

Wikis sémantiques : le peuplement d'ontologies pour tous ?

Alexandre Passant, Philippe Laublet

LaLIC, Université Paris-Sorbonne, 28 rue Serpente, 75006 Paris
nom.prenom@paris4.sorbonne.fr

1 Du wiki au wiki sémantique

Popularisé avec la mouvance Web 2.0, le wiki est un outil dont les innovations, plus que technologiques, se situent dans les usages qu'il introduit et la philosophie associée. Les wikis permettent la mise en place de sites Web que quiconque peut faire évoluer aussi bien en terme de contenu que de structure, via un système d'édition ouvert et non hiérarchisé et un usage intensif d'hyperliens pour la navigation et la création de pages, tout en minimisant les pré-requis techniques. L'ouverture et le consensus prennent ainsi le dessus sur les approches de construction de sites que proposaient il y a peu les portails Internet, gérés par une autorité centrale et référente. C'est d'ailleurs cette philosophie qui fait que même si les wikis trouvent des nombreux usages sur le Web mais également en entreprise, leur adoption dans ce contexte est parfois problématique, ceux-ci risquant d'aller à l'encontre des structures hiérarchiques classiques. Un autre aspect important des wikis est la diversité des rôles qu'ils peuvent jouer en fonction des besoins et intérêts de la communauté qui se l'approprie : capitalisation de connaissance, gestion de projets...

Dans la continuité des wikis, et faisant le pont entre le Web 2.0 et le Web Sémantique, les wikis sémantiques ont vu le jour. Ils permettent la création d'annotations sémantiques à partir de wikis, généralement modélisées sous forme de graphes d'annotations rattachés aux pages. L'objectif est ainsi de produire des connaissances précises et exploitables par des agents logiciels permettant des recherches plus précises que celles des wikis classiques et facilitant l'interopérabilité avec d'autres applications. On distingue principalement (Buffa *et al.*, 2008) :

- les approches où les pages sont associées aux concepts et instances d'une ontologie permettant ainsi leur construction collaborative. (ex : Semantic MediaWiki, SemperWiki, UfoWiki, OntoWiki...);
- celles qui s'attachent à modéliser, structurer, enrichir et utiliser des assertions au sujet des meta-données des pages (ex : SweetWiki, IkeWiki...).

La première catégorie de wiki peut permettre, par exemple, de répondre à des requêtes telles que "Ou se situe le LaLIC ?" et la seconde "Quelles sont les pages contribuées par untel les 10 derniers jours" alors que certains wikis combinent les deux approches (UfoWiki, IkeWiki...).

2 Peuplement d'ontologies avec un wiki sémantique

En matière d'Ingénierie des Connaissances (IC), les wikis sémantiques du premier type (i.e. permettant de créer des annotations vis-à-vis du contenu des pages) nous paraissent intéressants à étudier dans la mesure où leurs objectifs se rapprochent des outils de peuplement d'ontologies. Alors que cette démarche est en général complexe et réservée à un public restreint, ces wikis permettent de créer de nouvelles instances à partir de pages wiki tout en conservant les principes de ces outils et leur flexibilité.

Semantic MediaWiki (SMW) (Krötzsch *et al.*, 2006) est sans doute le wiki sémantique dont la démarche utilisateur se rapproche le plus de la philosophie wiki. La modélisation d'assertions RDF s'effectue par une extension de la syntaxe wiki, rendant le processus accessible aux habitués de celle-ci. On va ainsi écrire `Le LaLIC est basé à [[se_situe_a::Paris]]` pour produire les annotations RDF associées. SMW permet d'utiliser des ontologies pré-définies pour modéliser ces annotations, mais celles-ci peuvent également être implicites et émerger des pratiques utilisateurs. Malgré une simplicité accrue, cette ouverture peut nuire à la qualité des annotations sémantiques produites, rapidement confrontées à des problèmes d'hétérogénéité (`se_situe_en` et `se_situe_a` pour identifier la même relation).

