actuellement proposées. Les questionnements d'ordre éthique et juridique que soulèvent la gestion des données.

Mots-clés : Big Data, Data scientist, NoSQL, Hadoop, Big Data analytics, Open Data, Web des données

Programme :

- Big Data : une introduction aux enjeux, perspectives et applications
- Problématique des grandes bases de données (NoSql, NewSql)
- Big Data et modèle économique : le cas de l'intermédiation
- Open Data : données publiques ouvertes
- Big-Data Analytics : les bases de l'analyse de grands volumes de données
- Représentation et visualisation des données
- Trois TP informatique long (4h) sur la visualisation, sur Apache/Hadoop et sur le web des données (SparQL).

Compétences

A l'issue de ce MOS, l'étudiant sera capable :

- De connaître les enjeux, les opportunités et les problèmes éthiques soulevés par le big-data. - De réaliser des programmes Hadoop/Map-Reduce simples pour exploiter des données réparties.
- De manipuler des bases NoSql à l'aide d'un SGBD moderne (p. ex. Mongo-Db).

Modalités d'évaluations de l'unité d'enseignement :

Session 1 : Ecrit 2h 50% examen terminal + Rapport 50% contrôle continu

Session 2 : Pas de session 2

(Modalités données à titre indicatif, les modalités précises sont votées tous les ans par le CA sur proposition du CFVU)

Vos enseignants dans cette UE : S. DERRODE et intervenants extérieurs