



**HAL**  
open science

## Social bookmarking et tags structurés

Benjamin Huynh-Kim-Bang, Eric Dané

► **To cite this version:**

Benjamin Huynh-Kim-Bang, Eric Dané. Social bookmarking et tags structurés. Ingénierie des Connaissances - IC 2008, Jun 2008, Nancy, France. pp.111-122. hal-00359418

**HAL Id: hal-00359418**

**<https://hal.science/hal-00359418>**

Submitted on 6 Feb 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Social bookmarking et tags structurés

Benjamin Huynh-Kim-Bang<sup>1</sup>, Eric Dané<sup>2</sup>

<sup>1</sup> LORIA (UMR 7503 CNRS-INPL-INRIA-Nancy 2-UHP)

Campus scientifique

54506 Vandœuvre-lès-Nancy

benjamin.huynh-kim-bang@loria.fr

<sup>2</sup> Université René Descartes

45, rue des Saints-Pères

75270 Paris

eric.dane@univ-paris5.fr

**Résumé** : Nous présentons dans cet article les enjeux du rapprochement entre les techniques du web participatif (Web 2.0) et du web sémantique. Nous expliquons ensuite les limites des outils actuels explorant cette piste mais qui, malgré l'emploi de tags, conservent une part de rigidité liée aux ontologies sous-jacentes employées. Nous décrivons finalement l'ajout à un outil de social bookmarking de trois fonctionnalités permettant la structuration progressive de tags par une communauté tout en respectant la souplesse des tags classiques.

**Mots-clés** : social bookmarking, tags, ontologie, web 2.0, web sémantique

## 1 Contexte : rapprochement du web sémantique et du web participatif

Le web sémantique et le web participatif (ou web 2.0) sont deux tendances qui se sont développées ces dernières années et qui peuvent sembler s'opposer en apparence.

Le web participatif décrit l'implication croissante des utilisateurs sur Internet. Les expressions "Web 2.0"<sup>1</sup> ou "Read/Write Web" apparurent en 2004 pour marquer le passage d'un réseau où quelques entités diffusaient de l'information à un réseau où de plus en plus d'individus participaient en donnant leurs opinions et en produisant des contenus. Les blogs ou les wikis sont les outils illustrant le mieux le succès<sup>2</sup> de cette tendance. L'expression Web 2.0 est aussi reliée au procédé de tagging, où la participation des internautes est mise à profit pour organiser des documents à l'aide de mots-clefs, les tags, employés spontanément (Marlow *et al.*, 2006). Néanmoins, la masse croissante des utilisateurs souleva des difficultés pour organiser les contributions. Le procédé de tagging permet aux systèmes informatiques de classer des ressources précédemment décrites à l'aide de tags en regroupant les ressources décrites par les mêmes tags. Malgré

---

<sup>1</sup>O'Reilly : <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

<sup>2</sup>"Time's Person of the Year : You" : <http://www.time.com/time/magazine/article/0,9171,1569514,00.html>

ses avantages, le procédé de tagging rencontra des limites sur au moins trois points (Guy & Tonkin, 2006; Sen *et al.*, 2006; Passant, 2007) :

- problème typographique : les tags fonctionnent par rapprochement orthographique. Si un tag contient une faute d'orthographe ou si un utilisateur emploie la forme au pluriel et un autre utilisateur la forme singulière alors le système sera incapable de faire le rapprochement.
- problème d'ambiguïté : si un utilisateur emploie le tag "edf" les systèmes ne pourront déterminer directement si ce tag signifie "Electricité De France" ou "Equipe De France" ?
- problème de structure à plat : les tags ne peuvent être rapprochés que s'ils décrivent des ressources communes. Or comme le procédé de tagging à la conséquence de produire de nombreux tags, les utilisateurs désirent à partir d'un certains nombres de tags les organiser directement entre eux mais ont peu de moyens pour le faire.

Ces limites sont d'autant plus contraignantes que la seconde tendance, celle du web sémantique, offre des outils pour les dépasser.

