



**Observatoire astronomique de Strasbourg
Centre de Données astronomiques de Strasbourg (CDS)**

11 rue de l'Université
67000 Strasbourg

Proposition de stage (année universitaire 2015/2016) d'élève ingénieur en informatique (UTBM)

Big Data en astronomie (Hadoop, MapReduce, Spark, etc.)

Contexte: Le CDS est une équipe d'une trentaine de personnes (1/3 d'astronomes, 1/3 d'informaticiens et 1/3 de documentalistes) au sein de l'Observatoire astronomique de Strasbourg (env. 80 personnes). Il héberge plusieurs services de données astronomiques de renommée mondiale et il est un acteur majeur de l'IVOA (International Virtual Observatory Alliance) qui définit des standards et développe des prototypes pour l'interopérabilité des données et des services astronomiques. Le CDS mène une activité soutenue de R&D depuis de nombreuses années afin d'améliorer en permanence la qualité de ses services.

Encadrants côté CDS: André Schaaff (andre.schaaff@astro.unistra.fr), François-Xavier Pineau (francois-xavier.pineau@astro.unistra.fr)

Sujet du stage:

Dans le cadre de l'évolution importante des quantités de données que nous serons amenés à gérer dans les années à venir, nous souhaitons tester et prototyper des applications / implémentations dans le domaine du Big Data. Le CDS propose notamment un service de cross-match (identification croisée) de catalogues de données astronomiques. Ce type de traitement est potentiellement "lourd" et demande la mise en place de techniques appropriées pour assurer de bonnes performances et l'utilisation dans des services en ligne.

Nous avons une première expérience de la mise en oeuvre de Hadoop mais dans un environnement très restreint. Nous souhaitons aller plus loin en améliorant l'algorithme, en mettant en oeuvre cette technologie dans un environnement technique approprié (plusieurs noeuds) et en testant des jeux de données plus étoffés. D'autres technologies comme Spark semblent très prometteuses et nous souhaitons également réaliser des tests comparatifs à la fois avec Hadoop mais également avec notre service de cross-match "maison".

Profil recherché: étudiant intéressé par les aspects R&D, développement Java, éventuellement quelques notions de C, aptitude à apprendre un autre langage de programmation si nécessaire, bonne organisation pour gérer les différentes étapes

Le stage est rémunéré suivant les textes en vigueur (actuellement 508 euros par mois) au moment de la signature de la convention.

Liens:

<http://astro.unistra.fr/>, <http://cdsweb.u-strasbg.fr/>