

Titre : Explication de réconciliations de données obtenues par une approche logique et numérique.

Renseignements : [Fatiha.Sais@lri.fr](mailto:Fatiha.Sais@lri.fr)

Mots-clés : Web Sémantique, Ontologies, Réconciliation, Raisonnement.

Contexte : Deux problèmes importants se produisent lorsque l'on s'intéresse à l'intégration sémantique de données :

- (i) **la réconciliation de schémas** qui consiste à mettre en correspondance deux schémas (ou deux ontologies) et calculer les mappings entre les éléments (concepts et relations) des deux schémas ;
- (ii) **la réconciliation de données** qui consiste à mettre en correspondance des données provenant de différentes sources et détecter les données qui représentent la même entité du monde réel, i.e. la même personne, le même lieu, le même article, etc.

Ce sujet s'inscrit particulièrement dans le cadre de la réconciliation de données. Dans de précédents travaux [1], une approche de réconciliation de données a été développée. Cette approche, nommée LN2R, combine une méthode logique et une méthode numérique qui exploite la sémantique des connaissances représentées dans une ontologie, pour calculer les décisions de réconciliation.

Travail : l'objectif de ce travail est de permettre l'explication des décisions de réconciliation obtenues que ce soit par la méthode logique ou par la méthode numérique. Concernant la méthode logique, une explication peut revenir à la preuve qui a conduit à la déduction de la réconciliation (voir les travaux [2,3]). En revanche, pour la méthode numérique aucune formalisation n'est proposée, à notre connaissance, dans la littérature.

Il s'agit donc de :

1. Réaliser une étude bibliographique des approches existantes d'explication de réconciliation de schémas et de données.
2. Proposer un modèle d'explication des décisions de réconciliation obtenues par LN2R.
3. Utiliser les explications obtenues pour améliorer la qualité des résultats de la méthode LN2R.

Une implémentation d'un moteur d'explication est souhaitée.

[1] Combining a Logical and a Numerical Method for Data Reconciliation. Fatiha Saïs, Nathalie Pernelle and Marie-Christine Rousset. (2008)

[2] Explanation in the DL-Lite Family of Description Logics. Alexander Borgida, Diego Calvanese, and Mariano Rodriguez-Muro (2008)

[3] Web explanations for semantic heterogeneity discovery. Paulo Pinheiro Da Silva, Pavel Shvaiko, Fausto Giunchiglia, Fausto Giunchiglia, Deborah L. McGuinness, Deborah L. McGuinness (2005)

Encadrement : Nathalie Pernelle et Fatiha Saïs

Lieu : IASI-Gemo (INRIA-Saclay île de France) à Orsay (<http://gemo.futurs.inria.fr/>).

Début du stage : 01 mars 2009

Fin du stage : 31 juillet 2009 (ou le 1er septembre)

Financement : oui.