

Alignement de ressources sémantiques à partir de règles

Valentina Ceausu*, Sylvie Desprès*

*CRIP 5, Université Paris V
45 Rue des Saints Pères
75006 Paris
ceausu@math-info.univ-paris5.fr
sd@math-info.univ-paris5.fr

Résumé. Ce papier présente une approche automatique pour aligner des ressources sémantiques. L'alignement se traduit par la mise en correspondance des entités (termes, concepts, rôles) appartenant à des ressources d'un même domaine qui peuvent avoir des niveaux de formalisation différents. Les entités correspondantes sont de même nature et un coefficient caractérise leur degré de ressemblance.

L'approche proposée est fondée sur des règles d'appariement entre les entités des deux ressources. Dans une première phase, ces règles d'appariement sont identifiées empiriquement. Des algorithmes combinant les différentes règles identifiées sont ensuite définis afin d'établir des correspondances entre les entités des ressources considérées.

Ce papier présente un ensemble de règles d'appariement exploitant des éléments situés à différents niveaux conceptuels. Cet ensemble constitue un cadre pour l'alignement automatique des ressources sémantiques. Les résultats d'une première expérimentation qui a porté sur l'alignement de deux ressources du domaine de l'accidentologie sont également présentés.

1 Introduction

Avec l'émergence du web sémantique, l'exploitation des ressources sémantiques pour annoter des documents, personnaliser des services ou décrire des ressources disponibles sur le web est devenue essentielle. Créées par des communautés distinctes, existant déjà en grand nombre, les ressources sémantiques se différencient par des niveaux différents de formalisation et de conceptualisation engendrant un certain degré d'hétérogénéité.

Ainsi, l'utilisation conjointe des éléments (documents, services web) décrits par des ressources distinctes est soumise à leur mise en correspondance. Un nouveau problème émerge, qui concerne la mise en correspondance des ressources sémantiques.

Parmi les techniques proposées pour apporter des solutions à ce problème, l'alignement sera considéré dans ce papier. L'alignement établit des appariements entre les entités appartenant à deux ressources distinctes. Au-delà d'une certaine complexité, taille, ou nombre de ressources il est impossible d'établir manuellement ces appariements. La nécessité des méthodes automatiques (ou semi automatiques) pour l'alignement des ressources sémantiques est évidente.