Le web sémantique est basé sur les spécifications formelles de représentations partagées d'un domaine appelées ontologies (Gruber, 1993). Une fois les concepts, relations et objets définis de manière commune par les humains et compréhensible pour les systèmes informatiques, ces derniers peuvent raisonner dessus par inférence pour fournir des réponses puis services avancés (Berners-Lee *et al.*, 2001). Le peu d'applications à grande échelle supportées par des ontologies et des moteurs sémantiques réside au moins dans deux difficultés relevées par Downes (Downes, 2004) ou Shirky<sup>3</sup> :

- difficulté pour obtenir une vision commune : le processus classique pour réaliser une ontologie consiste à réunir les experts d'un domaine et de leur demander d'explicitier et rendre cohérent l'ensemble des concepts du domaine. La difficulté réside dans le fait de se mettre d'accord sur une vision exactement identique lorsque les concepts deviennent de plus en plus nombreux puis que le nombre d'utilisateurs augmente.
- difficulté pour annoter les ressources : pour pouvoir appliquer les raisonnements sur des ressources, les systèmes doivent avoir une description des ressources basée sur les ontologies réalisées précédemment. Or l'annotation détaillée de ressources est une tâche qui ne peut pour l'instant être totalement automatisé, qui reste coûteuse en temps et qui n'est pas forcément parfaite du fait des biais d'interprétation possibles chez les annotateurs.

Plusieurs oppositions distinguent la participation du web participatif et la rigueur du web sémantique. Cependant, leurs défauts se complétant, il existe un enjeu important à intégrer ces deux approches dans une voie intermédiaire et notamment en faisant fonctionner ensemble les tags et les ontologies. Plusieurs travaux et applications explorent aujourd'hui cette voie.

Des applications grands publics proposent de structurer des descripteurs de manière hiérarchique. F-spot<sup>4</sup> facilite ainsi la classification de photographies tandis que Wordpress<sup>5</sup> permet la classification de billets de blogs. Ces applications sont néanmoins

---

<sup>3</sup>Ontology is Overrated : [http://www.shirky.com/writings/ontology\\_overnated.html](http://www.shirky.com/writings/ontology_overnated.html)

<sup>4</sup>F-spot : <http://f-spot.org>

<sup>5</sup>Wordpress, application de gestion de blogs : <http://wordpress.org/>

limitées car elles visent principalement des contextes mono-utilisateurs où une seule personne crée et organise les descripteurs entre elles.

Un premier ensemble de travaux de recherche portent sur la mise en place d'une ontologie de tags afin de faciliter leur réutilisation par plusieurs systèmes (Gruber, 2007). Dans ce cas-là, la conceptualisation ne porte pas sur le sujets des tags (médecine, aviation, etc) mais sur le procédé de tagging lui-même (un Utilisateur taggue un Objet avec des Tags). Cette approche ne prend pas en compte d'éventuelles relations explicites entre les tags.

Un second ensemble s'intéresse à la création de relations entre les tags de plusieurs utilisateurs. Ces travaux portent sur des wikis sémantiques (Buffa & Gandon, 2006), des blogs sémantiques (Passant, 2007), l'annotation de photographies d'art (Trichet & Nizon, 2007) ou de documents pédagogiques (Bateman *et al.*, 2006).

## **2 Problématique : permettre la structuration sans perdre la participation**

Réunir la puissance sémantique et la souplesse nécessaire à la participation des utilisateurs est un enjeu important de ce que certains appellent le "Semantic Web 2.0" (Greaves, 2007) ou encore le "Web 3.0"<sup>6</sup>. Ce rapprochement semble d'autant plus possible que l'observation de l'usage des tags montre leur stabilisation (Halpin *et al.*, 2007). A nos yeux la plus grande difficulté porte sur les contraintes imposées aux utilisateurs et nécessaires à la structuration des tags. La participation des utilisateurs est un phénomène fragile, facile à voir disparaître si les coûts ergonomiques d'une application sont trop importants.

Or les derniers travaux cités ci-dessus, nous semblent conserver des ontologies des lourdeurs pouvant affaiblir la participation des utilisateurs. Les contraintes portent sur :

- une vision commune du domaine : les travaux cités ci-dessus conservent toujours sous-jacent une ontologie représentant l'unique modélisation du domaine employée par le système. Même si les tags permettent de rendre flexibles les représentations des différents utilisateurs, ceux-ci doivent toujours à un moment rattacher leurs visions à l'unique ontologie. En conservant cette ontologie sous-jacente, ces systèmes conservent la lourdeur de création d'une ontologie et alourdissent le travail des usagers devant se rattacher à l'ontologie.
- la cohérence entre les tags et leurs relations : une ontologie est une structure devant exprimer de manière explicite et cohérente les concepts d'un domaine. Les ontologies ne peuvent admettre de contradiction interne afin de permettre les raisonnements automatisés. En reposant sur des ontologies, les systèmes cités ci-dessus doivent alors laisser des experts manier l'ontologie et interdire la création de certaines annotations aux utilisateurs (par exemple dans OSIRIS (Trichet & Nizon, 2007)).

