

## **Indexation collaborative : entre gain informationnel et déperdition conceptuelle ?**

L'affectation de mots-clés à des contenus est une activité d'écriture qui s'est généralisée aux cycles de publication et consultation web : on tague<sup>1</sup> des photos, des vidéos, des morceaux de musique, des articles, des billets de blogs. Les utilisateurs collectionnent et partagent des ressources par l'intermédiaire de mots-clés (tags). Ce processus de marquage de l'information se matérialise par un système flottant de catégorisation par étiquettes, personnel et partagé, permettant un repérage individualisé de ressources hétérogènes publiées sur le web. L'ensemble des mots-clés générés par les utilisateurs est appelé communément « folksonomy ».

Nous nous intéressons ici aux « folksonomies larges » qui concernent la recherche d'informations repérées par un ensemble d'outils de marque-pages, signets ou « bookmarking ». Cette pratique collective a reçu diverses appellations dont celui d'indexation collaborative que nous retiendrons ici. Suivant l'architecture informatique, le système de catégorisation peut être totalement flottant, n'utilisant que les tags entrés par les utilisateurs (ex : Del.icio.us) ou proposer aux utilisateurs le rattachement de tags à des descripteurs (ex : Connotea) ou structures sémantiques (ex : ZigTag), ou bien calculer automatiquement ce rattachement (ex : Twine).

Lors d'une communication à caractère prospectif (BROUDOUX, 2006) sur les folksonomies et l'indexation collaborative (via Del.icio.us), différentes questions avaient été soulevées qui ont connu depuis une évolution significative.

Alors que l'existence potentielle de relations hiérarchiques entre les tags avait été mise en perspective par Kome<sup>2</sup> dès 2005, des algorithmes pour repérer les hiérarchies existantes ont été conçus pour optimiser la navigation (HEYMANN et GARCIA-MOLINA, 2006) et pour détecter les conceptualisations implicites partagées entre utilisateurs dans l'objectif de les formaliser en ontologies (JÄSCHKE, 2007). Un ensemble d'outils tenant compte de ces évolutions a été développé dont la première partie fera l'état.

La création d'une couche signifiante de tags par les utilisateurs pourrait améliorer les résultats des recherches sur les moteurs. Le placement du « tagging » dans la recherche d'informations sur le web et son utilisation par les interfaces d'interrogation de catalogues en ligne sera la deuxième question étudiée.

L'accumulation d'informations réalisée en collectif entraîne la création de réseaux thématiques spécialisés repérable à la précision du vocabulaire utilisé et qui peut indiquer des domaines d'expertise. L'introduction du tagging dans les pratiques professionnelles fera l'objet de la troisième partie.

### **1. L'offre de la nouvelle génération d'outils de marque-pages**

---

<sup>1</sup> Le terme « taguer » est ici préféré au terme « balisage » à consonance plus technique.

<sup>2</sup> Expérience réalisée par 10 étudiants en sciences de l'information sur 30 url et 647 posts indiquait des relations de subordination entre tags adjacents.

Aux côtés de Del.icio.us, s'est greffée entre 2005 et 2008 une large panoplie d'outils de marque-pages se spécialisant dans des directions différentes (citons Stumble pour ses recommandations et Diigo pour ses annotations) mais se positionnant de plus en plus clairement comme « réseaux sociaux » (Twine).

## **Comment les outils existants résolvent les problèmes de hiérarchies entre les tags**

- **L'utilisateur documente le système**

Comme Connotea et CiteUlike qui permettent de référencer des publications, Bibsonomy lancé début 2006 fait partie des SRM<sup>3</sup> et s'adresse plutôt à la communauté universitaire (orientée BibTeX). Il est historiquement le premier service de partage de signets à offrir à l'utilisateur l'introduction d'une hiérarchie au moment de l'enregistrement de ses tags. Et ceci par l'intermédiaire de flèches (-> et <-) établissant une ascendance entre mots-clés. Les mots-clés généralistes deviennent alors des concepts dans la terminologie de Bibsonomy sur lesquels on peut faire les mêmes recherches que les tags. Cependant malgré toutes ses qualités, un nombre suffisant de tags<sup>4</sup> et d'utilisateurs n'est pas encore atteint par ce service issu de l'Université de Kassel qui rassemble majoritairement des taggeurs de langue allemande.

Projet quasi-confidentiel de l'université d'Oslo lancé fin 2006, Fuzzy<sup>5</sup> se revendique socio-sémantique (ZACKLAD, 2005) et est développé à partir d'une ontologie basée sur un moteur de topicmap<sup>6</sup>. Ici, pour mémoriser une page web, l'utilisateur doit d'abord spécifier un ensemble de fonctionnalités qui servira pour la recherche à facettes : genre de la ressource (outil, vidéo, portail, etc.), contexte géographique, catégorie (fait, commerce, loisir, etc.), type de connaissances concerné (pourquoi, comment, qui quoi où, etc.) et son niveau de scopie (panoramique, détaillé). Ce n'est que dans un deuxième temps que le processus de balisage commence en choisissant dans le « pot commun » (liste alphabétique) ou parmi ceux que le tagueur utilise (liste). L'enregistrement d'une définition pour un nouveau tag comporte la possibilité d'indiquer sa source sous la forme d'une URL. Possibilité est donnée de générer automatiquement des tags supplémentaires (reliés dynamiquement à ceux existants mais que l'on a pas encore trouvés). Enfin, la gestion communautaire qui est mise en place (proche de celle d'Everything2 et des jeux d'aventure avec gain de points) acquiert un caractère démocratique avec la gouvernance du site établie par élection.

La hiérarchisation entre tags est une action indépendante qui se déroule a posteriori et suivant un ensemble d'inclusions horizontales ou verticale entre tags, de manière réversible, ce qui permet des tests. Il s'agit d'une partie très aboutie qui renseigne une ontologie construite dynamiquement et dont il est possible d'approuver ou de désapprouver les constructions proposées.

Mais le principal problème se retrouve à l'usage quotidien pour l'utilisateur qui veut se servir des signets qu'il a enregistrés : lorsqu'il veut les consulter, il n'a accès à aucune partie personnelle et est rapidement confronté à l'impossibilité de retrouver ce qu'il a enregistré.

