

L'indexation sociale au service des petites bibliothèques

Social tagging for small libraries

Samuel SZONIECKY (1)

(1) Paragraphe, Université Paris8
samuel.szoniecky@univ-paris8.fr

Résumé. Dans cette article nous nous interrogeons sur l'utilité de l'indexation sociale pour donner aux « petites » bibliothèques une plus grande visibilité dans les écosystèmes d'informations. Nous expliquons ensuite comment utiliser l'application Zotero pour mettre en place de façon simple, rapide et peu coûteuse une indexation sociale.

Mots-clés. Indexation sociale, bibliothèque, écosystème d'information.

Abstract. In this article we ask the utility of social tagging to give "small" libraries greater visibility in information ecosystems. We then explain how to use the tool Zotero to implement a simple, rapid and inexpensive social indexing.

Keywords. Social indexing, library, information ecosystems.

1 Petites Bibliothèques, Indexation sociale et intelligence collective

Face à des géants comme les bibliothèques nationales, les bibliothèques universitaires ou Google Book quelle place peut occuper dans l'écosystème d'information mondial qu'est le Web, une bibliothèque de quartier, celle d'un établissement scolaire voire même celle d'un particulier ? Ces « petites » bibliothèques ne possèdent pas forcément de système d'information centralisé ni même de site Web leur permettant de rendre visible leurs collections. Dès lors, on pourrait croire que c'est une affaire perdue d'avance mais c'est négliger la part importante que peut prendre l'indexation sociale dans la mise à disposition des données sur le Web et surtout l'utilité de ces « petites » bibliothèques dans le développement d'une intelligence collective.

Grâce au Web, il est aujourd'hui possible pour des millions de personnes de produire et diffuser des connaissances dans le monde entier. Pour organiser cette information on utilise un principe de base du classement d'une bibliothèque : décrire avec des mots-clés des documents pour ensuite pouvoir les retrouver grâce à ces mêmes mots-clés. Mais dans le cas du Web, vu le « déluge » d'information à traiter et son renouvellement continu, on fait appel à des robots pour automatiquement faire ce travail en utilisant des algorithmes plus ou moins complexes. On le sait, ces traitements automatiques ne sont pas sans problèmes (Rastier, 1991 ; Bachimont, 2007 ; Lévy, 2011). Notamment du point de vue technique, par exemple concernant le temps de calcul et les possibilités de

conservation, ou du point de vue théorique dans les rapports entre la forme d'une information et sa signification. Problèmes qui portent aussi sur des questions politiques et éthiques (Citton, 2010) car ces traitements sont en majorité la propriété d'entreprises privées qui gardent jalousement leurs « secrets de fabrication », ce qui ne permet pas de comprendre et d'évaluer les choix effectués par ces algorithmes. De plus, les algorithmes utilisés pour ces traitements fonctionnent pour un utilisateur moyen et ne prennent pas en compte les spécificités de tel ou tel profil, encore moins l'évolution des préoccupations d'un chercheur qui le soir peut ne plus avoir besoin des informations qui lui étaient nécessaires le matin. Pire encore, ces traitements automatiques sont parfois optimisés pour satisfaire une demande marketing comme par exemple le moteur de recommandation d'Amazon ou le moteur de recherche Google Books :

« Les ingénieurs qui manipulent l'algorithme de Google Books ont visiblement décidé de favoriser la recherche et le référencement de titres, la vente d'exemplaire, plutôt que la recherche dans l'ensemble du corpus, la mise en perspective de catégories, de mots-clefs... »¹

Pour répondre à ces problèmes du traitement automatique, des pratiques d'indexations sociales ont émergé sur le Web afin de ne pas laisser uniquement l'initiative aux machines et à ceux qui les construisent (Whittaker, 2011). Même si ces pratiques entraînent de nombreuses difficultés (Deuff, 2011), placer l'humain et ses facultés d'interprétation au cœur du processus de gestion des connaissances permettent de proposer des alternatives aux traitements automatiques pour fournir une indexation détaillée respectueuse d'une éthique correspondante à l'accès à toutes les connaissances pour tous.