Pour l'instant, il est difficile d'évaluer l'usage par les utilisateurs des travaux cités ci-dessus. D'un côté, le cadre encore rigide de ces annotations est peut-être possible dans

---

<sup>6</sup>Appel conférence OntoLex08 : [www.lrec-conf.org/lrec2008/IMG/ws/OntoLex08.pdf](http://www.lrec-conf.org/lrec2008/IMG/ws/OntoLex08.pdf)

certains contextes comme dans une entreprise aux incitations parfois fortes ou pour un blogger prenant le temps d'organiser ses propres billets de blogs. D'un autre côté, ces outils nous semblent encore trop rigides dans d'autres contextes comme pour une communauté d'utilisateurs aux liens souples et bénévoles ou pour l'annotation d'adresses web qui est généralement une activité plus fréquente et répétitive que l'annotation de ses propres billets de blogs.

### 3 Choix d'implémentation

Nous proposons une approche pour l'annotation collaborative soutenant la participation des utilisateurs. Nos travaux s'inscrivent dans le contexte du partage de signets (ou social bookmarking) que nous détaillerons dans la prochaine section.

Notre approche s'inspire des procédés basés sur les tags et les ontologies. Des tags nous conservent la souplesse, la liberté d'emploi et le phénomène d'émergence qui leur est lié. En cas de compromis à trouver, nous favoriserons toujours la conservation de la souplesse des tags à la recherche d'inférence puissante. En effet, un mécanisme d'inférence puissant est peu utile si aucun utilisateur n'a annoté de ressources avec des métadonnées supportant ces inférences. Pour ceci, nous conservons l'aspect privé des tags qui signifie que le créateur d'un tag est constamment libre de lui donner le sens qu'il désire, sans avoir à le relier à la communauté ou même à être cohérent avec lui-même. Des ontologies nous conservons l'intérêt d'inférences simples pour améliorer les recherches. Nous ne désirons pas conserver des ontologies le caractère stable mais plutôt son processus de construction négociée (Karapiperis & Apostolou, 2006; Braun *et al.*, 2007).

L'équilibre à trouver crée un nouveau type d'outil explorant la frontière entre intelligence collective (où les comportements sont statistiquement agrégés) et intelligence collaborative (où les acteurs discutent et collaborent). Ceci nécessite une liberté participative où les utilisateurs se voient proposer une gamme d'actions<sup>7</sup> différentes en fonction de leur niveau d'implication dans le groupe, mais en pouvant à tout moment varier ses niveaux. Cela nécessite aussi la possibilité pour chaque utilisateur de structurer ou re-structurer son organisation des données.

Enfin, notre cadre théorique se rapproche de celui du Web Socio-Sémantique (Zacklad, 2005), dans le sens où nous considérons les ontologies plus en tant qu'objets co-construits dans un cadre d'interactions humaines et moins pour leur seule puissance formelle de raisonnements.

### 4 Contexte d'implémentation

Nos travaux sont basés sur Scuttle<sup>8</sup>, outil de social bookmarking sous licence libre, et clone de l'application web de Del.icio.us. Grâce à sa licence libre, Scuttle est déployé sur une centaine de sites web en quelques années. Ceci signifie qu'il a été rodé d'un

---

<sup>7</sup>Ross Mayfield, "Power law of participation" : [http://ross.typepad.com/blog/2006/04/power\\_law\\_of\\_pa.html](http://ross.typepad.com/blog/2006/04/power_law_of_pa.html)