---

<sup>3</sup> Services References Management

<sup>4</sup> Une recherche simple sur self-efficacy (concept de la théorie sociale cognitive de Bandura) ramène quatre signets et trois publications pour Bibsonomy et 294 signets pour del.icio.us. (Autre exemple : folksonomy est taggué 1366 fois sur Bibsonomy et 83512 fois sur del.icio.us)

<sup>5</sup> Projet de Master de Roy Lachica et Dino Karabeg. 420 membres dont seulement 9 ont taggué plus de 50 pages web, l'immense majorité étant venue tester l'outil sans rien mettre en signet.

<sup>6</sup> TMCORE de NetworkedPlanet.

ZigTag, lancé en phase beta en avril 2008<sup>7</sup>, se présente comme un « outil intelligent » de bookmarking sémantique qui propose, au moment de l'introduction des tags, un choix de différentes significations<sup>8</sup> issu d'un thesaurus de deux millions d'entrées. Ici, l'utilisateur est clairement chargé de désambiguïser son vocabulaire à l'aide des mots-clés qui lui sont proposés et sur lesquels il peut exprimer un retour positif ou négatif. De même que Fuzzy, il peut enregistrer la définition de concepts mais leur hiérarchie lui reste cachée. Un moteur de recommandations offre à la consultation une liste de sites déterminée à partir des intérêts partagés par d'autres utilisateurs. Dans l'avenir, les auteurs prévoient de développer une api de manière à donner accès à l'ontologie co-construite par les utilisateurs.

- **Le système intègre les propositions des utilisateurs**

Twine, lancé fin 2007<sup>9</sup>, est un service de marque-pages orienté web sémantique. Il extrait automatiquement des « concepts » à partir des contenus qui lui sont soumis pendant que l'utilisateur ajoute ses mots-clés pour marquer les pages vues. S'il existe une page personnelle donnant accès à la liste des liens enregistrés, il n'y a aucun moyen pour l'utilisateur de visualiser ses « tags » pour les ré-organiser ou bien de constituer un réseau à partir du partage de ses mots-clés, ceux-ci restant invisibles après leur saisie.

La véritable richesse de Twine réside en fait dans les sept facettes de filtrage des résultats de la recherche (groupes, tags, personnes, places, organisations, types d'item, et personnes ayant utilisé le même tag). En réalité, Twine a repoussé à la marge la pratique du « tagging » ; certes l'utilisateur fournit le contenu mais cette activité n'est pas valorisée par le service et passe en arrière-plan, ce qui donne à penser que Twine ne joue pas sur le caractère potentiellement massif de l'indexation collaborative. En fait, le créateur de Twine est de l'avis que les outils de gestion spatiale de l'information sont en train de céder leur place aux outils de gestion temporelle de l'attention<sup>10</sup>. D'autres fonctionnalités comme l'intégration de documents<sup>11</sup> et la création de groupes « à la mode de Facebook » sont censées combler ce manque, ce qui pousse à classer Twine dans les services de type « réseau social ».

## **Quel processus ces outils sont-ils à même de soutenir ?**

- **L'aide à la recherche et à l'organisation**

L'exploration des quatre outils ci-dessus indique que la hiérarchisation des tags a essentiellement pour objectif de perfectionner l'organisation sémantique à l'intérieur du

---

<sup>7</sup> Par Reg Cheramy et Scott Montgomerie

<sup>8</sup> Par exemple, le terme article est défini selon quatre significations : article de journal, article de grammaire, article de commerce, article de loi

<sup>9</sup> Lancé par Radar Networks entreprise fondée en 2007 par Nova Spivack et toujours en phase bêta en août.

<sup>10</sup> « As our digital lives shift from being focused on the old fashioned desktop to the Web environment we will see a shift from organizing information spatially (directories, folders, desktops, etc.) to organizing information temporally (feeds, lifestreams, microblogs, timelines, etc.).

« Instead of being just a directory, the desktop of the future is going to be more like a feed reader or social news site. The focus will be on keeping up with all the stuff flowing in and out of the user's environment. The interface will be tuned to help the user understand what the trends are, rather than just on how things are organized.

« The focus will be on helping the user to manage their attention rather than just their information. This is a leap to the meta-level: A second-order desktop. Instead of just being about the information (the first-order), it is going to be about what is happening with the information (the second-order). »

<http://www.twine.com/item/11bshgkbr-1k5/the-future-of-the-desktop>

<sup>11</sup> Par téléchargement et publication.

système dans l'objectif d'effectuer des regroupements et des recommandations et non d'aider les utilisateurs à structurer leurs recherches en vue de futures actions.

La dernière version de Delicious (août 2008) a ajouté des fonctionnalités allant dans ce sens :

- la recherche dans les signets personnels et dans les signets du réseau de l'utilisateur,
- à l'enregistrement de la ressource, la proposition de signets rangés dans les bundles (regroupements de tags nommés par l'utilisateur).

Si elles facilitent la recherche, ces modifications apparaissent mineures dès qu'il s'agit de rechercher et d'organiser ses tags pour des actions.

Un seul outil, qui n'a connu pour l'instant qu'une existence de prototype répond au double enjeu de la recherche et de l'organisation de l'information et par l'utilisateur et par le système. Facetag (QUINTARELLI, 2006) a pour particularité d'associer les propriétés de hiérarchisation des tags à leur organisation en quatre facettes (type de ressource, thèmes, personnes, intentions), susceptibles d'être ensuite remaniées « éditorialement » pour améliorer la navigation dans le domaine. Chaque facette contient un ensemble de tags ordonné par nombre d'utilisateurs et date. Le signe + affecté à un tag est indicateur d'une hiérarchie qui peut être développée et laisser apparaître les tags qui lui sont reliés. Une liste de ressources reçoit les signets classés par ordre chronologique avec les tags qui lui sont affectés (classés par la couleur de la facette à laquelle ils appartiennent).

Les résultats d'une requête s'affichent dans les containers des facettes qui s'ajustent en conséquence en éditant la liste des ressources concernées. A chaque clic sur un tag, un filtrage additionnel supplémentaire affine les résultats. Cette recherche dynamique permet à l'utilisateur d'observer la hiérarchie des tags et leur appartenance à différentes facettes durant sa recherche dans les tags. Mais il n'est pas possible de sauvegarder les résultats pour une utilisation future dans d'autres environnements logiciels.

Le véritable intérêt pour l'utilisateur résiderait dans la possibilité de créer lui-même ses propres facettes pour l'action. Ainsi, pour l'étudiant qui prépare un exposé, le journaliste un article, le professeur un cours, la réorganisation dynamique des tags en vue de circonscrire des ressources pour des actions transitoires serait un réel apport de ces services.

