

## **Proposition de stage recherche en laboratoire 2016-2017**

### **Titre : Mise en Oeuvre d'un Plate-forme d'Analyse de Grandes Masses de Données - Application au trafic routier transManche**

#### **Description du sujet :**

Le projet TransManche knowledge est une initiative incluant l'ULCO (laboratoire LISIC et l'EILCO), l'Université de Lille (laboratoire Painlevé), le port de Calais et de nombreux organismes tels que les sociétés d'autoroutes, etc. Le but du projet est de disposer d'une plate-forme permettant l'agrégation et l'analyse des diverses données pouvant également être massifiées dans le but d'extraire une connaissance fine permettant entre autres de mieux réguler le trafic des véhicules au niveau TransManche (de la côte d'Opale vers l'Angleterre et Vice Versa).

#### **Problématique :**

Dans ce cadre, la problématique de ce sujet de master concerne la possibilité de disposer d'une plate-forme hébergeant des services destinés à l'analyse de données massifiées. Il s'agit de proposer des services permettant de faciliter entre autres :

- l'intégration de nouvelles sources de données alimentant un processus d'analyse ;
- la transformation ou mapping de données pour les mettre en conformité avec tel ou tel outil d'analyse ;
- le lancement d'une analyse dans une infrastructure distribuée ;
- des services de visualisation et d'exploration de données à base de réseaux, etc ;
- etc.

Le but n'est pas de redévelopper toutes ces fonctionnalités mais de proposer une manière de les encapsuler pour faciliter leur intégration, combinaison et coordination. L'idée est qu'il soit possible de disposer d'une sorte de *programmation in the large* dans laquelle les instructions de base sont les services usuellement utilisés dans le cadre de l'analyse de données massifiées ou Big Data.

#### **Solutions préconisées :**

Nous préconisons l'utilisation de la notion de web services et de business processus pour réaliser les attentes citées dans la partie problématique. La particularité de l'analyse des données massifiées nécessitera une attention particulière au niveau de la définition et du déploiement des différents services. Le candidat est appelé à intervenir dans une architecture type de traitement de données big data (hadoop de Apache), à proposer une intégration de la notion de web service à ce type de plate-forme et proposera une manière d'intégrer ces services dans les business processes qui seront codés en BPMN.

**Encadrant(s) :** Mourad Bouneffa

**email :** bouneffa@lisic.univ-littoral.fr