

Offre de sujet TER M1 IFI

Formation

Nom: Master MBDS
Adresse: 1645 route des Lucioles, Sophia Antipolis
Ville: Biot Code Postal: 06410
Site Web: www.mbds-fr.org

Contact

Nom: MOPOLO-MOKE Gabriel
Téléphone Fixe / Mobile: 0673212284
E-mail: Gabriel.Mopolo@gmail.com

Encadrement du TER : Gabriel MOPOLO-MOKE / Benjamin RENAUT vacataire MBDS

Etudiant(e)(s)

Nombre de postes : 1
Niveau étude : 4^{ème} année Informatique

Description du TER

Titre: Utilisation des outils de Business Intelligence Oracle pour l'analyse de données Hadoop et NoSQL

Résumé (20 lignes maximum):

Deux cours ont été construits au MBDS sur Hadoop et la base de données NoSQL Oracle. L'objectif de ce TER est dans un premier temps de construire une Base de Données SQL (DWH) avec des données volumineuses provenant de Hadoop et de la Base de Données NoSQL Oracle. Dans un second temps d'utiliser les outils de Business Intelligence intégrés dans le moteur Oracle 12c pour l'analyse avancée données. Il s'agit de l'utilisation de fonctions de Datamining pour la segmentation et la classification de données. L'objectif final est de créer une application de référence incrémentale autour de la solution Bigdata Oracle. Cette application servira de cas pour les cours Big Data du master.

Principaux objectifs:

- Construire la Base de données DWH Oracle 12c
- Utiliser les connecteurs Oracle Hadoop pour pomper les données vers la base DWH Oracle 12c
- Utiliser les connecteurs Oracle NoSQL pour pomper les données vers la base DWH Oracle 12c
- Utiliser les outils de la BI intégrés dans le moteur Oracle 12c

Principales activités:

- Trouver deux jeux de données volumineuses. Un jeu de données qui sera géré via HADOOP et un autre qui sera géré via Oracle NoSQL
- Mettre en œuvre le connecteur Oracle Hadoop
- Mettre en œuvre le connecteur Oracl NoSQL
- Utiliser les outils intégrés de la BI Oracle.

Technologies : SQL, PLSQL, NoSQL, JAVA

Autres Informations

Mots-clés : SQL, NoSQL, Big Data, Business Intelligence (BI), Dataware House (DWH), Data Mining

Durée : Durée du TER.

Priorité : La partie connecteur Hadoop sera traitée s'il du temps.