La recherche d'information est avant tout contextuelle et comporte aussi une bonne part de subjectivité traduit par des mots-clés d'humeurs (cool, funny, interesting, etc.) ou d'auto-référence (me). Les tags dédiés à l'action planifiée dans le temps (todo, toread, howto) traduisent une volonté d'organisation pratique chez les utilisateurs. Les tags qui ne se rattachent pas directement à un sujet sont intrinsèquement liés au temps et expriment une relation dynamique de l'utilisateur au document et constituent selon (KIPP, 2007) des réponses personnelles apportées aux documents.

Pour améliorer l'utilisabilité de ces systèmes, il faudrait donc pouvoir se resservir de ces tags significatifs d'une action en préparation.

## • Des outils et des processus cognitifs

Si on a pu dire un moment que le « bookmarking » était une opération à faible coût cognitif, force est de constater que celui-ci se charge au fur et à mesure de l'augmentation du nombre de signets, dès lors qu'il s'agit de retrouver l'information taguée.

D'autre part, on pourrait imaginer que les outils de gestion de savoir informel que sont les outils de bookmarking seraient à même de soutenir l'utilisateur dans ses conceptualisations.

Si on considère le taggage comme une action s'inscrivant dans un processus cognitif d'acquisition de connaissances, on y reconnaît facilement trois phases : accroître l'existant (utiliser un tag connu pour le repérage d'une nouvelle ressource), agencer le présent (organisation des tags en hiérarchies pour préparer de futures actions) ou créer de nouvelles structures permettant de nouvelles conceptualisations (ajout de nouveaux tags provoquant un ré-agencement de tags anciens).

L'apprentissage mobilise des structures où des schémas sont réactivés de manière à être modifiés par l'apprenant, selon un cycle d'acquisition de connaissances comprenant d'après (RUMELHART & NORMAN, 1976) :

- l'accrétion de faits et d'informations sur des structures de connaissances connues (accretion),
- l'ajustement des schémas connus aux nouvelles connaissances (tuning),
- la création de nouvelles structures de connaissances pour accueillir et interpréter du nouveau matériel (restructuring).

Il serait intéressant de s'interroger sur les processus favorisés par le tagging : création de nouvelles structures, ré-agencement de structures existantes ou simplement leur accroissement ? L'horizontalité des tags tellement vantée n'est-elle pas un obstacle au mécanisme de conceptualisation qui exige de pouvoir hiérarchiser soi-même l'information ?

## • Un processus d'auctorialisation

Nous faisons l'hypothèse que la catégorisation par la « mise en signets » entrerait dans un processus d'auctorialisation. Que le tagueur choisisse les mots issus de son vocabulaire personnel ou pioche dans celui qui est mis en commun, il « écrit ». Et il « réécrit » lorsqu'il modifie son expression. Cette écriture « devant » et « par » les autres n'est pas anodine, elle habitue progressivement le « marqueur de pages » à dévoiler ses intérêts. D'autre part, des constantes émanent de l'observation d'un réseau de 150 tagueurs dans Del.icio.us :

- les auteurs institués, c'est-à-dire tirant leur autorité d'organisations antérieures au web, divulguent rarement leur identité en indiquant un lien vers leur CV ou leur institution,
- les auteurs autoritatifs « non institués » ou auteurs « en devenir » s'identifient en indiquant leur nom ou/et en déposant un lien vers leur dispositif de publication (blog, etc.).

Ensuite, la durée d'observation de trois années montre un dévoilement progressif des utilisateurs anonymes qui abandonnent leur incognito ne serait-ce qu'en laissant de multiples traces sur le web. Ce peut être dû à l'ouverture d'un blog et/ou à la montée en puissance de l'activité des réseaux sociaux.

## 2. État de l'art

### **Orientation informatique des études disponibles**

On ne peut que constater la richesse des études informatiques<sup>12</sup> pour définir des algorithmes, spécifier des ontologies pour le tagging, explorer par des études quantitatives l'efficacité de l'accumulation d'informations par les folksonomies. Des groupes de travail se sont constitués (*TagsCommon*) réalisant des répertoires de hiérarchies (*Open Tag Directory*) ou développant des ontologies comme *Scot*<sup>13</sup> qui spécifie pour le web sémantique les principaux concepts et

---

<sup>12</sup> Celles-ci se trouvent à l'intersection de recherches sur la gestion des connaissances et les métadonnées, la génération d'ontologies, le datamining, la classification documentaire et les systèmes collaboratifs.

<sup>13</sup> Scot : Social semantic clouds of tags. <http://scot-project.org/>

propriétés nécessaires à la description du process de tagging. L'amélioration des résultats de la recherche d'informations (*Information retrieval*) par les pratiques d'indexation collaborative connaît plusieurs orientations dans la conception de nouveaux algorithmes :

- *Une direction sémantique* :
  - o la désambiguïsation des termes (problèmes de polysémie, homonymie, synonymie), différentes formes lexicales exprimant le même concept peuvent être ramenées à une seule expression (ex : pluriel/singulier), les erreurs dans l'orthographe des mots ne sont plus des obstacles à la recherche. La reconnaissance sémantique des termes comme l'amélioration de l'extension de requêtes par similarités avec des modèles de détection des synonymes (Clements, 2008)
  - o l'intégration de facettes ou de clusters dans les interfaces de recherche.
- *Une direction sociale* avec la prise en compte du réseau de tagueurs comme facteur d'amélioration des résultats de la recherche. Exemple, *ContextMerge*, nouvel algorithme (SCHENKEL et al., 2008) prend en compte, dans le classement des réponses aux requêtes le réseau social impliqué dans le partage des mot-clés. Il s'agit d'un algorithme de requête incrémental à deux dimensions d'extension : une extension « sociale » vers des utilisateurs « similaires » qui mesure l'intensité des connections entre utilisateurs et une extension « sémantique » qui calcule les liaisons entre les tags.
- *Une direction socio-sémantique* qui vise à relier le vocabulaire libre des taggeurs aux taxonomies du web sémantique. Par exemple, *Moat*<sup>14</sup> nouveau framework établit un pont entre les tags entrés librement par les utilisateurs et les données sémantiques (PASSANT & LAUBLET, 2008). Il élargit le graphe tripartite décrivant la structure (utilisateur, tag, ressource) en insérant une signification contextuelle pour chaque action de tagging (utilisateur, ressource, tag et sa signification). Ainsi, la signification globale du tag est composée de multiples interprétations en liaison avec les taggeurs eux-mêmes.

