

Un modèle métier extensible adapté à la gestion de dépêches d'agences de presse

Frédéric Bertrand*, Cyril Faucher*,
Marie-Christine Lafaye*, Jean-Yves Lafaye*, Alain Bouju*

*Laboratoire Informatique, Image, Interaction
Université de La Rochelle
Avenue Michel Crépeau, 17042 La Rochelle Cedex 1
prenom.nom@univ-lr.fr
<http://www-13i.univ-lr.fr>

1 Présentation

Dans le cadre du projet RIAM¹ « Relaxmultimédia » mené conjointement avec deux agences de presse (AFP et Relaxnews) nous présentons une approche destinée à gérer deux aspects d'un modèle métier défini avec UML : son extensibilité et la possibilité de naviguer entre les classes et les instances définies à partir de ce modèle. Nous montrons que la transformation du modèle UML en un schéma RDF sur lequel est utilisable SeRQL, un langage d'interrogation, présente des caractéristiques intéressantes pour gérer de tels aspects.

2 Extensibilité et navigation dans le modèle métier

Une dépêche fait généralement référence à un ensemble d'entités nommées qui ont été classées par l'IPTC² en quatre catégories génériques : *Party* (se scindant en deux sous-catégories *Person* et *Organisation*), *Location* représentant les lieux, *Event* représentant les différents types d'événement, et *Artifact* décrivant les objets (physiques ou virtuels) cités dans une dépêche. Pour affiner l'information il est possible de définir des sous-catégories à partir des catégories précédentes. De plus, chaque type d'entité peut être rattaché à un domaine, ceci pour traduire la notion de contexte associé à chaque information. Comme exemple, dans la dépêche (simplifiée) suivante :

Le film «La Guerre des Mondes», avec Tom Cruise, sort le 6 juillet 2005 en France.

on identifie le type *Movie* (sous-type d'*Artifact*, instance «La Guerre des Mondes»), le type *Country* (sous type de *Location*, instance «France»), le type *Release* (sous-type d'*Event*, instance «Sortie du film G. D. M. »), le type *Actor* (sous-type de *Person*, instance «Tom Cruise»).

¹Recherche et Innovation en Audiovisuel et Multimédia.

²International Press Telecommunications Council

Les types *Movie*, *Release* et *Actor* sont associés au domaine « *Arts, culture and entertainment* ». Des relations sémantiques peuvent également être établies entre entités. La conséquence de cette richesse sémantique est que, d'une part, le nombre de classes à définir pour prendre en compte les différentes entités ne peut être connu de manière exhaustive et, d'autre part, ce nombre est beaucoup trop important pour apparaître directement dans le modèle. L'approche que nous avons développée consiste à utiliser le mécanisme d'extension d'UML, les stéréotypes, pour créer un modèle métier générique en « marquant » les classes fondamentales *Party*, *Location*, *Event* et *Artifact* ainsi que les associations entre ces classes. La sémantique associée à ce marquage est que la classe possède N sous-types pouvant être créés dynamiquement. Cette sémantique est conservée lors de la traduction de notre modèle vers un schéma RDF (RDFS) définissant un modèle de navigation entre les dépêches. La raison principale qui a guidé notre choix vers la création d'un modèle RDF/RDFS est la disponibilité d'outils permettant d'interroger notre modèle aussi bien au niveau instance qu'au niveau classe.

La traduction UML→RDFS a été conduite selon un processus décrit par CraneField (2001) en utilisant une feuille de style XSLT pour convertir le format XMI³ représentant le modèle UML en un document RDFS/XML. Le modèle obtenu est géré par une base RDF, Sesame Broekstra et al. (2002), possédant un langage d'interrogation, SeRQL, fondé sur des expressions de chemin, permettant au sein d'une même requête d'exprimer des conditions sur les associations entre classes et les instances de ces classes et ces associations. Le peuplement de cette base a été réalisé par un outil de saisie s'interfaçant avec Sesame. L'organisation de l'information sous forme de triplets RDF et la présence de SeRQL permettent d'exprimer des requêtes qu'il est très complexe (voir impossible) d'exprimer avec SQL sur un modèle relationnel. La navigation a été développée avec une interface graphique permettant une visualisation modulaire (par déploiement partiel) du graphe de connaissance formé par les dépêches.

Pour conclure, nous soulignerons que l'utilisation d'UML permet de définir précisément le modèle de base et les points d'extension grâce à l'utilisation de stéréotypes. Il n'offre cependant pas de possibilités sur l'interrogation du modèle. Un modèle fondé sur RDFS, associé au langage SeRQL, permet cette interrogation à la fois sur les classes et les instances ce qui est utile pour permettre une navigation conjointe dans le modèle et dans les données.

Références

- Broekstra, J., A. Kampman, et F. V. Harmelen (2002). Sesame : A Generic Architecture for Storing and Querying RDF and RDF Schema. In *Proceedings of International Semantic Web Conference*, Sardinia, Italy, pp. 54–68.
- CraneField, S. (2001). Networked Knowledge Representation and Exchange using UML and RDF. *Journal of Digital Information* 1(8).

Summary

We present an approach intended to manage two aspects of a model defined with UML: its extensibility and the possibility of navigate between the classes and instances created from this model. We show that the transformation of UML model into a RDF schema on which is usable SeRQL, a request language, shows interesting features to manage such aspects.

³XML Metadata Interchange