<sup>8</sup><http://sourceforge.net/projects/scuttle/>

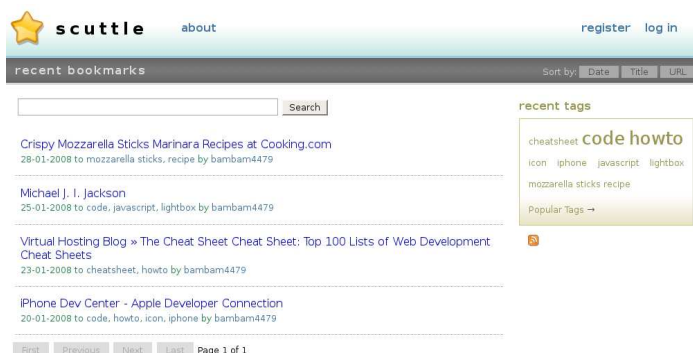


FIG. 1 – Visualisation des signets et des tags (Copie d'écran Scuttle)

point de vue ergonomique et que nous pouvons concentrer nos efforts sur de nouvelles fonctionnalités. Sa ressemblance avec le site Del.icio.us nous permettra éventuellement plus tard de comparer les usages des deux outils. Ceci signifie aussi que Scuttle possède déjà les fonctionnalités classiques liées au Web 2.0 comme les flux RSS, une API permettant à une autre application d'interroger Scuttle ou encore une extension s'installant dans le navigateur des utilisateurs pour accélérer l'ajout d'un signet.

Le fonctionnement de Scuttle est assez classique par rapport à d'autres outils de social bookmarking. Après s'être enregistré sur un site Scuttle, l'utilisateur peut ajouter un signet qui se compose d'une URL, d'un titre, d'une description et d'une zone de tags. La zone de tag est un champ d'une ligne où l'utilisateur ajoute les tags les uns à la suite des autres, sans contrainte autre que de les séparer par des virgules. Pour favoriser la ré-utilisation, le système lui rappelle en bas de la page les tags qu'il emploie le plus souvent.

L'utilisateur peut ensuite retrouver ce signet sur son espace personnel ou sur la page principal de Scuttle qui liste les derniers signets ajoutés. La facilité d'emploi des tags conduit rapidement à une grande quantité de tags. La navigation parmi eux s'effectue à l'aide de nuage de tags qui sont des listes de tags dont la taille de caractère est proportionnelle au nombre de signets qu'ils décrivent. Dans Scuttle, les pages listent généralement les signets sur la gauche et les tags dans des boîtes sur la droite. La figure 1 montre ainsi sur la droite les tags les plus récents.

Nos développements consistent à ajouter à Scuttle des fonctionnalités permettant aux utilisateurs de structurer les tags. Notre projet s'appelle SemanticScuttle<sup>9</sup>.

## 5 Première fonctionnalité : tags structurés

Cette fonctionnalité a pour but de dépasser les limites des tags référencés en introduction.

<sup>9</sup>SemanticScuttle est un projet libre, téléchargeable et documenté sur : <http://sourceforge.net/projects/semanticscuttle/>

Pour résoudre ces difficultés, les travaux cités précédemment relient les tags à des ontologies comme WordNet pour (Bateman *et al.*, 2006). Comme nous l'avons expliqué, ces travaux se retrouvent, d'un point de vue ergonomique, handicapés par l'emploi d'ontologie. De notre côté, nous proposons de structurer les tags tout en conservant la forme et l'esprit des tags.

Pour cela, nous laissons la possibilité aux utilisateurs de créer des relations entre tags et non entre tags et concepts. Ainsi, notre système ne possède pas de modèle caché relié à une quelconque ontologie réalisée a priori. Les relations sont de deux types : inclusion et synonymie. Ces relations se retrouvent sous d'autres noms dans la littérature comme "SubClassOf" et "SeeAlso" (Buffa & Gandon, 2006). Des outils comme OSIRIS (Trichet & Nizon, 2007) proposent des types de relations plus affinés. Cependant ces types plus complexes demandent des requêtes complexes que nous ne pensons pas les utilisateurs prêts à réaliser (de la même manière que les utilisateurs emploient très rarement le mode "recherches avancées" de Google).

Les tags et les relations entre eux appartiennent à leur créateur qui peut les manipuler librement et indépendamment des autres utilisateurs. De ce point de vue, les utilisateurs ne manipulent pas des objets communs mais des objets privés. En revanche, les utilisateurs peuvent observer et profiter des tags et structures de tags créés par les autres utilisateurs.

Lorsqu'un utilisateur crée des tags et des relations entre les tags, il ne doit jamais être contraint ou recevoir de message d'erreur. Le système doit ainsi supporter les incohérences de structure pour un utilisateur et entre utilisateurs lorsqu'il agrège les tags et les relations. Nous verrons comment des inférences simples supportent ces incohérences.