## Recherche d'informations sur le web

D'un autre côté, le lancement d'études comparatives avec les moteurs de recherche du web vise à mettre en évidence le potentiel des folksonomies dans la recherche ouverte d'informations ou ROI selon (ZACKLAD, 2007).

Des études récentes ont cherché à étudier le recouvrement entre l'indexation réalisée par les moteurs de recherche et l'indexation collaborative. (HEYMANN, 2007) a comparé les résultats de 9 mois de crawling de Del.icio.us (septembre 2006 à fin juillet 2007) avec l'indexation réalisée par les moteurs de recherche. L'objectif de l'étude était de quantifier la taille de l'information générée par les utilisateurs (mots-clés des pages mises en signets), et d'en regarder ses caractéristiques de manière à estimer si celle-ci était à même d'améliorer les résultats obtenus par les moteurs de recherche du web.

La comparaison des requêtes entre Yahoo, AOL et Del.icio.us indique que Del.icio.us apparaît pour 19% aux dix premiers résultats aux requêtes correspondantes et pour 9% aux cent premiers résultats. Ramené à la couverture web de del.icio.us, ce résultat mérite d'être remarqué. La taille de Del.icio.us a été estimée en juin 2007 à 115 millions de posts publics référençant 30-50 millions d'url uniques. Avec une moyenne de 120 000 posts par jour, la

---

<sup>14</sup> Moat : Meaning of a tag. <http://moat-project.org/index>

contribution s'avère faible par rapport au 1,5 million de billets postés par jour sur les blogs. Heymann juge impossible de prévoir la taille future de Del.icio.us, sa croissance étant moins dépendante de facteurs organiques que de stratégies marketing, ce qui est aussi le problème des outils cités plus hauts. Il admet cependant qu'actuellement les posts ne sont numériquement pas assez nombreux pour couvrir le classement obtenu par un crawl d'un moteur de recherche majeur sur le web.

Mais c'est le caractère inédit des pages postées sur Del.icio.us qui doit être remarqué : soit elles viennent juste d'être créées et ne sont pas encore indexées par les moteurs (25%), soit elles ont été modifiées récemment. Le contenu référencé y est nouveau : bien que 20 % de posts se réfèrent à des url postées plus de 50 fois, 30-40 % correspondent à des url postées sur Del.icio.us pour la première fois ; 1 post sur 8 correspond à un nouveau nom de domaine.

Del.icio.us apparaît comme une application massivement distribuée qui ne dépend pas d'un nombre faible de contributeurs, contrairement à Digg où les 100 premiers utilisateurs contrôlent 56% du contenu et exercent différentes formes de pression. L'observation sur une période d'un mois de la proportion de posts réalisée par les n premiers utilisateurs indique que 80% du contenu est contrôlé par les 100 000 premiers utilisateurs. Pour couvrir 50% du contenu posté, il faudrait recourir à plus de 30 000 utilisateurs.

Une autre étude comparative sur un jeu de documents indexés par Google, ICRA, DMOZ<sup>15</sup> et Del.icio.us menée par (NOLL & MEINEL, 2007) confirme le bon classement de Del.icio.us dans les résultats fournis par les moteurs : les documents les plus fréquemment taggués ont un PageRank (PR) élevé et dans le corpus étudié, 78,7 % du tagging est appliqué à seulement 18,2 % des documents dont le PR se situe entre 5 et 7. Il note que les informations apportées par les tags sont plus complètes que les métadonnées html et permettent de mieux situer les documents.

## **L'intégration des folksonomies aux interfaces publiques d'accès des catalogues de bibliothèques**

Lorsque Britta Gustafson publiait sur le blog de Del.icio.us en mai 2008<sup>16</sup> un billet intitulé « Qui a dit que les bibliothécaires et les enseignants n'aimaient pas les tags ? », elle dévoilait peut-être au « grand public » l'attrance des professionnels des sciences de l'information (informaticiens, bibliothécaires, enseignants et doctorants) pour ce qui apparaît aujourd'hui constituer un nouveau champ de recherches et d'expérimentations pour différentes communautés de l'*information retrieval* sur le web.

La récente étude menée par Tosca Consultants indique qu'une moyenne de huit Opacs<sup>17</sup> sur dix propose l'intégration de l'apport des usagers (tags, commentaires, évaluations) aux catalogues des bibliothèques en ligne. Quatre sur dix enrichissent les possibilités de recherche avec les annotations des lecteurs et s'en servent comme critères de restrictions ou facettes (MAISONNEUVE, 2008).

---

<sup>15</sup> Etude comparative sur 97578 URL portant sur les métadonnées d'auteurs (entrées par les auteurs des pages html et sur les répertoires) et les métadonnées de lecteurs (entrées sur les logiciels de bookmarking)

<sup>16</sup> <http://blog.delicious.com/blog/2008/03/who-says-librarians-and-teachers-dont-like-tags.html>

<sup>17</sup> Online public access catalog

- **De nouveaux intermédiaires entre lecteurs, libraires et bibliothécaires : Librarything**

L'exemple emblématique des nouveaux intermédiaires entre lecteurs et bibliothèques est LibraryThing (LT)<sup>18</sup> qui propose aux lecteurs depuis août 2005 de « cataloguer » leurs livres dans l'objectif de relier les utilisateurs en fonction de leurs listes à partir desquelles un système offre des recommandations de lectures. Le lecteur remplit une notice correspondant à chaque livre de sa bibliothèque (actuelle ou future) et renseigne de ses propres mots les « mots-clés ». Les tags fonctionnent ainsi comme des embrayeurs vers d'autres livres catégorisés de la même façon. Diverses techniques améliorent le tagging (tagmash, etc.). Comme tous les sites de réseaux sociaux, le site suppose l'inscription et la gestion d'un profil puis la création ou l'inscription dans des groupes publics ou privés (clubs ou géolocalisés). LT utilise la norme Z39.50 qui lui permet d'importer les notices ISBN des librairies (Amazon) et de 680 bibliothèques.

Deux profils très distincts d'utilisateurs font vivre LT : les lecteurs et ceux qui cherchent à les rejoindre (les bibliothécaires et les libraires).

Un examen des statistiques des groupes de discussion est une bonne indication de la réalité des échanges : ainsi le premier groupe en termes de membres est celui des bibliothécaires (*Librarians who LibraryThing* : 4901 membres et 5514 messages), suivi de trois groupes thématiques (science-fiction, bande dessinée et aventures fantastiques intitulé *Green dragon* qui comporte 1410 membres et a échangé 105 805 messages). Cette orientation littéraire désigne à première vue une population plutôt jeune et masculine.

