

# Sujet de Post-doctorat en informatique

## Mise en correspondance d'ontologies / Ontology mapping

**Contact :** Thomas Burger ([thomas.burger@univ-ubs.fr](mailto:thomas.burger@univ-ubs.fr)),

Jean-Marc Bienvenu ([bienvenu-jm@mgdis.fr](mailto:bienvenu-jm@mgdis.fr))

**Durée :** 1 an

**Lieu :** Lab-STICC (UMR 3192), équipe CID (Connaissance, Information, Décision), à Vannes (Université de Bretagne Sud)

**Encadrant universitaire :** Thomas Burger (<http://www-labsticc.univ-ubs.fr/~burger/>)

**Coopération :** MGDIS (PME éditrice de progiciels de gestion). L'embauche d'un ingénieur d'étude par MGDIS est en cours de planification afin de permettre de dégager le post-doctorant du travail d'ingénierie, d'implémentation et de transfert de technologie.

### **Objectifs**

**Objectifs opérationnels :** Mise en place d'outils permettant l'interprétation commune (via le système d'information mis en place par MGDIS) des méta-données statistiques issues de l'INSEE et issues d'EuroStat en dépit d'une structuration ontologique différente.

**Objectifs académiques :** Etude méthodologique de la mise en correspondance d'ontologies : (1) état de l'art des algorithmes existants, (2) propositions de nouveaux algorithmes efficaces et implémentables, (3) publications.

### **Problématique**

La structuration des grands corpus de données à un niveau sémantique par la création de métadonnées est gouvernée par la définition d'ontologies adaptées. Par ailleurs, la multiplication des ontologies dans des domaines différents mais concomitants, nécessite la définition d'interfaces ou de méthodes de conversions entre ces différentes ontologies. A titre d'exemple, les métadonnées associées aux grands corpus de données publiques au niveau national ou européen ne correspondent pas toujours à un même standard. Un système d'information souhaitant faire profiter à ses utilisateurs des données structurées selon ces différents standards doit permettre leur interopérabilité en garantissant que les différentes métadonnées ayant des contenus sémiotiques/sémantiques proches puissent être mis en correspondance.

De nombreuses méthodes de mise en correspondance automatique existe dans la littérature (cf. bibliographie), soit par des méthodes quantitatives (calcul de distance entre les concepts), soit logiques (exploitation des similitudes dans la structure des graphes représentant les ontologies), cependant, les résultats n'atteignent pas encore la fiabilité d'une mise en correspondance manuelle et fastidieuse. Une des difficultés majeures réside dans la potentielle différence de granularité entre les concepts manipulés dans chacune des ontologies à comparer. Ainsi, un concept relativement général peut correspondre dans une autre ontologie à un ensemble de concepts plus précis, certains ne recouvrant que partiellement la notion d'intérêt, mais permettant d'y ajouter une certaine nuance. Cette imprécision permet

justement de refléter la richesse des entités sémiotiques manipulées, néanmoins, la quantification de la pertinence d'une mise en correspondance devient beaucoup plus difficile.

Ainsi, nous proposons au candidat d'appréhender ce problème d'une manière originale, et ensuite de l'appliquer à un cas d'étude précis, correspondant aux besoins de la société MGDIS.

### **Profil**

Docteur en informatique ou en traitement de l'information, ayant des compétences sur les ontologies. En outre, de bonnes compétences en programmation ainsi qu'une aptitude au travail en équipe sont indispensables.

### **Références bibliographiques**

- [1] Référentiel Général Interopérabilité : <http://www.references.modernisation.gouv.fr/rgi-interoperabilite>
- [2] Recommandation INSPIRE : <http://inspire.jrc.ec.europa.eu/>
- [3] Statistical Data and Metadata eXchange : <http://sdmx.org/>
- [4] Eurostat : <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>
- [5] Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art. The Knowledge Engineering Review 18(1), 1-31.
- [6] Jérôme Euzenat and Pavel Shvaiko : Ontology matching, Springer-Verlag, 2007.
- [7] Rong Pan, Zhongli Ding, Yang Yu, and Yun Peng. A Bayesian Network Approach to Ontology Mapping, Proceedings of the Fourth International Semantic Web Conference, 2005.
- [8] A. D'Andrea and F. Niccolucci. "Mapping, Embedding and extending: pathways to semantic interoperability. The case of Numismatic collections", proceedings of the Semantic Interoperability in the European Digital Library workshop, 2008.
- [9] Standart W3C OWL : <http://www.obitko.com/tutorials/ontologies-semantic-web/web-ontology-language-owl.html>