

MuMie: Une Approche Automatique pour l'Interopérabilité des Métadonnées

Samir Amir, Ioan Marius Bilasco, Thierry Urruty et Chabane Djeraba

LIFL UMR CNRS 8022, Université de Lille1, Telecom-Lille1,
50 avenue Halley Parc scientifique de la Haute-Borne, Villeneuve d'Ascq, France
{samir.amir, marius.bilasco, thierry.urruty, chabane.djeraba}@lifl.fr

Résumé. Avec l'explosion du multimedia, l'utilisation des métadonnées est devenue cruciale pour assurer une bonne gestion des contenus. Cependant, il est nécessaire d'assurer un accès uniforme aux métadonnées. Plusieurs techniques ont ainsi été développées afin de réaliser cette interopérabilité. La plupart d'entre elles sont spécifiques à un seul langage de description. Les systèmes de matching existants présentent certaines limites, en particulier dans le traitement des informations structurelles. Nous présentons dans cet article un nouveau système d'intégration qui supporte des schémas provenant de langages descriptifs différents. De plus, la méthode de matching proposée a recours à plusieurs types d'information de façon à augmenter la précision de matching.

1 Introduction

L'omniprésence des ressources numériques pose le problème de l'efficacité de leur gestion, ce qui a généré un intérêt croissant pour l'utilisation des métadonnées destinées à améliorer la gestion de ces types de contenu. Les métadonnées peuvent véhiculer divers types d'information décrivant les types de contenu multimédia, les informations sémantiques de ces contenus et les caractéristiques des terminaux utilisant ces contenus.

Si on anticipe la croissance des métadonnées, il sera certainement de plus en plus difficile d'obtenir un accès uniforme aux objets multimédia en raison du nombre de communautés indépendantes de métadonnées. En effet, chaque communauté combine les termes à partir de plusieurs vocabulaires spécifiques, et utilise différentes structures pour organiser les métadonnées. L'interopérabilité des métadonnées devient donc un enjeu crucial.

L'intégration manuelle des métadonnées entraîne un coût temps assez important. En plus, elle doit être mise à jour à chaque apparition d'un nouveau format de métadonnées. Ainsi, plusieurs techniques automatiques ont été développées pour faciliter le processus d'intégration. Les techniques de *matching* jouent un rôle central dans le cadre de ces approches. Cependant, peu de systèmes de matching supportent des schémas issus de différents langages. On trouve notamment les systèmes suivants : Similarity Flooding (Melnik et al. (2002)), Cupid (Do et Rahm (2002)). Ces approches n'utilisent pas la majorité des informations structurelles et sémantiques (ex : propriétés d'équivalence, caractéristiques de généralisation, etc.), ce qui rend la détection des matchings complexes (n:m matching) difficile. De plus, un des inconvénients majeurs de ces méthodes est leur façon d'utiliser les informations structurelles.