L'examen des 38 groupes francophones totalisant 311 inscrits laisse transparaître la faible participation des quelque 3700 membres francophones inscrits sur LT<sup>19</sup> : un groupe sur 4 est constitué d'un seul membre et près de la moitié des inscrits l'est dans des mini-groupes. De plus, leur constitution est plus basée sur le langage (pratiquer le français ou regrouper les lecteurs québécois, suisses, belges et français) que sur une thématique de lecture (la science-fiction est cependant représentée). Cette faible participation est confirmée par une longue note de lecteur qui s'étonne d'avoir eu peu de contacts après avoir publié sa bibliographie de 200 livres (maximum atteint avant que le service ne devienne payant). Est-ce qu'une masse critique d'utilisateurs francophones n'a pas été atteinte ou bien est-ce que cette association automatique par affinités n'est pas suffisante pour qu'un groupe d'intérêt littéraire se crée ?

Les nuages de mots-clés générés par LT sont insérables dans les Opacs. L'adoption de LT en bibliothèque et son intégration par widgets (nuages de tags, mots-clés, commentaires) sur les interfaces d'interrogation documentaire apparaît comme un réel succès en Amérique du Nord au vu des multiples exemples d'intégration<sup>20</sup>. Mais LT apparaît surtout servir aux professionnels du livre pour faire de l'animation autour de la lecture : récupération des mots-clés de lecteurs publiés sous forme de nuages, ajout de liens vers des ouvrages similaires, de recommandations, commentaires et conseils.

---

<sup>18</sup> Statistiques d'août 2008 : membres : 473 908, livres catalogués : 30 263 858, tags ajoutés : 39 335 482.

<sup>19</sup> 40 si on ajoute les groupes d'aide (11 membres) et de traduction-maintenance (47 membres). (Nbre de groupes, Nbre de membres) = (11,1), (7,2), (6,3), (1,4), (1,5), (1,7), (3,10), (3,12), (1,13), (1,16), (2,18), (1,21), (1,37), (1,46), (1,116).

<sup>20</sup> Comme l'indique un récent bilan de Ellissa Kroski en juin 2008 : <http://www.slideshare.net/ellyssa/libraries-and-the-hive-mind-folksonomies-and-tagging>



Le réel succès de LT est la réalisation du branchement entre « folksonomies » et « vocabulaires contrôlés ». La prise de conscience que le flux d'informations en provenance des lecteurs-usagers peut enrichir la navigation dans les opacs et faciliter l'exploration du fonds a fait son chemin. La circulation entre les documents est ainsi renforcée par la folksonomy et peut faciliter la recherche de la « bête rare » dans un fonds existant. Cette intégration s'inscrit plus largement dans l'ouverture des interfaces de recherche documentaires aux services web<sup>21</sup>.

- **Participation des usagers à la description documentaire**

L'étude comparative de (SMITH, 2007) entre les nuages de mots-clés attribués à cinq livres sur LibraryThing et les vedettes-matières correspondantes de la Bibliothèque du Congrès a montré que les tags des utilisateurs indiquaient plus de sujets « latents » que ceux du catalogue, un lecteur étant capable de déceler de multiples sujets au sein d'un livre.

En ce qui concerne la génération des nuages de mots-clés, l'abus de synonymes dans les folksonomies peut toutefois provoquer la disparition de tags (calibrés pour ne représenter que les mots les plus fréquemment employés). D'autre part, la souplesse des entrées (pluriel, singulier) autorise une forme de recherche en langage naturel que les opacs interdisent, de même que les tags qui ne correspondent à aucun sujet sont considérés comme des nouvelles entrées potentielles.

Le projet Penntags de la bibliothèque de l'Université de Pennsylvanie est un exemple d'intégration des folksonomies en parallèle aux différentes entrées du catalogue. Plus qu'un nuage de tags, le service offre de véritables fonctionnalités de gestion de dossiers et de signets qui peuvent soutenir le processus d'une recherche sur opac. Une thèse a récemment étudié si la description documentaire pouvait être influencée par la façon dont les étudiants utilisent Penntags (ERIKSSON-ÅHL, 2007). L'étude quantitative a consisté à étudier si les meilleures contributions des usagers étaient dépendantes des sujets choisis. Si l'ensemble des apports à la description documentaire est pertinent dans 59% des cas, l'intérêt des usagers à améliorer le catalogue se concentre essentiellement sur les matières Langues, Littérature et Sciences humaines.

L'intégration des folksonomies sur les catalogues dans les bibliothèques améliore sans aucun doute la richesse des résultats de la recherche comme l'indique le catalogue expérimental de l'université de sciences et technologie de Hong-Kong<sup>22</sup>.

Cet apport des usagers à la description documentaire mérite de retenir l'attention. Loin de menacer les professions des documentalistes et bibliothécaires, il oblige à envisager l'élargissement des services aux lecteurs en tenant compte de leur littératie informationnelle ou de leur transliteracy (LE DEUFF, 2008)

### **3. Usages sociaux du tagging**

#### **Intégration professionnelle de l'indexation collaborative**

---

<sup>21</sup> Un fil rss permet d'être tenu au courant des dernières expérimentations : <http://www.rss4lib.com/library-labs.html>

<sup>22</sup> <http://catalog.ust.hk/catalog/>

Une véritable intégration professionnelle de l'indexation collaborative existe mais il est difficile de l'évaluer car le tagging est une activité récente de communication qui s'exerce de manière contextuelle pour divers besoins. En même temps, la recherche par des études quantitatives sur les usages sociétaux des systèmes est focalisée sur l'amélioration de la production d'informations. Les études d'usage en situation sont rares car supposent une observation en immersion et une pratique régulière des outils avec des temps longs d'activité, de réflexion et d'analyse.

- **Bookmarking dans l'enseignement**

Un récent projet d'Europe du Nord-Est (VUORIKARI, 2007) a expérimenté un outil de bookmarking pour les enseignants de 78 écoles, traduit pour quelques écoles pilotes en hongrois, tchèque, allemand et polonais. Les ressources pédagogiques sont accessibles par un portail mais le service restant en local en fait un outil de travail personnel. L'observation s'est portée sur une analyse sémantique des tags multilingues et sur les 13 000 interactions qui s'étaient déroulées sur le portail ayant résulté en 460 marque-pages et 600 tags. Les interactions sur 31 jours d'observation ont été classées par ordre décroissant d'activité : la consultation de l'item, la recherche, la consultation des détails, la consultation exhaustive, le bookmarking, la consultation des évaluations, l'évaluation et l'effacement des bookmarks. Sur 4641 interactions seulement 127 étaient consacrées au bookmarking, les enseignants passant plus de temps à la consultation et à la recherche approfondie plutôt qu'au repérage. Seulement 5% des items consultés ont été tagués. On retrouve 93% de tags factuels qui se décomposent en 79% associé à des thèmes, 14% à une catégorie plus fine et 7 % étant subjectifs (qualité des ressources ou ce que les utilisateurs en pensaient).

