

Compte-rendu de la conférence

Ingénierie des Connaissances 2008

L'édition 2008 de la conférence Ingénierie des Connaissances s'est tenue du 18 au 20 juin à Nancy (organisée par *Yannick Toussaint*), comme conférence indépendante pour la première fois depuis 2004. Elle a regroupé plus de 70 personnes venues de toute la France, qui ont pu pendant trois jours discuter des évolutions et de l'avenir de la discipline.

Sur 41 articles soumis (émanant d'équipes de recherche localisées en France et au Maghreb), 21 ont été acceptés en version longue (représentant un taux de sélection de 50%) et une dizaine ont été acceptés en posters. Ces travaux ont été présentés dans le cadre de 9 sessions thématiques décrites ci-après.

Deux conférenciers invités étaient prévus. *Ivan Herman*, du CWI (Amsterdam) est responsable de l'activité Web sémantique au W3C et nous a fait l'honneur de sa présence. Sa présentation dans un excellent français a fait l'état des lieux du Web sémantique et a été l'occasion d'une discussion très riche notamment sur les interrelations entre web sémantique et web social.

Jean-Max Noyer (Université Paris 7) devait présenter un point de vue plus orienté philosophie et sciences de l'information et de la communication sur « L'avenir des écritures aujourd'hui, dans les espaces-temps des mémoires numériques hypertextuelles ». Il n'a malheureusement pu être présent, et un débat improvisé a été mené par *Alain Mille* sur la thématique du paysage informatique français et de ses évolutions en cours (avenir de l'informatique au CNRS et à l'INRIA notamment), avec la participation de *Karl Tombre*, directeur du LORIA.

Un autre débat, organisé par *Sylvie Desprès* et *Nathalie Aussenac-Gilles* a permis d'évoquer la question de l'Ingénierie des Connaissances française dans ses rapports avec le monde anglo-saxon. Plusieurs doctorants, post-doctorants et chercheurs français à l'étranger ont participé à cette table-ronde, en présentiel ou à distance. Encore une fois, les débats furent riches. La question qui reste posée est celle de l'acceptation sans discussion du mode de recherche anglo-saxon, qui semble partagée par les différents expatriés, et la préservation d'une éventuelle spécificité française, notamment dans la manière de penser les problèmes, peut-être plus interdisciplinaire.

Membre du bureau du Gracq et du comité de programme depuis de nombreuses années, *Rose Dieng-Kuntz* n'avait pu nous rejoindre à Nancy. Elle s'est éteinte quelques jours après la conférence. La communauté Ingénierie des Connaissances sait ce qu'elle lui doit et ne l'oubliera pas.

Session 1 – « Diffusion de connaissances médicales » animée par Jean Charlet

La session « Diffusion des connaissances médicales » a commencé par un exposé de J. Bouaud (B. Séroussi et A. Brizon, co-auteurs) sur l'impact de la réactualisation de recommandations de pratiques cliniques sur l'évolution d'une base de connaissances. Le travail présenté consiste, à la suite d'autres papiers des mêmes auteurs sur les guides de bonnes pratiques (GBP), à étudier l'évolution de ces guides dans le temps. Ils utilisent pour cela deux GBP, portant sur la prise en charge des tumeurs superficielles de la vessie, aux années 2002 et 2004. L'article a surtout montré la forte évolution qui a eu lieu entre 2002 et 2004 avec, par exemple, un recouvrement faible, puisque plus de la moitié des recommandations 2002 (59 %) sont considérées obsolètes en 2004 et près des 2/3 des recommandations 2004 (67 %) sont considérées comme nouvelles. Les auteurs notent, pour expliquer cela, un domaine (l'écriture des recommandations) jeune et en forte évolution. La prise en compte des recommandations 2007, pas disponible au moment de l'étude, pourrait être riche d'enseignements.