D'autres outils assistent (ou restreignent) l'utilisateur lors de la production d'annotations. Ceci permet de s'assurer que les données respectent les modèles ontologiques sous-jacents et que les bases de connaissances produites sont ainsi cohérentes, les requêtes pouvant ensuite tirer bénéfice de cette cohérence globale. Ainsi, IkeWiki (Schaffert, 2006) suggère des relations à partir des ontologies introduites dans le système lors de la pose de liens entre pages, les méta-données des pages étant modélisées séparément. Kaukolu (Kiesel, 2006) permet via un système d'alias d'utiliser du texte classique pour modéliser les assertions sans avoir à appréhender une syntaxe particulière mais surtout sans être confronté directement aux URIs ou espaces de noms. ACEWiki (Kuhn, 2008) utilise quand à lui ACE, langue naturelle contrôlée (anglais) pour permettre l'édition des pages et la production d'annotations liées. Ici, l'ontologie, et non seulement les instances, peut évoluer via le wiki en fonction des annotations, et un raisonneur permet de vérifier la consistance de l'ontologie et des instances créées.

Une autre approche pour valider la qualité des assertions est d'utiliser des formulaires pour créer celles-ci en fonction d'ontologies sous-jacentes. Dans OntoWiki (Auer *et al.*, 2007) par exemple, chaque page wiki est en réalité une vue virtuelle au dessus d'une base de connaissance associée à une ou plusieurs ontologies ce qui fait de l'outil un système efficace pour éditer collaborativement une base d'annotations RDF. Différents plug-ins de visualisation viennent ensuite offrir des possibilités avancées d'utilisation de ces annotations (vue calendrier ...). Dans notre approche avec l'outil UfoWiki (Pasant & Laublet, 2008), les administrateurs de wikis définissent des templates de pages ou chaque page correspond à une classe d'ontologie du domaine et les champs de formulaires sont associés à des propriétés et relations. Ces templates sont ensuite utilisés par les utilisateurs lors de la création de pages, et permettent la production automatique d'annotations RDF en fonction du contenu des pages. Ces annotations sont ensuite utilisées pour offrir des possibilités de requêtes et de navigation avancées au sein même du wiki.

3 Usage en entreprise

Nous avons pu expérimenter au sein d'EDF R&D l'utilisation de wikis sémantiques, en complément de l'utilisation de wikis classiques. Notre démarche a été motivée par le besoin de capitaliser des connaissances au sujet d'entreprises et de leurs domaines d'activités et permettre des requêtes que les wikis classiques ne peuvent satisfaire. Nous avons ainsi d'une part produit des ontologies dédiées, relativement légères, et d'autre part développé UfoWiki, un nouveau système de wiki sémantique avec une approche d'annotations basée sur des formulaires, et capitalisant également les méta-données des pages avec SIOC. L'utilisation de formulaires, définis par l'administrateur comme évoqué précédemment, nous a semblé un bon compromis entre les approches souples à la SMW et le tout-formulaire à la OntoWiki. A chaque page créée par les utilisateurs finaux correspond donc au moins une instance d'ontologie du domaine et un certain nombre de propriétés et relations associées à cette instance. A minima, car si une relation pointe vers une instance qui n'existe pas, celle-ci sera immédiatement créée et ses différentes propriétés pourront par la suite être alimentées au sein d'une nouvelle page wiki, si besoin par d'autres utilisateurs.

L'utilisation de ce prototype, auprès d'une vingtaine d'utilisateurs qui pour la plupart étaient novices avec les wikis s'est avérée encourageante. Plus de 200 instances ont été créées, fortement interconnectées, notamment à travers des objets transverses comme leur secteur d'activité. Ce peuplement d'ontologies, et la création de la base de connaissance associée nous a donc permis de proposer des services avancés. Par exemple, pour chaque entreprise, une requête intégrée au sein même de la page wiki associée va interroger la base de connaissance pour lister l'ensemble des entreprises évoluant dans le même secteur d'activité ou un secteur plus spécifique.