On voit que pour cette profession, l'évaluation et la sélection des ressources est une étape très importante et le tagging (à valeur positive ou négative) n'intervient que comme une manipulation qui termine l'estimation des moyens mis à disposition.

Plus généralement, l'insertion pédagogique du bookmarking (repérable aux tags-acronymes de noms de cours) fait partie des pratiques d'enseignants dans les domaines liés à la recherche d'information et à l'étude des nouveaux médias. Citons l'intégration récente d'un « passe-liens » dans la plate-forme d'enseignement à distance Claroline<sup>23</sup> qui pourrait signaler un élargissement progressif de ces nouvelles pratiques pédagogiques à l'ensemble du corps enseignant.

- **Tagging en entreprise**

L'intégration professionnelle du tagging dans l'activité quotidienne de multiples salariés travaillant en réseau est devenue une réalité pour les entreprises expertes dans les technologies de communication.

Au sein d'IBM (qui a lancé ces trois dernières années de multiples études sur le tagging au sein de son organisation), les usages de « bookmarking » de quatre services ont été étudiés (Signets, Répertoire d'adresses, Blogs, Activités) par (MULLER et al., 2007). Pendant un an, sur 4987 utilisateurs identifiés, un repérage a été réalisé pour vérifier si les 28 460 tags utilisés conservaient les mêmes significations dans l'utilisation des différents services malgré l'hétérogénéité des ressources. Ceci afin de vérifier si les conditions étaient réunies pour partager des concepts, exigence nécessaire à l'entreprise qui mise sur les travaux en collectifs. Il s'agissait également d'observer l'émergence des descripteurs du savoir commun au sein de l'organisation.

---

<sup>23</sup> Réalisée en 2008 par Marie Leproust et Xavier Roussel de l'Université de Versailles-Saint-Quentin.

L'étude menée cherchait à anticiper le fait que des utilisateurs sont susceptibles de ne pas utiliser les mêmes mots pour décrire les mêmes concepts.

L'analyse des tags a révélé un faible recouvrement des tags en commun pour les utilisateurs au sein de chacun des services, les usagers gardant un vocabulaire personnel tout au long des utilisations et cela malgré le fait qu'ils exerçaient les mêmes tâches en utilisant les mêmes services.

La conclusion est que le système doit encourager la réutilisation des tags sans toutefois chercher à normaliser leur entrée de manière à conserver la diversité du vocabulaire. Cependant lors de leur enregistrement dans la base de données, les différentes formes possibles d'une même expression devraient pouvoir être stockées ; de la même manière au moment de la recherche d'informations, une extension de la recherche au vocabulaire avoisinant devrait pouvoir améliorer les résultats.

- **Gérer la communauté au sein de l'entreprise**

Toujours au sein d'IBM, plus récemment, ce sont les rôles endossés par les taggeurs qui ont été mis en évidence par (THOM-SANTELLI, 2008) à partir d'une étude qualitative sur le tagging (33 utilisateurs – dont douze femmes – répartis dans six pays) sur quatre services constitués par un site de marque-pages, un répertoire d'adresses, un blog et un répertoire de podcasts. Là encore, le contexte d'utilisation est déterminant dans l'identification des conduites. Ainsi, cinq rôles sociaux ont été identifiés :

- **community-seeker** : les tags servent à chercher et identifier des communautés que les taggeurs cherchent à incorporer ; les tags étant choisis pour articuler des connections avec les autres ;

- **community-builder** : lorsque la communauté n'existe pas, il faut la créer et les taggeurs veillent à ce que le vocabulaire choisi puisse permettre à d'autres de les rejoindre, quitte à adopter un vocabulaire commun pour des activités spécifiques : les pratiques de tagging confirment qu'il s'agit d'être reconnu comme membre actif d'une communauté et en même temps autogérer sa carrière au sein d'une communauté de pratiques ;

- **evangelist** : ils servent de connecteurs à ceux qui partagent les mêmes intérêts au sein de l'organisation ; non seulement l'évangéliste utilise tous les moyens mis à disposition pour diffuser ses messages mais n'hésite pas non plus à rechercher de l'information qui donnera du sens à l'échange entre les membres. Le tagging devient une stratégie d'amélioration de la réputation du groupe et de défense des intérêts de la communauté ;

- **publisher** : c'est celui qui est responsable de la dissémination à diverses cibles. Comme pour le constructeur de communautés et l'évangéliste, le publisher est intéressé par la valorisation des contenus par les tags de manière à accroître l'audience (commerciaux, consultants, etc.) sans pour autant adhérer personnellement à la communauté, il visera à choisir des tags susceptibles d'être choisis pendant la recherche d'infos ;

- **small team leader** : il emploie une terminologie pour ses tags dans l'objectif d'être d'abord compris par les membres de son équipe (nom de projet, acronyme, etc) ; il a une activité variable en termes de tagging comme les autres membres de son équipe qui trouve une explication dans le fait que les membres entretiennent souvent des rapports sur le terrain.

Le souhait de maintenir une consistance dans l'appellation des tags est une motivation partagée par l'ensemble des acteurs d'abord pour retrouver leurs ressources et ensuite pour les partager. Parmi les recommandations qu'émettent les auteurs de l'étude pour l'amélioration des services en ligne, notons :

- Le besoin des usagers de pouvoir visualiser l'ensemble des tags utilisés dans tous les systèmes pour faciliter leur ré-emploi dans chaque service. Le fait de pouvoir

consulter les tags par système leur était peu utile et les nuages de tags sont apparus inappropriés pour leurs tâches.

- Le développement d'outils afin de faciliter la réalisation d'objectifs en fonction des rôles sociaux. En particulier, ceux qui distinguent audience et communauté : taguer pour gérer l'appartenance à un groupe n'est pas taguer pour émettre en direction d'un public.
- Enfin, offrir des recommandations de haut niveau – en fonction des rôles – aux étapes de sélection des tags et de leur recherche pourrait améliorer l'efficacité de ces systèmes.

## Conclusion

L'intégration du tagging dans différents types de services et dans les pratiques professionnelles confirme qu'une pratique d'écriture liée à la catégorisation est née et se développe sur le web.