Le 2^e article, « Mise à disposition des spectres d'activité des antibiotiques pour le médecin généraliste : méthode et résultats » de J. Nobécourt et C. Duclos, décrit le développement et l'évaluation d'un outil, disponible sur Internet, permettant à des médecins de consulter de façon rapide des spectres antibactériens et donc de choisir l'antibiotique à donner à un patient. L'article montre tout l'intérêt d'un tel travail qui tire partie d'un document d'origine déjà bien structuré. Le travail a nécessité la construction manuelle d'une ontologie identifiant, par exemple, les espèces bactériennes, les genres bactériens ou des regroupements de bactéries. Un article typiquement IC, combinant réflexion conceptuelle, inscription dans les usages des utilisateurs et évaluation (en cours). La question de la maintenance des connaissances contenues dans l'outil rejoint les travaux de l'article précédent.

Session 2 – « Interrogation de graphes de connaissances » animée par Philippe Laublet

La présentation effectuée par Fabien Gandon a montré un exemple de comment il était possible de gérer la distribution des requêtes de manière distribuée sur plusieurs serveurs dans le contexte des technologies du Web sémantique. Cette distribution souvent nécessaire, avec des bases d'annotation RDF, nécessite d'identifier les serveurs susceptibles d'avoir des éléments de réponse, de gérer les problèmes de performances et de recomposer les résultats à partir des réponses partielles. L'application à la base de ces réflexions provient d'un projet européen SevenPro visant à développer des outils et des technologies qui permettent de fournir des rendus 3D dans des tâches de conception. Mais, les réflexions présentées peuvent s'étendre à de nombreuses autres applications. Parmi les questions posées, l'une portait sur comment l'approche pouvait s'étendre quand les ontologies sous-jacentes aux différentes bases RDF étaient différentes

L'autre exposé présenté par Nicolas Moreau du Lirmm à partir du papier écrit avec Michel Leclère propose une réflexion approfondie sur les types de réponses obtenues dans les systèmes de requêtes fondés sur un des formalismes à base de graphes. Les auteurs proposent de différencier les notions de réponse par sous-graphes de la base de celles de réponses par création de graphes réponses. Ils proposent différents éléments de formalisation permettant de mieux caractériser les solutions proposées

Session 3 – « Systèmes à base de connaissances ontologiques » animée par Fabien Gandon

Une approche ontologique pour formaliser la connaissance experte dans le modèle du contrôle de conformité en construction. (Anastasiya Yurchyshyna, Catherine Faron-Zucker, Isabelle Mirbel, Birahim Sall , Nhan Le Thanh et Alain Zarli). Cette présentation a proposé une approche pour l'assistance au contrôle semi-automatique de la conformité d'un projet de bâtiment par rapport à un ensemble de normes techniques. Le format XML de l'IFC est traduit en RDF en suivant une ontologie. Cette traduction pourrait par ailleurs faire l'objet d'un profil GRDDL. La bibliothèque des contraintes de ces normes est représentée sous forme d'un ensemble de requêtes SPARQL annotées en RDF. Ces annotations des requêtes permettent non seulement de piloter la vérification (notamment au travers d'heuristiques) mais aussi de documenter le résultat de la vérification et ainsi de produire un rapport intelligible.

Les ontologies du prototype LUISA, une architecture fondée sur des Web Services Sémantiques pour les ressources de formation. (Monique Grandbastien, Benjamin Huynh-Kim-Bang, Anne Monceaux). Cette contribution décrit l'utilisation de schéma du web sémantiques dans une architecture distribuée destinée à assister la gestion des ressources de formation et des cursus en ligne. Deux contextes d'application sont mis en contraste pour l'utilisation de ce système : (1) l'évaluation des besoins en formation chez un industriel (2) la gestion d'une formation académique en bureautique. L'article donne un retour sur la faisabilité, l'utilisation et la valeur ajoutée des technologies du web sémantique pour la gestion de ressources de formation. Il montre aussi l'intérêt et les difficultés des modèles partagés.

Session 4 – « Conception d'ontologies » animée par Florence Le Ber

Cette session était composée de deux présentations. La première, présentée par Isabelle Mirbel, portait sur la construction d'une ontologie à l'usage des communautés du logiciel libre. Cette ontologie se focalise sur la description de la communauté de pratiques constituée autour d'un logiciel, et s'appuie sur - et complète - des ontologies développées par ailleurs. La discussion a porté sur la définition des rôles et des interactions entre les membres d'une telle communauté, ainsi que sur l'exploitation des forums existants à l'aide de l'ontologie construite. La deuxième présentation, par Thibaud Mondary, portait sur la comparaison de méthodes de construction d'ontologie à partir de textes. A partir de ces comparaisons, il précise les différentes étapes de la phase de conceptualisation menant à la construction d'une ontologie. Le débat qui a suivi a souligné l'intérêt de cette comparaison, tout en proposant d'affiner les notions mises en jeu.