4 Discussion et conclusion

A travers toutes ces approches, il convient donc de se poser la question du rapport simplicité d'utilisation / qualité des assertions (en terme de modèle) afin de trouver le juste milieu entre respect de la philosophie Wiki et IC. Si notre approche à l'avantage d'offrir des informations de qualité, elle contraint l'utilisateur en terme d'annotations puisque les relations utilisables sont définies, certes démocratiquement, par les administrateurs du wiki. On peut imaginer laisser les utilisateurs eux-mêmes faire évoluer les formulaires, voire les ontologies sous-jacentes, mais le problème d'hétérogénéité peut à nouveau se poser. Il est cependant probable qu'une sémantique émergente apparaisse, comme (Mika, 2005) le constate avec l'émergence d'ontologies à partir de folksonomies. Cependant, cette émergence ne peut avoir lieu que lorsqu'une masse critique d'utilisateur est atteinte, les premiers utilisateurs devant donc faire face à ces problèmes avant qu'un consensus éventuel ne prenne place. Ceci reste problématique dans un contexte d'entreprise où l'accès à une information de qualité est primordial.

Malgré tout, un aspect important à considérer est la manière dont ces wikis font évoluer les approches de constitution de bases de connaissances : ces processus jusque là complexes et réservés à un nombre restreint d'experts sont ainsi rendus disponibles à quiconque souhaite y apporter sa contribution. Sans nier les problèmes posés qu'il

sera nécessaire d'étudier (conflits d'édition, cohérence des instances par rapport aux contraintes définies dans les modèles, personnalisation de l'information...), on peut alors imaginer des ontologies que tout le monde peut peupler et faire évoluer rapidement, bénéficiant ainsi du principe d'édition participative en masse. Nous avons pu constater ce phénomène au sein d'EDF mais aussi sur le Web avec l'exemple du portail semanticweb.org, basé sur SMW, qui permet aux activistes du Web Sémantique de peupler une base de connaissance générale relative aux personnes, projets et langages du domaine et dont l'évolution n'aurait pu être aussi réactive avec une approche cloisonnée. De plus, ce processus de constitution et d'évolution des ressources permet de bénéficier des principes d'*architecture of participation* : l'utilisation du wiki par un ou un groupe d'utilisateur dans l'objectif de satisfaire des objectifs personnels (création d'une instance) va permettre à terme à l'ensemble de la communauté de bénéficier de ces évolutions (lien avec d'autres instances et élaboration du graphe global de connaissances), par un système, sans le qualifier de viral, que l'on nommera expansif.

A la question posée en titre de cet article, nous répondrons donc majoritairement oui, puisque, via des usages relativement simples, les wikis sémantiques peuvent permettre la production massive, ouverte et collaborative d'instances d'ontologies de domaine, et leur évolution. Cependant, il faut garder à l'esprit qu'en fonction de l'outil, le peuplement peut s'avérer plus ou moins complexe, et plus ou moins cohérent, mais également que de nouvelles problématiques se posent telles que celles évoquées précédemment.

Références

- AUER S., DIETZOLD S., LEHMANN J. & RIECHERT T. (2007). Ontowiki : A tool for social, semantic collaboration. In *Workshop on Social and Collaborative Construction of Structured Knowledge (CKC2007)*.
- BUFFA M., GANDON F. L., ERETEO G., SANDER P. & FARON C. (2008). Sweetwiki : A semantic wiki. *Journal of Web Semantics*, 6(1), 84–97.
- KIESEL M. (2006). Kaukolu : Hub of the semantic corporate intranet. In M. VÖLKE & S. SCHAFFERT, Eds., *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics*.
- KRÖTZSCH M., VRANDECIC D. & VÖLKE M. (2006). Semantic mediawiki. In *Proceedings of the 5th International Semantic Web Conference, ISWC 2006, Athens, GA, USA, November 5-9, 2006*, volume 4273 of *Lecture Notes in Computer Science*, p. 935–942 : Springer.
- KUHN T. (2008). AceWiki : Collaborative Ontology Management in Controlled Natural Language. In *Proceedings of the 3rd Semantic Wiki Workshop (SemWiki2008)*, Tenerife, Spain.
- MIKA P. (2005). Ontologies are us : A unified model of social networks and semantics. In *Proceedings of the 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10*.
- PASSANT A. & LAUBLET P. (2008). Towards an Interlinked Semantic Wiki Farm. In *Proceedings of the 3rd Semantic Wiki Workshop (SemWiki2008)*, Tenerife, Spain.
- SCHAFFERT S. (2006). Ikewiki : A semantic wiki for collaborative knowledge management. In *1st International Workshop on Semantic Technologies in Collaborative Applications (STICA'06)*, Manchester, UK.