

Proposition de contrat postdoctoral sur l'acquisition automatique de métadonnées de ressources pédagogiques par analyse de descriptions textuelles

Contexte : Établir des métadonnées sur des ressources, quelles qu'elles soient, est une tâche souvent considérée comme peu gratifiante, fastidieuse et consommatrice de temps. Elle est donc malheureusement, dans la majorité des cas ignorée lors de l'acquisition de la donnée et/ou de sa mise à disposition. De fait, ces données ne sont pas réellement valorisées car peu documentées et donc délaissées par les moteurs de recherche.

Par exemple, il existe un très grand nombre de ressources pédagogiques (ouvrages, vidéos, cours en ligne, ...) accessibles aux professeurs ou aux étudiants qui ne sont pas réellement utilisées voire totalement inconnues de des utilisateurs potentiels. En effet, les auteurs (qui sont très souvent les professeurs eux-mêmes) n'ont pas été sensibilisés à cet aspect « métadonnées ». L'objectif de notre projet est de construire une base de métadonnées décrivant et pointant vers ces ressources, interrogeable à distance.

Objet du contrat : Le travail à réaliser dans le cadre du contrat postdoctoral portera principalement sur la création automatique et a posteriori de métadonnées sur des ressources pédagogiques. Il s'agit, à partir de données informelles décrivant ces ressources, descriptions moissonnées sur des sites spécialisés, de « reconstruire » et/ou compléter des métadonnées afin de décrire formellement et le plus précisément possible les ressources concernées. Pour cela, d'une part, un modèle de métadonnées devra être proposé et d'autre part, des mécanismes d'analyse des descriptions informelles mais aussi, éventuellement, des ressources elles-mêmes (cas de ressources textuelles) sera mis en place afin de renseigner automatiquement une base de métadonnées.

Ce post-doctorat sera réalisé au laboratoire ICube, en collaboration avec le laboratoire LiLPa. Ces deux laboratoires rassemblent des chercheurs travaillant à la fois dans le domaine de l'informatique et dans divers domaines des sciences du langage, incluant la linguistique informatique.

Candidature : Les candidat(e)s pour ce poste doivent être titulaire d'une thèse en informatique avec de très bonnes connaissances en traitement automatique des langues, fouille de textes et recherche d'information.

Une très bonne compréhension de la langue française est indispensable : les ressources à modéliser et les publics concernés seront quasi exclusivement francophones.

La candidature doit comporter un CV détaillé (en français) incluant une liste de publications, une description de la thèse et son adéquation avec le sujet du post-doc ainsi qu'une lettre de motivation. Les coordonnées d'au moins deux référents sont à joindre.

Date limite pour postuler : 05/07/2018

Date prévisionnelle de début de contrat : Septembre 2018

Durée : 12 mois (prolongeable de 6 mois)

Contacts : Pierre Gancarski (gancarski@unistra.fr) et Delphine Bernhard (dbernhard@unistra.fr)

Téléphone : 03 68 85 45 76

Salaire : de l'ordre de 2100 € net/mois négociable.