Sur la forme, les relations entre tags doivent, comme les tags, pouvoir être créées à la volée au moment de l'annotation des URL. Des pratiques cherchant à organiser les tags à leur création, sont effectuées de manière intuitive par les utilisateurs d'outils de social bookmarking qui par exemple ajoutent des " " devant certains tags pour qu'ils apparaissent en haut de la liste des tags (Guy & Tonkin, 2006). En revanche, notre outil doit proposer aussi une interface pour nettoyer et ré-organiser ses tags en dehors des signets.

Nous proposons donc la création à la volée de relations entre tags grâce à un simple caractère liant les tags. Le caractère symbolisant l'inclusion est le « > » et celui signifiant la synonymie est le « = ». Par exemple, imaginons qu'un utilisateur désire tagguer une photographie en ligne de la tour Eiffel.

Il peut la décrire à l'aide des tags :

```
europe>france>paris, type>paysage, monument>toureffel
```

Le système va interpréter cette succession de tags comme une relation d'inclusion orientée entre les tags :

- *europe* et *france*
- *france* et *paris*
- *type* et *paysage*
- *monument* et *toureffel*

Le signet apparaît ensuite sous une forme simplifiée ne contenant que les feuilles des branches créées. C'est à dire :

paris, paysage, toureiffel

Imaginons que l'utilisateur désire ajouter un autre signet pointant vers la photo en ligne du Sacré-Coeur à Paris. Il décrira peut-être ce signet par :

paris, monument>sacrecoeur

Le système ajoute la relation d'inclusion entre les tags *monument* et *sacre-coeur*. L'affichage de tags sous ce signet devient :

paris, sacrecoeur

L'avantage de l'inclusion est par la suite d'améliorer la recherche de signets en permettant au système d'augmenter facilement les requêtes. Nous avons pour cela développé un modeste compléteur de requête basé sur cette relation d'inclusion. Si un utilisateur recherche les signés taggés *monument*, le système complètera grâce aux relations d'inclusion en une requête des signets taggés *monument, toureiffel et sacrecoeur*.

Imaginons maintenant qu'au moment de tagguer sa seconde ressource sur le Sacré-Coeur, l'utilisateur ne se souvienne plus si la catégorie parente est nommé *monument* ou *monuments* au pluriel. Nous avons alors implémenté la relation de synonymie qui permet de créer un rapprochement entre deux termes à la volée au moment de tagguer. Dans le doute l'utilisateur peut annoter le second signet par :

paris, monuments>sacrecoeur, monuments=monument

Le système interprète les deux derniers tags comme la création d'une relation de synonymie entre *monuments* et *monument*. Dans le cadre de la recherche de signets, cela signifie qu'une requête sur le tag *monuments* sera complété comme étant une requête sur *monuments, monument, toureiffel et sacrecoeur*. C'est à dire que pour une requête sur un tag, le système complète avec les tags synonymes et aussi inclus dans les tags synonymes de manière récursive. Bientôt, le système pourra aussi interpréter des successions de tags plus complexes comme :

monuments=monument>sacrecoeur

Comme nous le mentionnions précédemment, nos travaux ne cherchent pas à produire forcément des structures de tags cohérentes mais avant tout à améliorer la recherche à partir des tags. Ainsi, dans l'esprit des tags le système ne va pas signaler à l'utilisateur qu'un tag n'existe pas ou qu'une relation est impossible vis à vis du reste du système. Imaginons qu'un utilisateur crée un cycle d'inclusion avec les trois relations *europe>france, france>paris et paris>europe*. Si cette relation peut sembler illogique, pour notre système elle ne pose pas de problème. Lorsque l'utilisateur recherchera les ressources taggées *france*, le système complètera la requête en *france, paris, europe*. Le côté flou des tags permet de gagner en flexibilité et rapidité pour les utilisateurs. Nous pensons qu'en ajoutant les relations entre les tags, le système doit conserver cette flexibilité quitte à conserver une part de flou ou d'erreur.

Cependant, le système peut aider de manière non-intrusive l'utilisateur à nettoyer les relations fausses ou à en préciser d'autres. Le système propose deux interfaces de manipulation des relations : une simplifiée et une plus avancée.



## tags structurés (plus)



FIG. 2 – Boîtes affichant les tags structurés

L'interface simplifiée affiche les tags de manière arborescente selon la relation d'inclusion et affichant les uns à côté des autres les tags reliés par une relation de synonymie. L'illustration est un menu apparaissant sur le côté droit de la page dès que des relations structurées impliquent les tags recherchées (comme la boîte « recent tags » de la figure 1). Cela signifie que l'utilisateur peut modifier les relations entre les tags à partir de la même page où les signets impliqués par ces tags apparaissent sur le côté gauche. Ceci nous semble important pour faciliter une amélioration progressive des structures de tags. Entre parenthèses se trouvent les commandes permettant d'ajouter (add) ou d'effacer une relation (del).