En effet, face à l'omniprésence des géants du Web qui cherchent à homogénéiser au niveau mondial l'information pour pouvoir l'exploiter plus facilement, il devient important de cultiver et de préserver des spécificités informationnelles. Pour dire autrement, le point de vue global que proposent les grandes institutions publiques ou privées n'est pas le seul niveau d'organisation des connaissances possibles. La bibliothèque, de part son implantation physique dans les territoires offre un cadre fertile pour rassembler les individus afin qu'ils expriment leurs diversités par rapport à une organisation qui se veut globalement universelle. Pour ce faire, nous verrons comment mettre en place un processus simple d'indexation sociale dans l'écosystème d'information Web et comment la croissance de ce processus peut être assurée par un travail collaboratif. Notre but n'est pas de décrire ou d'analyser un système d'information pour gérer la totalité d'une bibliothèque ou d'un centre documentaire², nous voulons juste montrer comment, avec peu de moyen, on peut profiter de l'implantation territoriale d'une bibliothèque pour promouvoir des initiatives visant à rendre disponible sur le Web les ressources des « petites » bibliothèques. L'objectif consiste à développer une intelligence collective autour de l'indexation sociale pour participer à une redocumentarisation du monde (Pédaque, 2007).

2 Qu'est-ce que l'indexation sociale ?

Comme souvent dans le cas des technologies de l'information et de la communication, les pratiques qui émergent sont affublées de néologismes qui, comme des drapeaux, vont permettre de les reconnaître, s'y attacher ou au contraire les combattre. Le cas de l'indexation sociale ne fait pas exception à la règle puisqu'on trouve plusieurs termes, chacun mettant l'accent sur une dimension particulière de cette pratique.

2.1 Social bookmarking

Le social bookmarking décrit le processus collectif de mise en mémoire (Drechsler, 2009). Cette pratique est à la frontière entre le Web social et le Web de l'information (Limpens, 2010), il est en revanche assez loin de la sphère du Web documentaire de part l'utilisation libre des tags qui se différencie trop de l'utilisation de vocabulaire formalisé habituellement à l'œuvre dans les pratiques des professionnels de la documentation (thésaurus) ou du traitement informatique (base de données).

¹ <http://lafeuille.blog.lemonde.fr/2010/11/15/les-limites-de-google-books/> [Consulté 6 août 2012]

² Voir par exemple : <http://www.ori-oai.org/> [Consulté 2 août 2012]

En effet, l'indexation sociale utilise le plus souvent un vocabulaire de description qui n'étant pas normalisé le rend aussi difficile à traiter automatiquement que le langage naturel. Plus encore, l'utilisation d'acronyme rend son traitement automatique impossible. Voici par exemple une liste de tags, extrait de Delicious³, dont la signification est impossible à déduire automatiquement, ni même par un individu autre que celui qui l'a créé : x_ref125sm, vyaep, yycdata, q&a... Les difficultés d'analyse des tags concernent aussi le traitement multilingue des descriptions puisque les tags sont exprimés dans une langue particulière. Plus encore, pour un même corpus on peut trouver des tags qui renvoient au même concept mais exprimé avec des tags différents. Par exemple le tag « ontologie » et « ontology » renvoie au même concept mais exprimé d'une façon différente.

Au delà de ces problèmes liés globalement aux limites de traitement du langage naturel, l'indexation sociale est utile pour construire de nouveaux dispositifs de gestion de la mémoire collective mais aussi et peut-être surtout, pour mettre l'individu face à ses propres modes de fonctionnements cognitifs. En pratiquant l'indexation sociale, les utilisateurs vont pouvoir développer une gestion autonome des connaissances qui comme le remarque P. Lévy, est un enjeu pédagogique majeur :

