

Les moteurs de wikis sémantiques : un état de l'art

Thomas Meilender^{*,**}, Nicolas Jay^{**}, Jean Lieber^{**}, Fabien Palomares^{*}

^{*}A2ZI - 61 ter rue de Saint-Mihiel - 55200 Commercy
{prénom.nom}@a2zi.fr,

^{**}UHP-Nancy 1 – LORIA (UMR 7503 CNRS-INPL-INRIA-Nancy 2-UHP)
{prénom.nom}@loria.fr

Résumé. Cet article est un état de l'art sur les moteurs de wiki sémantique, en particulier sur leur utilisation des technologies du Web sémantique. Les principales notions liées aux wikis sémantiques sont d'abord présentées. Ensuite, plusieurs projets actifs de moteurs de wiki sont comparés selon différents points de vue. Finalement, des recommandations sont données pour le choix d'un moteur de wiki. En conclusion, les auteurs s'interrogent sur les perspectives des wikis sémantiques telles que la faible interopérabilité de certains moteurs.

1 Introduction

Imagines en 1995 par Ward Cunningham, les wikis sont des sites Web permettant la création et l'édition collaborative de contenus de manière simple (Leuf et Cunningham, 2001), dont l'exemple plus populaire est Wikipédia. En parallèle, le développement du Web sémantique a ouvert de nouvelles perspectives : on cherche à créer un Web intelligent où les données sont transformées en connaissances exploitables par des machines. Ainsi sont nés les wikis sémantiques, dont la particularité consiste à formaliser le contenu des articles et leurs relations. De nombreux moteurs de wiki sémantique ont vu le jour.

Cet article compare ces différents moteurs, en particulier au niveau de l'emploi des technologies du Web sémantique. La section 2 définit le concept de wiki sémantique et donne des applications concrètes. La section 3 présente les principaux moteurs de wiki sémantiques actifs et la section 4 les compare. Enfin, la section 5 discute les résultats mis en évidence.

2 Qu'est-ce qu'un wiki sémantique ?

Les wikis traditionnels reposent généralement sur un ensemble de pages éditables, organisées en catégories et reliées par des liens hypertextes. Ces systèmes se caractérisent par leur simplicité d'utilisation et leur interactivité. Les wikis sont créés et maintenus grâce à des systèmes de gestion de contenus, les moteurs de wiki. Ceux-ci proposent un format de données spécifique, les *wikitexts*, pour permettre la structuration, la mise en pages et les liens entre les articles. Toutefois, une limite de l'exploitation des wikis s'illustre par le questionnement des données contenues dans ces pages. La recherche se limite généralement à de la reconnaissance de chaînes de caractères, sans prendre en compte le sens des termes. Ce manque peut être comblé par l'intégration du Web sémantique, que (Berners-Lee et al., 2001) définissent



FIG. 1 – Différences de représentation entre un wiki (1) et un wiki sémantique (2)

comme une extension du Web qui vise à structurer les données en vue d'en représenter le sens. Les wikis sémantiques sont donc nés du rapprochement des wikis et du Web sémantique. Selon (Berners-Lee et Fischetti, 1999), un wiki sémantique est un wiki amélioré par l'utilisation des technologies du Web sémantique. Plus particulièrement, le contenu est organisé en pages éditables et indexables. Cependant, le wiki sémantique ne se limite pas au texte en langage naturel. Il formalise les ressources et les liens entre celles-ci comme l'illustre la figure 1. Comme mentionné dans (Buffa et al., 2007), on peut distinguer deux approches dans la conception des moteurs. La première, appelée *wikis for ontologies*, concerne le plus grand nombre des moteurs. Cette approche considère les pages comme des concepts ou des instances et les liens typés comme des propriétés. Il se dessine ainsi une ontologie dont la précision est renforcée par la catégorisation des concepts. Dans ce cas, la structure des données est généralement très souple et permet une grande liberté pour l'utilisateur mais ne garantit pas l'utilisabilité des ontologies résultantes. La deuxième approche, *ontologies for wikis*, suppose généralement une ontologie pré-existante. Selon cette approche, le but du moteur de wiki consiste à fournir les outils permettant son peuplement par l'ajout d'instances et parfois de classes. Cette approche apparente le wiki à un éditeur de métadonnées, permettant de peupler l'ontologie. Ainsi, les moteurs de wiki de cette catégorie sont le plus souvent destinés à des domaines spécifiques, plus facilement formalisables.