Session 5 – « Web sémantique et web 2.0 » animée par Raphaël Troncy

Cette session a fait écho à l'atelier « IC 2.0 - Vers une ingénierie sociale des connaissances » (<http://apassant.net/home/2008/05/ic/>) qui s'était tenu la veille de la conférence en se demandant dans quelle mesure les usages du Web 2.0 font évoluer les pratiques d'IC.

A. Passant et P. Laublet, dans « Ontologies pour le Web 2.0 », présentent le modèle MOAT, une ontologie simple qui définit le sens d'un tag d'une folksonomie. Ce modèle est ensuite combiné avec l'ontologie SIOC pour annoter sémantiquement les contenus des billets d'un blog. Une application de peuplement d'ontologies à partir d'un wiki sémantique est alors proposée, permettant une recherche sémantique des connaissances capitalisées dans le wiki.

B. Huynh-Kim-Bang et E. Dané, dans « Social bookmarking et tags structurés », proposent l'outil SemanticsScuttle, un outil de gestion sémantique des signets. Une syntaxe simple permet de structurer à la volée les tags décrivant les ressources. Un processus collaboratif permet en outre de décrire les signets au sein d'une communauté. Une évaluation de l'outil est en cours.

F. Limpens, F. Gandon et M. Buffa, dans « Rapprocher les ontologies et les folksonomies pour la gestion des connaissances partagées : un état de l'art », proposent finalement un état de l'art exhaustif sur les différentes approches permettant d'extraire des structures à partir de folksonomies, de guider le processus de tagging à l'aide d'ontologies, ou encore d'utiliser conjointement ces deux types de structures de connaissances. Les auteurs préconisent finalement un rapprochement naturel entre folksonomies et ontologies plutôt qu'une opposition: l'usage d'une folksonomie a tendance à mettre en avant des structures récurrentes que l'on peut formaliser, tandis que la maintenance d'une ontologie peut se nourrir des folksonomies associées.

Session 6 – « Extraction de connaissances à partir des textes » animée par Nathalie Aussenac

Cette session a réuni deux articles présentés par des chercheurs du LIPN, sur des approches très différentes. Dans la première communication, les connaissances extraites des textes permettent de repérer des fragments de texte en fonction de leur rôle dans une argumentation. Amanda Bouffier et ses collègues ont réalisé un travail s'appuyant sur une analyse linguistique fine au niveau du discours et débouchant sur un « système d'analyse et de structuration automatique des guides de bonnes pratiques cliniques ». L'approche consiste à repérer et à traiter des fragments de texte en relation au sein de ces guides : des conditions (définissant un contexte) et des actions préconisées dans ce contexte. Le texte est alors représenté par un arbre. Dans le deuxième exposé, les connaissances extraites des textes sont formalisées en vue de générer des animations en 3D, et ce dans le domaine des accidents de la route. Nicolas Kamenoff propose un processus d'extraction basé sur une représentation des connaissances orientée objet afin de représenter des informations spatiales et temporelles sur les entités repérées dans les textes. Grâce à des raisonnements non monotones, le système est capable de corriger ou valider ce qui est extrait des textes, et de gérer ainsi des informations implicites dans les textes ou la polysémie des termes.

Session 7 – « Instrumentation de pratiques à base de connaissances » animée par Sylvie Després

La session intitulée « Instrumentation de pratiques à base de connaissances » qui s'est déroulée vendredi matin a donné lieu à deux présentations.