En haut du menu se trouve la commande « (plus) » permettant d'accéder à une vue plus complète des relations où sont listées toutes les relations entre les tags sous la forme d'une succession de tag1, type de relation, tag2. L'utilisateur peut alors supprimer précisément des relations ou en créer de nouvelles.

Les tags structurés offrent de nouvelles perspectives pour l'affichage des tags que ne possèdent pas les tags classiques. Or il est démontré que les tags affichés aux utilisateurs ont une influence importante sur les tags qu'ils emploieront par la suite (Sen *et al.*, 2006). Les outils classiques comme Scuttle renvoient aux utilisateurs leurs activités en affichant les tags selon des critères de récence et fréquence. A partir des relations entre les tags, nous avons implémenté trois autres visualisations des tags ordonnés selon :

- le nombre de descendants. Le système affiche d'abord les tags incluant le plus de tags et sous-tags etc.  $a > b$ ,  $a > c$ ,  $b > d$  donne une descendance de trois tags à  $a$ .
- la profondeur maximale des branches. Le système affiche d'abord les tags possédant les branches les plus longues.  $a > b > c > d$  donne une profondeur de trois à  $a$ .
- le nombre maximal de modification des descendants. Un ajout ou un effacement compte pour une modification. Le système affiche d'abord les tags dont la somme des modifications des descendants est la plus grande, ce qui peut signifier un thème particulièrement actif.

Les résultats des expérimentations nous diront si mixés aux critères classiques, ces nouveaux modes de visualisations des tags facilitent leur organisation et leur usage pour les utilisateurs.

Pour l'instant nous avons considéré uniquement la situation où un seul utilisateur taggue et construit une structure de tags. SemanticScuttle permet l'annotation par plu-

sieurs utilisateurs. Si chaque utilisateur construit sa propre structure de tags, est-il possible de voir émerger une structure générale pour toute la communauté ? A la différence des travaux basés sur les ontologies, dans SemanticScuttle, les utilisateurs ne travaillent pas directement sur des concepts communs. Cependant, comme les tags qui travaillent sur des étiquettes partageables, nous faisons l'hypothèse que les structures de tags d'un utilisateur pourront être utiles à un autre utilisateur. Des tags classiques provenant de plusieurs utilisateurs peuvent facilement être agrégés pour donner une vue d'ensemble, par exemple dans un nuage de tags. Le même processus peut se produire avec les tags structurés par des relations. Si pour un utilisateur, le tag A inclut le tag B et que pour un autre utilisateur le tag A inclut le tag C, alors la page principale agrégeant leurs visions affichera le tag A incluant le tag B et le tag C. Ainsi cette agrégation conduit petit à petit à l'émergence d'une vision commune. Avec plusieurs utilisateurs, on peut ensuite imaginer que la structure générale ne montre que les tags et les structures, employés par plus d'un certain nombre d'utilisateurs afin d'éviter les structures isolées employées par un seul utilisateur.

Les relations structurées ont été implémentées et sont en cours d'expérimentation. Nous proposons maintenant deux autres fonctionnalités en cours d'implémentation. Celles-ci sont techniquement moins difficiles à implémenter mais offrent aux utilisateurs plus de moyen de structurer ensemble les tags et signets ajoutés.

## **6 Seconde fonctionnalité : espaces privés, espace commun, et espace de négociation**

Selon nous, le phénomène de structuration se joue à la limite entre espace privé (propre à un utilisateur) et espace commun (éditable par tous). Comme vu précédemment, les tags appartiennent généralement à chaque utilisateur alors que les applications employant des ontologies les rendent communes à tous. On pourrait dire que les tags correspondent à un espace privé et les ontologies sont un espace public commun. Cette notion d'espace privé des tags nous semblent alors avoir favorisé leur succès. En effet, il semble plus simple d'ajouter un tag de manière spontanée si ceci ne porte pas à conséquence pour les autres utilisateurs, que d'en ajouter un qui ira modifier le travail des autres utilisateurs. D'un autre côté, un espace public facilite la rencontre entre les opinions et peut être ouvert pour les utilisateurs plus investis dans la participation.

