



OFFRE DE STAGE Master Recherche

Etude et implémentation d'une solution Spatial Big Data Warehouse : Application aux données agro-climatologiques

PRESENTATION DES ETABLISSEMENTS DE RECHERCHE D'ACCUEIL

En particulier, l'Equipe COPAIN, d'Irstea de Clermont Ferrand, conçoit et gère des systèmes d'information, tels que des entrepôts de données ou des systèmes de gestion de connaissance, adaptés aux problématiques agro-environnementales.

CONTEXTE DU STAGE

Aujourd'hui à l'échelle internationale, nationale et régionale différentes initiatives en faveur de l'agriculture numérique sont déployées (par exemple, le Laboratoire d'Innovation Territoriale pour les grandes cultures de Limagne-Val d'Allier) Dans ce contexte d'innovation, des *systèmes d'information de gestion agricole* (SIGA) ont été développés. Les SIGA sont des systèmes pour la gestion des opérations techniques, financiers, d'inventaire, de traçabilité, de reporting, etc. Prendre en compte les facteurs climatiques est une tâche incontournable pour créer un véritable instrument utile pour aider les agriculteurs à analyser les choix de gestion. En fait, les variables *climatologiques*, telles que la température de l'air, l'humidité relative, le rayonnement solaire, la pluie, etc., interagissent constamment avec les plantes et les parasites. Avec l'avènement des technologies des fermes intelligentes (équipées de capteurs), des échelles spatio-temporelles plus détaillées deviennent possibles. Cela générera des données spatio-temporelles massives (*Spatial Big Data*). Par conséquent, une analyse plus précise pourrait être effectuée pour une agriculture de précision innovante et durable. Cependant, les SIGA présentent deux limites principales : (i) ils ne permettent pas l'analyse des données historiques, ce qui est nécessaire pour comprendre les bonnes pratiques durables (ii) ils n'intègrent pas dans l'analyse les données climatiques.

DESCRIPTION DU STAGE

Pour atteindre un tel objectif, nous proposons un système Big Data basé sur les technologies des entrepôts de données spatiales (EDS) et de l'SOLAP. Les EDS et le SOLAP peuvent être un support efficace pour l'analyse des activités agricoles avec des données climatiques. Ces systèmes sont des technologies permettant l'analyse en ligne de grandes quantités de données.

La conception des EDS dans un contexte de Spatial Big Data est une problématique ouverte car :

1. *Les différentes technologies de bases de données spatial big data offrent toutes des fonctionnalités différentes et parfois incomplètes*
2. *Il n'y a pas un modèle des EDS Big expérimenté avec succès pour les données agro-climatologiques.*

PLAN DE TRAVAIL :

1. Définir les fonctionnalités et caractéristiques nécessaires pour les EDS Big
2. Etudier les solutions existantes (NoSQL, NewSQL, plateformes)
3. Définir un modèle de EDS Big pour les données agro-climatologiques.
4. Implémentation du modèle

Les compétences en informatique recherchées sont :

- Entrepôts de données, et OLAP (i.e Mondrian)
- Technologies Big Data (bases de données – ex : Cassandra, MongoDB, etc. – plateformes –ex : Hadoop)
- Java
- Les compétences en *Systèmes d'Information Géographique* seront considérées comme *un plus*

Compétences autres :

- Capacité à travailler en équipe et aussi en autonomie
- Capacité à lire et comprendre des articles scientifiques en Anglais

INFORMATION PRATIQUES

Le stage se déroulera sur 6 mois à partir de avril-mai 2018.

L'étudiant sera accueilli dans les locaux d'Irstea à Clermont Ferrand.

La rémunération est d'environ 500€ (selon les règles officielles de calcul de gratification des stages)

L'encadrant est : **Sandro Bimonte**, Chercheur en Informatique à l'Irstea de Clermont-Ferrand

Lettre de motivation et CV sont à transmettre à : sandro.bimonte@irstea.fr;