

Sujet de Stage de Recherche Master 2 SEP 2017-2018

SUJET SOUS RESERVE D'OBTENTION DE FINANCEMENT (soumis à appel d'offre MNHN ATM)

Unité et équipe d'accueil : UMR7205– ISYEB

Titre : Mettre en oeuvre et évaluer l'apport des ontologies sur des bases de connaissances taxonomiques

Directeur de stage : Régine VIGNES LEBBE

Sujet :

1) PROBLÉMATIQUE – OBJECTIF

Il s'agit d'évaluer concrètement la valeur ajoutée du modèle RDF utilisé dans les ontologies informatiques en termes d'accès et de raisonnement sur une base de connaissances visant à comparer les données morphologiques sur un groupe de taxons. Les contenus utilisés ont déjà été formalisés avec le système expert Xper3. Cette plateforme, largement utilisée par les systématiciens, est une priorité de l'UMR ISYEB. Le projet permettra de comparer son modèle de connaissances et les fonctionnalités qu'elles offrent aujourd'hui avec les modèles et raisonnement apportés par des ontologies en RDF. Il guidera de futures évolutions de Xper3 et pourrait rendre plus accessible aux systématiciens la création d'ontologies et le partage de connaissances pour la comparaison de taxons.

2) MÉTHODE

Calendrier prévisionnel :

mois 1 : formation conceptuelle et pratique sur les ontologies et le modèle RDF. Bibliographie et mission à Montpellier.

Prise en main de la base de connaissance choisie sous Xper3.

mois 2 : Familiarisation avec la morpho-anatomie du groupe et la terminologie.

Traduction du modèle descriptif Xper3 en triplets RDF

mois 3 : en collaboration avec les systématiciens du/des groupe(s) taxonomique(s) choisi(s), préparation d'une liste d'interrogations sur chaque modèle, mise en œuvre.

mois 4 suite de la mise en œuvre. Comparaison des points forts/points faibles de chaque modèle.

Analyse des résultats à la lumière de la bibliographie sur l'utilisation d'ontologies biologiques.

mois 5 : fin de la rédaction du mémoire qui sera débutée au fur et à mesure des étapes.

3) FAISABILITÉ

Au sein des UMR CR2P et ISYEB de nombreuses bases de connaissances ont été réalisées avec Xper3. Nous disposons donc de contenus bien structurés et bien maîtrisés au sein de nos équipes. Ainsi, en fonction de l'étudiant et de son bagage préalable, nous sélectionnerons une ou deux de ces bases pour ce travail.

La plateforme Xper3 est accessible en ligne (www.xper3.fr) et l'ingénieur informatique qui en assure le développement participera à l'encadrement de l'étudiant. Le modèle de représentation des connaissances est documenté.

Ce travail utilisera également les langages RDF. Une collaboration à distance avec une collègue de Montpellier, spécialiste des ontologies, est prévue, avec une courte mission sur place en début de projet.

4) PORTÉE

Le GDR SémanDiv « Sémantique de la Biodiversité » s'est mis en place début 2017 pour une durée de quatre ans, associant écologues, systématiciens et bio-informaticiens. Ce GDR a pour missions de développer une synergie entre différents acteurs au niveau national pour établir une base conceptuelle et stimuler les recherches dans ce domaine.

Ce travail apportera pour le GDR SemanDiv un exemple concret concernant la Systématique.

5) AUTRES COLLABORATEURS IMPLIQUÉS DANS LE PROJET :

Adeline Kerner (CR2P, Xper3 et)

Isabelle Mougenot (Univ. de Montpellier, spécialiste des ontologies)

Sylvain Bouquin (ISYEB, développeur Xper3)