« Une des fonctions les plus importantes de l'enseignement, du primaire aux différents cycles universitaires, sera donc de favoriser chez les étudiants la croissance durable de capacités autonomes de gestion personnelle des connaissances. Et cette gestion personnelle doit être conçue dès l'origine comme un processus élémentaire à partir duquel peuvent émerger les processus distribués d'intelligence collective qui l'alimentent en retour » (Lévy, 2011, p. 102)

2.2 Folksonomies

Le terme de folksonomie est un néologisme inventé par (Vander Wal, 2007) à partir des termes anglais « folk » et « taxonomy ». Il désigne des pratiques collectives d'indexation de pages Web à l'aide de mots-clefs appelés « tags » (Monnin, 2009). Ceux-ci sont choisis librement par chaque usager selon des règles explicites ou non. Les folksonomies sont fondées sur un triptyque composé par l'objet étiqueté, l'étiquetage, et l'identité de l'étiqueteur (Pirolli, 2009). L'objectif est de faire émerger de ces multiples descriptions une forme de consensus ou au contraire une situation de controverse. Ces pratiques se sont développées grâce à des outils facilitant le processus de description par tag (tagging). On peut citer par exemple www.delicious.com ou apparu plus récemment www.diigo.com qui offre l'avantage de « tagger » non seulement une page mais aussi un contenu à l'intérieur de la page.

Le principe des folksonomies consiste à partager avec d'autres utilisateurs le résultat de ses propres choix d'indexation pour comparer les différents points de vue sur une même ressource. Prenons par exemple, le tag « bibliothèque » utilisé pour référencer les pages Web sur ce sujet. Avec Delicious, à la manière d'un répertoire de signets, on peut récupérer l'ensemble des urls indexées avec ce tag, mais plus encore on peut savoir qui sont les autres utilisateurs qui ont indexé les mêmes pages. En analysant l'historique de cette indexation collective, on peut connaître les utilisateurs qui ont trouvé intéressantes les mêmes ressources avec

3 <http://www.delicious.com/> [Consulté 2 août 2012]

les mêmes mots-clefs mais aussi ceux qui diffèrent pour une même ressource. En bref, les folksonomies apportent des informations pertinentes notamment pour :

- trouver des interlocuteurs en accord avec un point de vue,
- dégager un consensus autour d'une description, (Loustau, 2011)
- évaluer l'activité de veille d'un individu par rapport à la fréquence ou à la précocité de son activité d'annotation,
- créer un système de recommandation (Kim, 2010)

Les données produites par les outils de folksonomie peuvent aussi être utilisées pour orienter un groupe d'individu vers des ressources particulières. C'est le cas par exemple du tag « DHYP » qui permet aux enseignants du département hypermédia de l'université Paris 8 de fournir une liste de sites utiles pour les étudiants. Des enseignants fournissent ainsi des liens aux étudiants en utilisant ce tag associé avec d'autres tags dédiés à la description du site. Un agrégateur de flux développé par B. Rieder permet ensuite de visualiser le résultat de cette pratique collective d'annotation⁴.

2.3 Social tagging

Si les folksonomies sont avant tout dédiées aux descriptions des ressources présentes sur le Web, le social tagging a un champ d'action beaucoup plus large. On pense par exemple aux initiatives mise en place par les chercheurs pour recenser le nombre de papillons dans les jardins afin d'analyser les contraintes d'évolution de ces animaux nécessaires pour la fertilisation de l'agriculture⁵. Ainsi, les bénévoles qui souhaitent participer à cette entreprise d'intelligence collective, identifient et comptent le nombre de papillon présent dans leur jardin en remplissant les fiches d'identification et en les envoyant ensuite aux chercheurs qui pourront ensuite les analyser.

En bref, le Social Tagging mobilise la bonne volonté des individus pour effectuer une multitude de descriptions simples qui pour des raisons de budget ou d'organisation ne pourraient pas être menées par un seul individu ou une institution. En ce sens, le *social tagging* est une pratique de « crowd sourcing » appliquée à la description de ressources par des tags. Il permet de faire émerger une intelligence collective en mettant en commun les capacités de discernement des individus. Dans le cadre d'une bibliothèque, grâce à cette intelligence collective, on pourrait envisager une indexation des contenus non plus seulement au niveau du livre mais au niveau de chaque page d'un ouvrage, voire même de chaque phrase.

