

Sujet de Master 2 Recherche 2012-2013

Alignement sémantique pour la gestion de données personnelles

Encadrant principal (H.D.R.) : Philippe LAMARRE
courriel : Philippe.Lamarre@liris.cnrs.fr
tel. : 04 72 43 72 84

Co-encadrant : Elöd EGYED-ZSIGMOND
courriel : elod.egyed-zsigmond@insa-lyon.fr
tel. : 04 72 43 ????

Mots-clés : Ontologie, Qualification sémantique, Formulaires.

Contexte

Actuellement, pour accéder aux informations qu'il a saisies dans un formulaire web, un utilisateur dépend entièrement du tiers auquel il les a transmises. De plus, elles lui sont restituées sous une forme qu'il ne maîtrise pas.

Dans le cadre d'un projet, nous développons un outil orienté utilisateur, d'usage facile que ce soit en environnement personnel ou professionnel, offrant à chacun la possibilité de gérer ses données en utilisant ses propres ressources. Cela passe par la mémorisation des données et leur mise à disposition pour différents usages. Des opérations naturelles, mais actuellement inenvisageables deviennent ainsi possibles : la recherche dans les informations transmises (quoi, à qui, quand, pourquoi, etc.) ; des alarmes pour signaler à l'utilisateur la présence d'informations sensibles avant l'expédition ; les saisies multiples, souvent longues et fastidieuses, peuvent être évitées grâce à la réutilisation de données déjà saisies ; etc. Pour cela, il est nécessaire de "comprendre" les formulaires en ligne afin de bien tracer les données diffusées pour une réutilisation inter formulaire.

Objectif

L'objectif de ce stage est d'étudier les modèles et les méthodes d'alignement de champs de saisie d'un formulaire par rapport aux concepts d'une ontologie donnée et leurs attributs.

Approche et démarche

Dans un premier temps, on travaillera sur une méthode d'alignement manuel 'off line' des champs d'un formulaire par rapport à une ontologie. Cette partie doit permettre d'acquérir une bonne compréhension des notions liés aux ontologies et à l'annotation sémantique.

Dans un deuxième temps, une approche semi-automatique 'on line' sera envisagée [ES07, E⁺04, BBR11]. L'objectif est de permettre à l'utilisateur de qualifier sémantiquement un formulaire en même temps qu'il le renseigne, ce en tirant partie des informations déjà saisies et déjà qualifiées sémantiquement.

Résultats attendus

- Résultats théoriques : étude d'un mécanisme de qualification sémantique basé sur les instances.
- Résultats pratiques : développement d'un outil de qualification sémantique pour des formulaires web

Positionnement local

Cette thèse est proposée au sein des l'équipes BD et DRIM du LIRIS. L'équipe BD centre ses activités sur la gestion déclarative des données dans les nouvelles applications incluant les services, les flux et les environnements dynamiques. Les problématiques liées au stockage, à la gestion et à l'analyse des données personnelles d'un utilisateur sont tout à fait dans le champ d'investigation. De son côté, l'équipe DRIM. . .

Références

- [BBR11] Zohra Bellahsene, Angela Bonifati, and Erhard Rahm, editors. *Schema Matching and Mapping*. Springer, 2011.
- [E⁺04] J. Euzenat et al. State of the art on ontology matching. Technical Report KWEB/2004/D2.2.3/v1.2, Knowledge Web, 2004.
- [ES07] Jérôme Euzenat and Pavel Shvaiko. *Ontology matching*. Springer-Verlag, Heidelberg (DE), 2007.