Les wikis sémantiques ont déjà été mis à l'épreuve dans différents domaines d'applications. Parmi ceux-ci, *Wikitaable* (Cordier et al., 2009) est un wiki sémantique utilisé dans la gestion de recettes de cuisine. Celles-ci sont annotées et rendues utilisables par un moteur de raisonnement à partir de cas, qui peut proposer des adaptations. Dans un autre contexte, celui d'EDF R&D (Passant et Laublet, 2008), un wiki sémantique est utilisé pour capitaliser les connaissances sur les entreprises et leurs secteurs d'activité et permet la classification des entreprises par des réponses à des requêtes complexes.

3 Présentation des moteurs de wiki sémantique actifs

Cette partie se concentre sur les moteurs disponibles en ligne et faisant l'objet de mises à jour régulières. Un panorama plus complet est disponible dans (Meilender et al., 2010).

Le projet **KiWI** (*Knowledge In Wiki*, ex-**IkeWiki**) (Schaffert, 2006) est l'illustration de l'approche *ontologies for wikis*. Il se présente comme un outil de gestion collaborative des connaissances. Il nécessite une ontologie pré-existante. La gestion des connaissances se fait grâce au *framework* Jena, qui sert de *RDFstore* et de moteur SPARQL. Le système propose une granularité avancée dans la gestion des utilisateurs, afin qu'il soit possible de confier à des experts des tâches plus complexes que celles confiées à des novices. Une autre fonction intéres-

sante est l'import de pages depuis *Wikipédia*, dont Ikewiki gère les annotations postérieures. Son interface utilise les technologies Ajax, qui permettent une grande interaction avec l'utilisateur. Deux éditeurs sont proposés : un pour les métadonnées comportant un mécanisme d'auto-complétion, et un WYSIWYG pour le contenu.

À l'opposé, **Semantic Mediawiki** (Völkel et al., 2006) est très représentatif de l'approche *wikis for ontologies*. C'est une extension de *MediaWiki*, le moteur utilisé par Wikipédia. Afin de conserver une syntaxe simple, il intègre l'édition des triplets RDF à son wikitext et permet ainsi de créer facilement des liens typés et des attributs. Citons en particulier l'extension *Halo*¹ qui propose des formulaires, de l'auto-complétion, un éditeur WYSIWYG, l'intégration de fichiers audios ou vidéos mais aussi l'intégration d'un point d'accès SPARQL. Intégrant cette extension ainsi que de nombreuses extensions sous licence GPL, **SMW+**² est une version commerciale augmentée de Semantic Mediawiki. **SWOOKI** (Rahhal et al., 2009), un autre dérivé de Semantic MediaWiki, propose de résoudre les problèmes des wikis liés à la centralisation des données en intégrant une architecture pair-à-pair.

AceWiki (Kuhn, 2009) simplifie l'édition des connaissances en proposant une alternative aux triplets RDF : il utilise le langage ACE (*Attempto Controlled English*), un fragment de l'anglais, qui peut être facilement analysé grammaticalement et traduit en OWL ou en SWRL³ afin de peupler l'ontologie. Un éditeur prédictif guide l'utilisateur dans la construction des phrases et des tests de consistances sont effectués à chaque modification. De plus, ACE est également utilisé comme langage de requêtes.

D'une manière différente, **OntoWiki** (Auer et al., 2006) se fixe également comme principal but de faciliter l'acquisition et la présentation des données. Pour ce faire, il propose d'efficaces formulaires et intègre les triplets RDF directement dans le texte en wikitext selon une syntaxe appropriée. En complément, il dispose de fonctions intéressantes telles que la navigation à facettes, des statistiques avancées sur les utilisateurs ou l'intégration de services Web. De plus, il propose une gestion des communautés.

