
Un système d'intégration de métadonnées dédiées au multimédia

Samir Amir

*Laboratoire d'Informatique Fondamentale de Lille
Université des Sciences et Technologies de Lille
Cité scientifique, Bâtiment M3, 59655 Villeneuve d'Ascq, France.
samir.amir@lfl.fr*

MOTS-CLÉS : intégration des métadonnées, alignement sémantique, système de médiation, multimédia

KEYWORDS: metadata integration, semantic alignment, mediation system, multimedia.

Avec l'explosion du web sémantique et multimédia et le besoin de disposer de connaissances directement interprétables par les machines, plusieurs standards de descriptions des métadonnées sont apparus ces dernières décennies (MPEG-7, MPEG-21, TV-Anytime, etc.). Ils permettent de décrire un large éventail d'entités : des contenus multimédia disponibles sur le Web (image, vidéo, audio, etc.), des entités physiques transmettant ou consommant ces contenus (réseaux, téléphone mobile, télévision, etc.), des services accédant à ces contenus (moteurs de recherche, services d'adaptation, etc.) et finalement les profils des utilisateurs consommant ces contenus.

Ces standards de métadonnées sont souvent créés par des communautés multimédia différentes ayant des besoins informationnels et une manière d'encoder les métadonnées spécifiques. Ceci est à l'origine d'une grande hétérogénéité relative à l'utilisation des syntaxes, termes, ontologies et schémas de codage différents.

Les différents clients (utilisateurs, services) traitant ces standards de métadonnées disponibles sur le Web souhaiteraient à moindre coût être capables d'interpréter toutes les informations que ces standards contiennent afin de pouvoir les exploiter de manière unifiée. Cependant, cet objectif est loin d'être réalisable à cause du fait que l'utilisation de nombreux standards existants exige des connaissances spécifiques. Pour faire face à ces limitations, nous proposons une plateforme d'intégration des métadonnées dédiées au multimédia qui permet de gérer les hétérogénéités des différents standards.

La plateforme proposée contient une première couche qui est un schéma virtuel nommé schéma médiateur. Le schéma médiateur constitue une interface entre les utilisateurs et les métadonnées disponibles dans les différents standards, le schéma médiateur contient idéalement toutes les entités requises par l'utilisateur. Ces entités constituant le schéma médiateur sont alignées aux entités auxquelles elles

correspondent sémantiquement dans les différents standards de métadonnées. Le schéma de médiation présente l'avantage de ne pas nécessiter chez l'utilisateur une connaissance a priori sur les concepts des standards avec lesquels il interagit. Ses requêtes sont interprétées depuis le schéma de médiation « instance », et puis sont reformulées au niveau du médiateur en tenant compte des règles de traduction. Ces règles sont définies en fonction des alignements sémantiques effectués entre le schéma médiateur et ceux des autres standards. Les principaux défis sont :

- La sélection et l'adaptation si nécessaire de l'approche la plus prometteuse de l'état de l'art (Kalfoglou *et al.*, 2003) (Erhard *et al.*, 2001) et qui est adéquate avec les schémas de métadonnées dédiés au multimédia. Une spécificité de ce type de schéma est d'être souvent très riche syntaxiquement et sémantiquement.
- La conversion de tous les standards vers un langage commun pour la description des ontologies afin de faciliter l'alignement sémantique, cette conversion doit être faite avec le minimum de pertes structurelles et sémantiques (Chrisa et al 2007).
- La reformulation des requêtes des utilisateurs en fonction d'alignements complexes effectués entre le schéma médiateur et les standards existants. Cette requête doit être capable de retourner l'objet voulu par l'utilisateur en tenant compte du fait que, souvent ses caractéristiques sont décrites par des attributs appartenant aux standards différents. Autrement dit, a priori il est nécessaire d'identifier toutes les informations exprimées dans des standards différents et décrivant un objet donné.

Les trois points mentionnés précédemment constituent les principaux objectifs de notre travail de thèse. La plateforme que nous comptons définir et réaliser constituera une avancée significative, relative à la résolution des problèmes d'hétérogénéité des métadonnées dédiées au multimédia.

Bibliographie

- Kalfoglou Y., Marco S., "Ontology mapping: the state of the art", *The Knowledge Engineering Review*, vol. 18, 2003, p. 1-31.
- Erhard R., Philip B., "A survey of approaches to automatic schema matching", *The Very Large Databases Journal*, vol. 10, 2001, p. 334-350.
- Chrisa T., Stavros C., "Interoperability of XML Schema Applications with OWL Domain Knowledge and Semantic Web Tools". In *6rd International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE07)*, Vilamoura, Portugal Nov 27-29, 2007.