Bien que le bookmarking ne puisse que dans une moindre mesure affiner les résultats de la recherche par les moteurs traditionnels sur le web – compte-tenu du faible recouvrement des tags sur le volume total indexé – cette pratique reste pertinente dans la mesure où le vocabulaire concerne des thématiques déterminées et/ou des communautés identifiées.

### De nouveaux services

L'exploration des nouveaux services de marque-pages indique une maturité dans les fonctionnalités proposées et des avancées significatives du côté de l'automatisation des structures sémantiques. Cependant, les outils lancés récemment en phase bêta entrent en concurrence avec ceux qui sont déjà implantés et leur adoption est dépendante de facteurs économiques et sociétaux. Il leur sera difficile de fidéliser de larges publics, compte-tenu de la profusion existante de l'offre et des stratégies marketing à développer pour apparaître intéressants aux yeux des internautes. Ces outils pourraient cependant trouver un avenir dans les communautés partageant des savoir faire et ayant intérêt à partager et utiliser un vocabulaire commun.

Les fonctionnalités participatives liées à l'enregistrement du vocabulaire doivent être préservées :

- entrée libre des tags,
- examen des tags employés par les membres du réseau,
- retour sur les tags pour correction du lexique employé.

L'organisation des tags pour l'action nécessite leur gestion et cette partie du développement logiciel apparaît aujourd'hui délaissée au profit de l'automatisation des recommandations. Ceci pourrait mener à une forme de mort sociale de l'outil. Pour éviter les déperditions conceptuelles, il faut que l'utilisateur puisse ranger ses signets dans des structures amovibles, personnalisables et intégrables dans des hiérarchies taxonomiques.

L'objectif devrait donc être de passer au stade suivant qui serait d'aider l'utilisateur à organiser ses tags pour l'action, aussi nous préconisons de :

- rendre réalisables différents modes d'associations ressource/action (todo, howto, toread), ressource/évaluation (interesting) à choisir par l'utilisateur,
- proposer des modes d'organisation temporaire des concepts puisqu'un concept est susceptible de servir à de multiples organisations,
- autoriser des commentaires privés et publics sur ces organisations éphémères,
- améliorer l'organisation personnelle des tags par des regroupements reliés dynamiquement aux abonnements par tag,

- soigner la visualisation graphique des hiérarchies et inventer leur navigation (graphes dynamiques).

### **Rechercher de l'information c'est explorer**

Dans le cadre du bookmarking, la caractéristique exploratoire du tagging est démontrée et c'est ce process qu'il faut soutenir et encourager. Le fait qu'une part importante des pages tagguées sont nouvelles ou nouvellement modifiées et qu'elles ne sont pas encore indexées par les moteurs incite à penser que les usagers sont dans une tactique de repérage des informations. Le réseau de tagguateurs est alors utilisé comme filtre, signalant l'émergence d'idées, le lancement de nouveautés, l'apparition de thèmes de discussions, la concentration de l'attention sur des objets.

Dans le cadre plus général du tagging, l'intégration des folksonomies aux catalogues en ligne des bibliothèques donne à l'utilisateur des moyens d'exploration du fonds, en même temps qu'elle le valorise.

### **Pratiques professionnelles du tagging**

Le partage de vocabulaire « au travail » est une pratique récente et vise différents objectifs inhérents à la communauté professionnelle. Le tagging en entreprise concerne des utilisateurs identifiés dont les rôles au sein des groupes de travail sont repérables.

Fin 2007, était rajoutée aux pages des résultats de Google l'indication de la présence de signets indexés et commentés sous Stumble ; par ailleurs GoogleLabs explore actuellement l'ajout des commentaires « sociaux » aux résultats des recherches<sup>24</sup>. En conclusion finale, on peut dire aujourd'hui que nous assistons à l'interconnexion de trois domaines liés à la recherche d'informations : le monde documentaire professionnel fondé sur une description normalisée des connaissances et sur l'autorité des auteurs, le monde des usagers connectés qui introduit le filtrage collaboratif comme élément décisif de l'opération tri-classement, et celui des moteurs de recherches reposant sur la densité des liens hypertextes et la notoriété de réseau.

---

<sup>24</sup> <http://www.google.com/experimental/a840e102.html>

## Références

Broudoux, E. (2006). « Folksonomies et indexation collaborative. Rôle des réseaux sociaux dans la fabrique de l'information ». (Document en ligne sur <DocForum. <http://www.docforum.tn.fr/documents/23&24nov06SavResPar06InterBroudouxE.pdf>>)

Clements, M., de Vries, A. P., and Reinders, M. J. (2008). Detecting synonyms in social tagging systems to improve content retrieval. In Proceedings of the 31st Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Singapore, Singapore, July 20 - 24, 2008). SIGIR '08. ACM, New York, NY, 739-740. DOI= <http://doi.acm.org/10.1145/1390334.1390479>

Crecelius, T., Kacimi, M., Michel, S., Neumann, T., Parreira, J. X., Schenkel, R., and Weikum, G. 2008. Social recommendations at work. In Proceedings of the 31st Annual international ACM SIGIR Conference on research and development in information retrieval (Singapore, Singapore, July 20 - 24, 2008). SIGIR '08. ACM, New York, NY, 884-884. DOI= <http://doi.acm.org/10.1145/1390334.1390558>

Eriksson-Åhl, C. (2008). PennTags: A quantitative empirical study of read/write OPAC. url : <http://hdl.handle.net/2320/2289>

Gendarmi D., Lanubile F., « Community-driven ontology evolution based on folksonomies », OTM Workshops 2006, LNCS 4277, pp. 181 – 188, 2006. (Document en ligne sur [http://cdg.di.uniba.it/cdg/gendarmi/papers/gendarmi\\_lanubile\\_COMINF06\\_id31.pdf](http://cdg.di.uniba.it/cdg/gendarmi/papers/gendarmi_lanubile_COMINF06_id31.pdf))

Heymann, P., Garcia-Molina, H., (2006). Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems, Technical Report InfoLab. (Document en ligne sur <http://dbpubs.stanford.edu/pub/2006-10>).

Heymann, P., Koutrika, G., P., Garcia-Molina, H., (2008). Can Social Bookmarking Improve Web Search?, First ACM International Conference on Web Search and Data Mining (WSDM'08), 2008.