Des wikis sémantiques dédiés à certains domaines ont été créés comme **BOWiki** (Bacher et al., 2008) pour la biomédecine et **SWiM** (Lange, 2008) pour les mathématiques. BOWiki contrôle la consistance de l'ontologie à chaque ajout et rejette l'unité de connaissance le cas échéant. Il se pose ainsi en complément du travail de l'expert. Pour sa part, SWiM implémente IkeWiki en ajoutant un support du langage OpenMath, un langage XML de représentation des équations afin de faire de SWiM un outil capable de gérer l'édition collaborative de définitions de symboles mathématiques et leur apparence.

Certains moteurs proposent une représentation des connaissances différent. Assez peu documentée, **Subleme**⁴ est une application minimaliste pour le partage d'informations. Tout y est considéré comme des pages (documents texte, utilisateurs, etc.) liées par des liens typés. Il n'utilise pas les notions de classes ou d'instances mais gère des relations du type « est instance de ». Bien plus complet, **TaOPis** (Schatten et al., 2009) propose l'édition d'ontologies sur un modèle orienté objet. Pour ce faire, il utilise la logique décidable F-logic et intègre un moteur d'inférences Flora-2, qui utilise une syntaxe spécifique pour les requêtes.

1. <http://www.projecthalo.com/>

2. <http://wiki.ontoprise.de>

3. <http://www.w3.org/Submission/SWRL/>

4. <http://code.google.com/p/subleme/>

Les moteurs de wikis sémantiques : un état de l'art

Quelques systèmes commerciaux sont apparus, dont **Knoodl**⁵ et **Wikidmart**⁶ sont des exemples. Knoodl se présente comme un outil de développement de bases de connaissances. L'idée de communautés est mise en avant : chacune d'entre elles dispose de son propre wiki, intégrant son vocabulaire spécifique. Pour sa part, Wikidmart entre dans le cadre d'une application Web sémantique plus lourde dont le but est de proposer des services pour la gestion commerciale et humaine des entreprises. Dans les deux cas, il semble que les fonctions les plus intéressantes soient payantes, et les sociétés restent discrètes sur les moyens mis en œuvre.

4 Comparaison des projets actifs

Pour comparer les différents projets actifs, nous nous sommes appuyés sur plusieurs critères, répartis dans quatre catégories. La première catégorie concerne l'approche conceptuelle où nous cherchons à savoir si les concepteurs se fondent sur une approche *wikis for ontologies* ou une approche *ontologies for wikis*. La deuxième catégorie permet de juger de l'emploi des technologies du Web sémantique alors que la troisième catégorie traite de l'utilisabilité du système. Enfin, la dernière partie regroupe des informations complémentaires sur les projets.

Le tableau de la figure 2 confronte les systèmes avec les critères précédemment énoncés. Il est intéressant de noter que, d'un point de vue global, SMW+, KiWI et Knoodl sont les plus complets. On remarque également que les principales fonctions attendues d'un wiki sont remplies par la plupart des systèmes et que l'implémentation des standards du Web sémantique est en progression. Toutefois, l'import d'ontologies n'est proposé que par Knoodl, dont les possibilités exactes ne sont pas documentées. Toutefois certains systèmes intéressants comme AceWiki ou TaOPis souffrent du caractère expérimental de leur intégration dans ce type de comparaison.

5 Conclusion

Nous avons présenté un état de l'art sur les moteurs de wiki sémantique. Définis comme des outils de gestion des connaissances, ils peuvent devenir l'épine dorsale du Web sémantique et sont déjà mis en œuvre dans des applications concrètes. Désormais, plusieurs défis s'ouvrent : la complexité de l'annotation (du point de vue de l'utilisateur non-expert), l'interopérabilité, la migration de l'existant, la centralisation et l'intégration des communautés.