Dans Scuttle, pour l'instant, chaque utilisateur peut décrire le signet qu'il dépose. Cette description lui appartiendra et personne ne pourra la modifier. Un autre utilisateur enregistrant le signet créera une nouvelle description. Ces descriptions sont des espaces privés où chaque utilisateur peut exprimer son avis personnel. Nous prévoyons alors aussi l'intérêt d'un espace commun de description. Un espace de consensus où les utilisateurs s'accordent sur une description commune. Nous sommes en train d'implémenter un espace sous la forme d'une page éditable par tous, comme une page wiki. Nous ajoutons enfin un espace de négociation sous la forme d'un thread de forum. Ce thread correspond à la page discussion de wikipedia où sont négociés les modifications conflictuelles d'article. Pour une page sauvegardée par les utilisateurs 1 et 2, ceci donnera :

- URL du signet

- Description commune (espace public) avec lien vers thread de forum
- Descriptions utilisateur 1 (espace privé)
- Descriptions utilisateur 2 (espace privé)

Cette présentation se retrouvera dans la description possible des tags employés. Del.icio.us permet déjà de décrire ses propres tags. La notion d'espace de description commune et de négociation nous semble primordial pour la construction collaborative d'une représentation d'un domaine.

## 7 Troisième fonctionnalité : progressivité des descriptions

La dernière fonctionnalité porte sur le dépôt de nouveaux signets et met en avant l'aspect structurant de SemanticScuttle. La caractéristique de base d'une communauté est que tous les membres ont des engagements différents. Nous pensons comme Ross Mayfield que l'apport des sites web 2.0 est de favoriser la participation de tous et de permettre son exploitation. Wikipedia met en avant l'importance de laisser des mécanismes où chacun peut participer à la hauteur de son engagement (Endrizzi, 2006). Nous sommes en train d'implémenter dans SemanticScuttle un moyen de déposer une relation sans être identifié et sans le décrire. Cette action est la plus simple action possible sur l'outil après la lecture. Ensuite, SemanticScuttle possède une page listant tous les signets ajoutés de manière anonyme et ouverte à tous. Des utilisateurs ayant plus de temps ou plus engagés sur le site ont alors la possibilité d'enregistrer un signet dans leurs signets s'ils la trouvent intéressantes en ajoutant une description et des tags. Plus tard, un nouvel utilisateur pourra y ajouter au signet une description un peu plus travaillée dans l'espace commun.

Le processus correspond alors à :

- Un utilisateur peu engagé dépose une URL
- Un utilisateur un peu plus engagé parcourt régulièrement les URL sans descriptions pour tagger celles intéressantes.
- Un utilisateur encore un peu plus engagé décrit un peu plus longuement le signet dans l'espace commun.

Ce cas d'utilisation met en avant l'outil en tant que processus de filtrage de données brutes (une URL) vers son intégration dans le système de connaissance de la communauté. Ce processus collectif de structuration est possible si le système laisse assez de souplesse aux utilisateurs pour participer à leurs niveaux et sans trop de contrainte pour que leurs actions personnelles soient réellement personnelles et non pas une réplique exacte d'un point de vue autoritaire supérieur.

## 8 Conclusion et perspectives

Nous venons de présenter une approche souple et collective de structuration des tags pour l'annotation de signets.

Les résultats des expérimentations en cours dans une communauté de 1900 professeurs de génie mécanique<sup>10</sup> nous permettront d'évaluer nos hypothèses sur la facilité de structurer des tags à la volée ou d'appliquer le processus de description collaboratif. Nous pourrions alors comparer nos résultats avec d'autres outils en cours de développement. Nous pensons notamment à Wikia Search<sup>11</sup>, moteur de recherche réalisé par le fondateur de Wikipedia et basé aussi sur la collaboration des utilisateurs dans l'organisation des résultats.

Pour les expérimentations, nous sommes plus particulièrement attentifs à trois points. Premièrement, nous sommes attentifs à l'utilisabilité des relations entre tags et aux interfaces pour les organiser. Nous pensons ainsi tester des visualisations alternatives en liant SemanticScuttle à des outils de visualisations<sup>12</sup>. Deuxièmement, nous sommes attentifs aux interactions humaines dans la communauté autour de l'outil. Comment SemanticScuttle va compléter la liste de diffusion actuellement employée pour partager des ressources ? Est-ce que SemanticScuttle devra intégrer un mécanisme ouvert de modération comme Wikipedia ? Enfin, nous sommes attentifs aux termes et relations utilisés pour organiser les ressources. Nous attendons plus de données afin de comparer les structures de tags réalisées par les enseignants aux taxinomies des programmes officiels de l'éducation nationale.