Hunter, J., Khan, I., and Gerber, A. 2008. Harvana: harvesting community tags to enrich collection metadata. In Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (Pittsburgh PA, PA, USA, June 16 - 20, 2008). JCDL '08. ACM, New York, NY, 147-156. DOI= <http://doi.acm.org/10.1145/1378889.1378916>

Jäschke, R., Hotho, A., Schmitz, C., Ganter, B., Stumme, G. (2008). Discovering shared conceptualizations in folksonomies. Web Semantics: Science, Services and Agents on the World Wide Web. Volume 6, Issue 1. Pages 38-53.

Hotho, A., Jäschke, R., Schmitz, C., Stumme, G. (2006). BibSonomy: A social bookmark and publication sharing System. In Aldo de Moor, Simon Polovina, and Harry Delugach, editors, Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on conceptual structures, Aalborg, Denmark. Aalborg University Press.

Kipp, M. (2008). @toread and Cool : Subjective, affective and associative factors in tagging. Proceedings of the 36th annual conference of the Canadian Association for Information

Science (CAIS), University of British Columbia, Vancouver, Canada.

Kome H. Sam. « Hierarchical subject relationships in folksonomies ». Master's Paper for the M.S. in I.S degree. November, 2005. 32 pages. Advisor: Jane Greenberg. <http://hdl.handle.net/1901/238>

Lachica, R., Karabeg, D. Metadata creation in socio-semantic tagging systems : towards holistic knowledge creation and interchange.

Lebreton, C. (2008). Bibliothèques, tags et folksonomies. Mémoire d'étude de Conservateur de bibliothèque. <http://www.enssib.fr/bibliotheque-numerique/document-1750>

Le Deuff, O. (2008). « Le Ka documentarisé et la culture de l'information » in (Broudoux E., Chartron G., dir.). Traitements et pratiques documentaires : vers un changement de paradigme ? Actes de la 2<sup>e</sup> conférence Document numérique et Société, ADBS.

Mitis-Stanzel, I. (2008). Social Tagging in Bibliotheken. Master of Science (MSc) in Library and Information Studies. Vienne, Autriche.

Maisonneuve, M., (2008). Le catalogue de la bibliothèque à l'heure du web 2.0. Etude des opacs de nouvelle génération (Tosca Consultants). ADBS.

Mathes A. « Folksonomies - Cooperative classification and communication through shared metadata. Computer mediated communication » - LIS590CMC. Graduate School of Library and Information Science. University of Illinois Urbana-Champaign. December 2004. <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>

Muller, M. J. 2007. Comparing tagging vocabularies among four enterprise tag-based services. In Proceedings of the 2007 international ACM Conference on Supporting Group Work (Sanibel Island, Florida, USA, November 04 - 07, 2007). GROUP '07. ACM, New York, NY, 341-350. DOI= <http://doi.acm.org/10.1145/1316624.1316676>

Murison, J. 2005. Messageboard topic tagging: user tagging of collectively owned community content. In Proceedings of the 2005 Conference on Designing For User Experience (San Francisco, California, November 03 - 05, 2005). Designing For User Experiences, vol. 135. AIGA: American Institute of Graphic Arts, New York, NY, 5.

Noll, M. G. and Meinel, C. 2007. Authors vs. readers: a comparative study of document metadata and content in the www. In Proceedings of the 2007 ACM Symposium on Document Engineering (Winnipeg, Manitoba, Canada, August 28 - 31, 2007). DocEng '07. ACM, New York, NY, 177-186. DOI= <http://doi.acm.org/10.1145/1284420.1284465>

Passant, A., Laublet, P. (2008). Meaning of a tag : a collaborative approach to bridge the gap between tagging and linked data. LDOW2008. Beijing, China.

Quintarelli, E., Resmini, A., Rosetti, L. (2007). Facetag : integrating bottom-up and top-down classification in a social tagging system.

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., and Millen, D. R. 2007. Getting our head in

the clouds: toward evaluation studies of tagclouds. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 995-998. DOI= <http://doi.acm.org/10.1145/1240624.1240775>

Rumelhart, D. E., & Norman, D. A. - « Accretion, tuning, and restructuring: Three modes of learning ». In J.R. Anderson (Ed.) *Semantic Factors in Cognition* (pp. 37-53).

Schenkel, R., Crecelius, T., Kacimi, M., Michel, S., Neumann, T., Parreira, J. X., and Weikum, G. 2008. Efficient top-k querying over social-tagging networks. In Proceedings of the 31st Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Singapore, Singapore, July 20 - 24, 2008). SIGIR '08. ACM, New York, NY, 523-530. <http://lsirpeople.epfl.ch/smichel/publications/sigir2008.pdf>

Smith, T. (2007). *Cataloging and You : Measuring the efficacy of a folksonomy for subject analysis* in Lussy, Joan, Eds. *Proceedings 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Milwaukee, Wisconsin.

Thom-Santelli, J., Muller, M. J., and Millen, D. R. 2008. Social tagging roles: publishers, evangelists, leaders. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 1041-1044. DOI= <http://doi.acm.org/10.1145/1357054.1357215>

Vuorikari, R. (2007). Can social information retrieval enhance the discovery and reuse of digital educational content?. In *Proceedings of the 2007 ACM Conference on Recommender Systems* (Minneapolis, MN, USA, October 19 - 20, 2007). RecSys '07. ACM, New York, NY, 207-210. DOI= <http://doi.acm.org/10.1145/1297231.1297276>

Wu, H., Zubair, M., and Maly, K. 2007. Collaborative classification of growing collections with evolving facets. In *Proceedings of the Eighteenth Conference on Hypertext and Hypermedia* (Manchester, UK, September 10 - 12, 2007). HT '07. ACM, New York, NY, 167-170. DOI= <http://doi.acm.org/10.1145/1286240.1286289>

Zacklad, M. (2005). *Introduction aux ontologies sémiotiques dans le Web Socio Sémantique*. In *actes de la conférence Ingénierie des Connaissances 2005*, Nice.

Zacklad, M. (2007). *Classification, thésaurus, ontologies, folksonomies : comparaisons du point de vue de la recherche ouverte d'information (ROI)*, In: CAIS/ACSI 2007, 35e Congrès annuel de l'Association Canadienne des Sciences de l'Information. *Partage de l'information dans un monde fragmenté : Franchir les frontières, sous la dir. de C. Arsenault et K. Dalkir*. Montréal : CAIS/ACSI, 2007. (Document en ligne sur < [http://www.cais-acsi.ca/proceedings/2007/zacklad\\_2007.pdf](http://www.cais-acsi.ca/proceedings/2007/zacklad_2007.pdf)>)