Les systèmes que nous avons présentés proposent des éléments de réponse à ces problématiques. L'adaptation du wikitext dans Semantic MediaWiki et OntoWiki, les formulaires et l'auto-complétion dans IkeWiki/KiWI et le langage contrôlé dans AceWiki facilitent l'acquisition des connaissances. De plus, l'emploi des standards du web sémantique tels que OWL et SPARQL tend à se généraliser. De plus, IkeWiki propose l'import de pages depuis Wikipédia, ce qui constitue une avancée pour la reprise des données.

Toutefois, quelques failles demeurent. On regrettera entre autres l'impossibilité pour la plupart des systèmes d'importer des ontologies qui nuit à l'intégration avec d'autres systèmes sémantiques. De même, les wikitexts et les langages de requêtes ne sont pas uniformisés.

Pour finir, on constate que l'approche *wikis for ontologies* ne garantit pas l'homogénéité de la base de connaissances. À l'heure actuelle, la seule solution semble être le contrôle manuel des ajouts par un ingénieur des connaissances.

5. <http://knoodl.com>

6. <http://www.zagile.com/products/wikidmart.html>

Approche		Outils du Web sémantique										Fonctions wiki					Statut		
Wikipédia	Wikipédia for ontology	Langage de représentation	Export ontologie	Import ontologie	Système de requête	SPARQL endpoint	Moteur inférence	Stockage des données	Aspects-temps	Syntaxe	WYSIWYG	Forma-tiltre	Auto mode d'édition	Gestion version	Gestion utilisateurs	Language	Licence	Editeur	Version
•	•	OWL, SMWRL	•	•	Texte, requêtes avancées	•	Pellet	Texte	Dans le texte	ACE	•	•	Editeur prédictif	•	•	Java	LGPL	University of Zurich	alpha
	•	OWL	•	OBO format	Texte, requêtes avancées	•	Pellet	DB	Inconnu	wikitext	•	•	•	•	•	PHP	GPL	Max Planck Institute	stable
	•	RDF, OWL	•	•	Texte, SPARQL, requêtes avancées	•	Jena	RDFstore + DB	Séparées	wikitext	•	•	Auto-complétion	•	•	Ajax, Java	BSD	KOM European Project	stable
	•	Unknown	•	•	Texte, SPARQL, requêtes avancées	•	•	RDFstore	Dans le texte	wikitext	•	•	•	•	•	Java	Inconnu	Reveljox	stable
	•	OWL, cswarv, XML	•	•	Texte, SPARQL	•	Sesame	•	Dans le texte	wikitext	•	•	Auto-complétion	•	•	Java	LGPL	University of Wurzburg	beta
	•	RDF(S), OWL	•	•	Texte, SPARQL, requêtes avancées	•	POWEL	RDFstore	Séparées	U+know	•	•	Auto-complétion	•	•	PHP	GPL	University of Leipzig	stable
	•	RDF(S), OWL	•	•	Texte, SPARQL, requêtes avancées	•	•	DB	Dans le texte	wikitext	•	•	•	•	•	PHP	GPL	University Karlsruhe	stable
	•	RDF(S), OWL	•	•	Texte, SPARQL, requêtes avancées	•	KAON 2	DB, RDFstore	Dans le texte	wikitext	•	•	•	•	•	PHP	GPL	Onepages GmbH	stable
	•	RDF	•	•	Texte	•	•	HTML pages	Séparées	wikitext	•	•	•	•	•	Inconnu	GPL	Inconnu	Inconnu
	•	RDF(S), OWL	•	•	Texte, SPARQL, requêtes avancées	•	•	DB	Séparées	wikitext	•	•	•	•	•	Java, ASCL	GPL	KWARSK research group	alpha
	•	F-Logic	•	•	Texte, FLogic-2, avancées queries	•	FLogic-2	DB	Dans le texte	wikitext	•	•	•	•	•	Python, PHP	GPL	University of Zagreb	alpha
	•	Inconnu	•	•	Texte, SPARQL, requêtes avancées	•	•	DB	Dans le texte	wikitext	•	•	•	•	•	PHP	AGPL	Zaggle, inc	stable

FIG. 2 – Récapitulatif des moteurs de wiki et de leurs fonctionnalités.

