

PROPOSITION DE TER

TITRE :

Développement de méthodes de mesure de similarité dédiées au mapping d'ontologies biomédicales.

DESCRIPTION :

De plus en plus d'ontologies sont développées dans différents domaines, leur utilisation devenant de plus en plus nécessaire pour le Web sémantique, l'intégration de données.

Dans les sciences de la vie, un nombre très important d'ontologies ont été développées, dans le but de couvrir des domaines spécifiques (par exemple Gene Ontology (GO), Foundational Model of Anatomy (FMA), Adult Mouse Anatomy (MA), &). La plupart de ces ontologies sont accessibles via le site Web OBO, <http://www.obofoundry.org/>.

De nos jours, il y a un nombre important de projets qui nécessitent l'utilisation de plusieurs de ces ontologies. Le réseau d'excellence européen VPH (Virtual Physiological Human <http://www.vph-noe.eu/>) est l'un d'eux.

Afin de pouvoir utiliser de manière intégrée plusieurs ontologies, celles-ci doivent être reliées par des correspondances sémantiques (mappings). On parle de « mapping » d'ontologies. Etant donné la taille et le nombre d'ontologies biomédicales, il est important de pouvoir disposer de systèmes de mapping automatiques (ou semi-automatique).

Nous avons développé un système de mapping d'ontologies, appelé OMIE (Bouzeghoub et al. 2008)). OMIE est un système Interactive et Extensible permettant d'utiliser et de combiner plusieurs méthodes (appelées « matchers ») pour mesurer la similarité entre les concepts d'ontologies. Plusieurs méthodes génériques ont été développées et intégrées au système. Elles sont réparties en trois principaux groupes : les matchers lexico-syntaxiques, les matchers structurels ou topologiques et les matchers sémantiques.

Le but de ce TER est de développer de nouvelles méthodes plus spécifiques au domaine biomédical. Nous nous intéresserons en particulier à développer un matcher lexical basé sur UMLS (Unified Medical Language System, <http://www.nlm.nih.gov/research/umls/>), et un matcher sémantique qui permettrait de prendre en compte des informations contenues dans les sources de données biomédicales.

Ce TER se déroulera dans le cadre du projet du réseau d'excellence européen VPH, et les méthodes développées seront testées et appliquées aux ontologies utilisées et développées dans ce projet.

Références :

A. Bouzeghoub, A. Elbyed, F. Tahi, « OMIE: A tool for Ontology Mapping within an Interactive and Extensible environment), International Workshop of Data Integration in the Life Sciences (DILS). Dans LNCS/LNBI proceedings, 2008.

ENCADREMENT :

Le stage sera encadré par Fariza Tahi de l'équipe AMIS-Bio du laboratoire IBISC et de Amel Bouzeghoub de l'Institut Telecom SudParis, en collaboration avec Randy Thomas de l'équipe DYNAMIC du laboratoire IBISC.

LIEU :

Le stage se déroulera au laboratoire IBISC, dans 1 équipe AMIS-BIO (<http://amisbio.ibisc.fr/>).
2^{ème} étage, tour Evry 2, 523 place des Terrasses. Evry.

CONTACT :

Fariza Tahï : fariza.tahi@ibisc.univ-evry.fr

Tel : 01 60 87 38 20

Fax : 01 60 87 37 89