Finalement, comme concluent (Trichet & Nizon, 2007), nous pensons que "les tags définis et partagés par les communautés sont profondément porteurs de sens et donc vecteurs des évolutions des domaines de connaissances sous-jacents." Dans cette optique, nous aimerions poursuivre l'exploration de ce qui est généralement caché dans les ontologies à savoir la dynamique d'une approche collective structurante.

## Références

- BATEMAN S., BROOKS C. & MCCALLA G. (2006). Collaborative tagging approaches for ontological metadata in adaptive elearning systems. In *Fourth International Workshop on Applications of Semantic Web Technologies for E-Learning (SW-EL 2006) in conjunction with 2006 International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2006)*, p. 3–12, Dublin, Ireland.
- BERNERS-LEE T., HENDLER J. & LASSILA O. (2001). The semantic web. *Scientific American Magazine*.
- BRAUN S., SCHMIDT A., WALTER A., NAGYPAL G. & ZACHARIAS V. (2007). Ontology maturing : a collaborative web 2.0 approach to ontology engineering. In *16th International World Wide Web Conference (WWW 2007)*, Banff, Alberata, Canada.
- BUFFA M. & GANDON F. (2006). Sweetwiki : semantic web enabled technologies in wiki. In *International symposium on Wikis*, p. 69–78, Odense, Denmark : ACM.
- DOWNES S. (2004). Resource profiles. *Journal of Interactive Media in Education*, **5**.
- ENDRIZZI L. (2006). L'édition de référence libre et collaborative : le cas de wikipedia.
- GREAVES M. (2007). Semantic web 2.0. *IEEE Intelligent Systems*, **22**, 94–96.

---

<sup>10</sup><http://www.listepgm.org>

<sup>11</sup>Wikia : <http://alpha.search.wikia.com/>

<sup>12</sup>Visualisations pour Del.icio.us : <http://www.solutionwatch.com/252/visualizing-delicious-roundup/>

- GRUBER T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, p. 199–220.
- GRUBER T. (2007). Ontology of folksonomy : A mash-up of apples and oranges. *International Journal on Semantic Web & Information Systems*, **3**.
- GUY M. & TONKIN E. (2006). Folksonomies : Tidying up tags ? *D-Lib Magazine*, **12**.
- HALPIN H., ROBU V. & SHEPHERD H. (2007). The complex dynamics of collaborative tagging. In *WWW07*, p. 211–220, Banff, Alberta, Canada : ACM.
- KARAPIPERIS S. & APOSTOLOU D. (2006). Consensus building in collaborative ontology engineering processes. *Journal of Universal Knowledge Management*, **1**, 199–216.
- MARLOW C., NAAMAN M., BOYD D. & DAVIS M. (2006). Ht06, tagging paper, taxonomy, flickr, academic article, to read. p. 31–40, Odense, Denmark : ACM.
- MOLINARI G., BAKER M., SÉJOURNÉ A. & LUND K. (2005). Les graphes argumentatifs interactifs sur internet : sont-ils plus efficaces comme média du débat ou comme moyens de le représenter ? In *EIAH 2005*, p. 201–211, Montpellier.
- PASSANT A. (2007). Using ontologies to strengthen folksonomies and enrich information retrieval in weblogs : Theoretical background and corporate use-case. In *International Conference on Weblogs and Social Media*, Boulder, Colorado, USA.
- SEN S., LAM S. K., RASHID A. M., COSLEY D., FRANKOWSKI D., OSTERHOUSE J., HARPER F. M. & RIEDL J. (2006). tagging, communities, vocabulary, evolution. In *CSCW06*, p. 181–190, Banff, Alberta, Canada : ACM.
- TRICHET F. & NIZON Y. (2007). Une plateforme de développement d’espaces webs sémantiques communautaires dédiés au partage de ressources multimédia. In *Ingénierie des Connaissances (IC’2007)*, Grenoble.
- ZACKLAD M. (2005). Vers le web socio sémantique : introduction aux ontologies sémiotiques. In *Ingénierie des Connaissances 2005*, Nice.