

Projet : calcul de statistiques RDF par Hadoop Map/Reduce

Etant donné un jeu de données RDF (en format ntriples), écrire un ou des programmes MapReduce qui calculent (et les faire exécuter par Hadoop) :

1. Le nombre d'apparitions de chaque URI et chaque littéral,
 - dans les champs « sujet » des triples RDF
 - dans le champ « propriété » des triples RDF
 - dans le champ « objet » des triples RDF
2. Les 10 propriétés les plus fréquentes dans le jeu de données
3. Les 10 classes apparaissant le plus fréquemment dans le jeu de données
4. Les 10 sujets ayant le plus grand nombre de propriétés distinctes spécifiées.