La première présentation « Cartographie sémantique auto-organisée d'un référentiel de connaissances partagées » a été faite par Jean Villerd du centre de recherche LGI2P de l'Ecole des Mines d'Alès. Elle était consacrée à la visualisation d'un référentiel de connaissances se voulant le plus intuitive possible afin de faciliter la navigation. La solution proposée consiste en une approche cartographique dynamique fondée sur l'application d'un jeu de forces contrôlées associées à des lentilles. Les questions ont essentiellement porté sur la notion de lentilles et ont mentionné la complémentarité avec le langage Fresnel établi pour la visualisation des graphes RDF

La seconde présentation « Formalisation des connaissances capitalisées en conception inventive » a été présentée par Alexis Bultey de LGECO de l'INSA de Strasbourg. Le sujet de la présentation concernait l'analyse des bases de connaissances associées à la théorie des problèmes inventifs (TRIZ), qui sont les plus utilisées dans la phase de résolution. Les principales questions étaient centrées sur l'évaluation des différentes ontologies construites.

Session 8 – « Traces, documents, inscriptions de connaissances » animée par Nathalie Souf

La session sur les traces, documents, inscriptions de connaissances comportait trois présentations. La première présentation (*Ingénierie des traces numériques d'interaction comme inscriptions de connaissances* par Julien Laflaquière, Yannick Prié et Alain Mille), présente le cadre conceptuel des Systèmes à base de Traces et explicite la notion de trace modélisée associant une collection d'observés avec un modèle explicite de cette collection. Le statut en tant qu'inscription de connaissances de telles traces est souligné et une démarche de modélisation de la trace est proposée. La seconde présentation proposée par Leila Yahiaoui (*Redocumentation des traces d'activité médiée informatiquement dans le cadre des transactions communicationnelles*, avec Yannick Prié et Zizette Boufaïda) appuie ses recherches sur des travaux présentés en première et en dernière présentation lors de la session. Elle situe un travail de documentation de l'activité médiée informatiquement qui repose sur des traces d'interaction de l'utilisateur avec le système, en présentant le cadre conceptuel des transactions communicationnelles issu des travaux de M. Zacklad. Les étapes de documentation de l'activité par les traces, de documentarisation et de redocumentation des traces sont ensuite exposées et illustrées par un scénario de redocumentation de traces issu de l'exploitation d'un outil de gestion de ressources bibliographiques sur internet. La dernière présentation (*Intelligence qualitative dans la gestion de risque : le rôle de la documentarisation et des Systèmes d'Organisation des Connaissances*) est proposée par Manuel Zacklad, Jean-Pierre Cahier, Aurélien Bénel, L'Hédi Zaher, Christophe Lejeune, et Chao Zhou. Il situe la gestion des situations à risque complexes, dans le cadre de ses recherches sur le web socio-sémantique, des notions de meta-sémiotiques et du protocole HyperTopic.

Session 9 – « Apprentissage et adaptation » animée par Alain Mille

Le premier travail présenté (*Familles génératrices de règles d'adaptation pour assister leur acquisition semi-automatique*, par Matthieu Tixier, Fadi Badra et Jean Lieber), très technique, ouvre des voies très intéressantes pour théoriser, formaliser et in fine outiller le processus d'acquisition des connaissances d'adaptation dans le processus même du raisonnement à partir de cas. Ces travaux résonnent avec d'autres travaux nationaux et internationaux tendant à intégrer les différentes phases du raisonnement à partir de cas comme un raisonnement unifié.

La notion de profils d'utilisateur n'est pas nouvelle pour la personnalisation de la navigation mais le second papier (*Vers une personnalisation de la navigation par l'apprentissage de profils utilisateurs* par Khaled Khelif, Yassine Mrabet et Rose Dieng-Kuntz) a été l'occasion de présenter une réalisation originale d'un système de construction des éléments du profil à partir des fichiers logs liés à l'utilisation d'un portail.

Le dernier travail présenté (*Enrichissement d'ontologies : maintenance de la consistance et évaluation de la qualité* par Rim Djedidi et Marie-Aude Aufaure) propose une approche permettant de « tracer » l'évolution d'ontologies pour en faciliter la maintenance et mettre en évidence les choix qui ont mené à des évolutions apportées. Il s'agit aussi d'un environnement pour faciliter la résolution de contradictions en suggérant des opérations sur les ontologies qui peuvent les résoudre. La traçabilité est en soi une documentarisation du processus et offre de facto une aide potentielle très intéressante pour les opérations de maintenance voire d'exploitation.

Yannick Prié
LIRIS – UNIVERSITE CLAUDE BERNARD LYON 1
Président du comité de programme d'IC'2008