

Méthodologie d'Évaluation Intelligente des Concepts Ontologiques

Lobna Karoui*, Marie-Aude Aufaure*
*Supelec, Plateau de Moulon, Gif-sur-Yvette
{Lobna.Karoui, Marie-Aude.Aufaure}@supelec.fr
http://www.supelec.fr/ecole/si/pages_perso/karoui.html

Résumé. Un des problèmes majeurs dans la gestion des ontologies est son évaluation. Cet article traite l'évaluation des concepts ontologiques qui sont extraits de pages Web. Pour cela, nous avons proposé une méthodologie d'évaluation des concepts basée trois critères révélateurs : « le degré de crédibilité »; « le degré de cohésion » et « le degré d'éligibilité ». Chaque critère correspond à un apport de connaissance pour la tâche d'évaluation. Notre méthode d'évaluation assure une évaluation qualitative grâce aux associations de mots ainsi qu'une évaluation quantitative par le biais des trois degrés. Nos résultats et discussions avec les experts et les utilisateurs ont montré que notre méthode facilite la tâche d'évaluation.

1 Introduction

L'évaluation d'une ontologie est une tâche difficile. Ceci explique l'absence de méthodes d'évaluation standard ou de mesures d'évaluation servant à valider l'ontologie. Dans cet article, nous focalisons notre intérêt sur l'évaluation des concepts de l'ontologie de domaine (appelés concepts ontologiques) qui sont extraits des pages Web. Nous travaillons sur les documents HTML écrits en français, dans le domaine du tourisme. Dans un travail précédent, nous avons défini un contexte structurel qui tient compte du document HTML et développé un algorithme de clustering afin de bien rassembler les mots sémantiquement proches (Karoui et al, 2006). Le résultat de ce travail était constitué de classes de mots pour lesquelles les experts ont du réaliser une lourde tâche d'évaluation et d'étiquetage. Pour aider ces derniers et faciliter l'interprétation sémantique de ces classes (concepts), nous avons défini une méthode d'évaluation basée sur trois critères révélateurs. Ces derniers sont le degré de crédibilité, le degré de cohésion et le degré d'éligibilité. Le degré de crédibilité exploite deux types de contextes : un contexte linguistique et un contexte documentaire. En se basant sur ces deux types de contextes, nous calculons le degré de crédibilité associé à chaque classe de mots et à chaque contexte. Le degré de cohésion calcule le degré de rapprochement des mots d'une classe en utilisant les documents du web. Le degré d'éligibilité sélectionne ou suggère le mot de la classe qui peut être son concept ou qui peut orienter le raisonnement vers le futur concept approprié. Dans ce qui suit, nous détaillerons ces critères ainsi que leur rôle.

2 Les critères révélateurs de l'évaluation intelligente des concepts

Les critères révélateurs assistent l'expert du domaine durant la tâche d'évaluation. Ces critères sont : le degré de crédibilité: le caractère de ce qu'on croit ; le degré de cohésion: le caractère d'une chose dont toutes ses parties sont réunies avec une relation logique entre ses éléments et sans aucune contradiction ; le degré d'éligibilité : le caractère d'un mot qui com-