Les moteurs de wikis sémantiques : un état de l'art

Remerciements

Les auteurs tiennent à remercier les relecteurs pour leurs remarques pertinentes.

Références

- Auer, S., S. Dietzold, et T. Riechert (2006). Ontowiki - a tool for social, semantic collaboration. In *5th Int. Semantic Web Conference, ISWC 2006, Athens, USA, November 5-9, 2006*, pp. 736–749. Springer.
- Bacher, J., R. Hoehndorf, et J. Kelso (2008). Bowiki : Ontology-based semantic wiki with abox reasoning. In *3rd Semantic Wiki Workshop*.
- Berners-Lee, T. et M. Fischetti (1999). *Weaving the Web : The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor* (1st ed.). Harper San Francisco.
- Berners-Lee, T., J. Hendler, et O. Lassila (2001). The semantic web. In *Scientific American Magazine*, pp. 29–37.
- Buffa, M., G. Erétéo, et F. L. Gandon (2007). Wiki et web sémantique. In F. Trichet (Ed.), *IC 2007 : 18e J. Francophones d'Ingénierie des connaissances 2007, Grenoble, France*, pp. 49–61. Cepadues.
- Cordier, A., J. Lieber, P. Molli, E. Nauer, H. Skaf-Molli, et Y. Toussaint (2009). Wikitaaable, un wiki sémantique utilisé comme un tableau noir dans un système de raisonnement à partir de cas textuel. In *17ème atelier de Raisonnement à Partir de Cas*.
- Kuhn, T. (2009). How controlled english can improve semantic wikis. In *the Fourth Workshop on Semantic Wikis, European Semantic Web Conference 2009, CEUR Workshop Proceedings, 2009*.
- Lange, C. (2008). Mathematical semantic markup in a wiki : The roles of symbols and notations. In *3rd Semantic Wiki Workshop*.
- Leuf, B. et W. Cunningham (2001). *The Wiki way : quick collaboration on the Web*. Addison-Wesley.
- Meilender, T., J. Lieber, N. Jay, et F. Palomares (2010). Les moteurs de wiki sémantique : un état de l'art. In *Rapport technique*.
- Passant, A. et P. Laublet (2008). Wikis sémantiques : le peuplement d'ontologies pour tous ? In *Atelier (IC 2.0) des 19e J. Francophones d'Ingénierie des connaissances 2008, Nancy, France*.
- Rahhal, C., H. Skaf-Molli, et P. Molli (2009). Swooki : Un Wiki Sémantique sur réseau Pair-à-Pair. *Ingénierie des Systèmes d'Information 14*(1).
- Schaffert, S. (2006). Ikwiki : A semantic wiki for collaborative knowledge management. In *15th IEEE Int. Workshops on Enabling Technologies : Infrastructures for Collaborative Enterprises (WETICE 2006), 26-28 June 2006, Manchester, United Kingdom*, pp. 388–396. IEEE Computer Society.
- Schatten, M., M. Čubrilo, et J. Ševa (2009). A semantic wiki system based on f-logic. In *19th Central European Conference on Information and Intelligent Systems, Varaždin, Croatia*.
- Völkel, M., M. Krötzsch, D. Vrandečić, H. Haller, et R. Studer (2006). Semantic wikipedia. In *15th int. conference on World Wide Web, WWW 2006, Edinburgh, Scotland, May 23-26, 2006*.

Summary

This paper is a survey about semantic wikis and the use of semantic Web technologies by wiki engines. First, it presents the main notions related to semantic wikis. Then, several active semantic wiki projects are compared, according to various viewpoints. Finally, recommendations are carried out to provide a guideline for the choice of semantic wiki engine, given a set of needed features. In conclusion, the paper argues about semantic wiki issues such as the weakness of some engine interoperability.