

## *Développement d'une plate-forme informatique de gestion et d'analyse documentaire*

*Société TecKnowMetrix – <http://www.tkm.fr>*

Contact : Loïc MAISONNASSE - Directeur Recherche & Développement – [lm@tkm.fr](mailto:lm@tkm.fr)

TecKnowMetrix est une jeune société innovante issue de l'université Pierre Mendès France et de l'INRA basée sur Voiron, à proximité de Grenoble. Elle conçoit et développe des méthodes et des outils d'analyse de l'information scientifique et technologique contenue dans des bases de données de brevets et publications.

La combinaison de ces méthodes et outils permet à partir d'un recensement exhaustif donc volumineux d'informations objectives (corpus de brevets, de publications scientifiques, d'articles de presse, etc.) de délivrer une vision synthétique d'un environnement technologique complexe, à l'usage des dirigeants d'entreprises, des responsables de R&D, des investisseurs et des élus de collectivités locales.

Ces informations sont utiles dans le pilotage des projets de recherche des laboratoires publics ou des groupes industriels innovants mais leur usage est rendu complexe par la diversité des sources, l'hétérogénéité des formats et les volumes d'informations extrêmement importants qui doivent être manipulées.

TecKnowMetrix a développé une plate-forme logicielle qui permet de ramener en un format commun des sources hétérogènes de données. Cette plate-forme permet ensuite d'appliquer des logiciels d'analyse sémantique, de *datamining*, de *textmining*, de cartographie de l'information que TecKnowMetrix utilise et développe pour restituer à ses clients une vision synthétique et directement utilisable d'un domaine technologique donné.

Techniquement, la plate-forme TecKnowMetrix est basée sur un couplage des langages de programmation PHP, JavaScript et FLEX associés à une base de données MySQL. Et la plupart des logiciels d'analyse de corpus développés chez TecKnowMetrix sont dans ces langages.

L'objectif de ce stage est principalement de poursuivre le développement de la plate-forme TecKnowMetrix, en abordant un ou plusieurs des points suivants :

- Améliorer la navigation dans les bases documentaires, par la mise en place d'index de mots-clés permettant la réalisation de recherche avancée (proximités, mots similaires, etc...);
- Développer des méthodes de classification supervisée ou semi-supervisée et des méthodes d'extraction d'information, pour faciliter la classification et la visualisation des documents de la plateforme ;
- Poursuivre le développement des outils de cartographie de corpus, améliorer leur intégration dans la plate-forme TecKnowMetrix et proposer de nouvelles cartographies ;
- Développer une application autonome et sécurisée fonctionnant en local sur un unique poste et permettant des analyses de corpus de documents ;
- Travailler à la sécurisation de la plate-forme TecKnowMetrix ainsi qu'à l'optimisation de son exécution et à la compression des données présentes dans nos bases ;
- etc.

Le stage se déroulera dans les locaux de la société TecKnowMetrix à Voiron (38). Des solutions de logement pourront être proposées au stagiaire si ce dernier le souhaite. Une rémunération sera proposée, son montant est à négocier.