3 Ranger, retrouver, naviguer, pour soi et les autres

Quelque soit l'appellation qu'on donne aux pratiques d'indexations sociales, elles relèvent toutes d'une mise en commun des capacités de catégorisation humaines pour un but bien défini. Les outils développés pour pratiquer cet exercice d'intelligence collective, permettent de réels succès tant par le nombre de

⁴ <http://hypermedia.univ-paris8.fr/> cliquer sur :: [delicious stream](#) [Consulté 2 août 2012]

⁵ <http://www.noeconservation.org/index2.php?rub=12&srub=31&ssrub=98&sssrub=100&goto=contenu> [Consulté 2 août 2012]

personnes concernées que par la quantité des ressources indexées. Toutefois, l'utilisation de vocabulaire libre rend parfois obscure et même inutile les efforts de chacun pour parvenir à une description commune. Face à ce danger de « babélisation », il apparaît nécessaire de tirer profit de l'expérience des métiers de la bibliothèque pour rendre encore plus efficace l'indexation collective (Alvarez-Cavazos, 2005).

3.1 Les standards de classification

Le premier point sur lequel, il est facile d'agir pour rendre l'indexation sociale plus efficace en terme d'interopérabilité, consiste à utiliser des outils de classification comme les taxonomies ou les thésaurus. Grâce à leur vocabulaire de référence, ces outils vont permettre la mise en place d'une terminologie standardisée permettant ensuite une interrogation plus fine des banques de données. Voir par exemple tout les travaux effectués autour de standards Marc ou BCR (<http://www.loc.gov/standards/>, <http://www.loc.gov/marc/>), mais aussi la spécification des normes ISO 2146:2009 (modèle) et ISO 8459:2009 (vocabulaire) ISO 20775:2009 (schéma XML), ISO 25964 (thésaurus). Dès lors, l'apparition d'un rapprochement entre le métier des bibliothèques et celui du Web, laisse envisager l'apparition de formats communs (Hudon, 2011), voir le rapprochement entre l'indexation sociale et les représentations structurées des connaissances (Limpens, 2010 ; Huynh-Kim-Bang, 2009 ; Zaklad, 2010).

Ces classifications utilisent des « metadatas », c'est à dire des grilles de description d'une ressource. Dans le monde des bibliothèques, la grille de description du Dublin Core⁶ est bien connue mais on peut modéliser autant de grille de description que l'on souhaite. Pour une classification qui a pour vocation de fournir des données pour les écosystème d'informations, il convient d'enrichir Dublin Core avec des grilles de description plus facilement exploitable dans le Web notamment pour l'utilisation de gadget ou de « mashup » comme googlemap, simile timeline, microdata...

Pour une organisation simple et générique, ces grilles de description peuvent être globalement organisées par rapport à quatre dimensions fondamentales de temps, de lieu, de personne et de concept.

- temps : outre les spécifications du Dublin Core concernant la définition des dates⁷, le format de simile timile⁸ semble particulièrement bien adapté pour décrire des informations temporelles même s'il existe d'autres formalismes utilisables suivant les besoins⁹.
- lieux : le standard de description des lieux qui s'impose aujourd'hui est le langage KML¹⁰ qui avec un format XML permet de représenter sur une carte tous les objets géographiques en deux et trois dimensions. L'intérêt du KML vient aussi des outils gratuits (Google Earth, Google Map) qui permettent de dessiner des cartes de façon simple et rapide.

6 <http://dublincore.org/> [Consulté 27 février 2012]

7 <http://library.princeton.edu/departments/tsd/katmandu/html/dccdate.html> [Consulté 2 août 2012]

8 http://code.google.com/p/simile-widgets/wiki/Timeline_EventSources [Consulté 2 août 2012]

9 <http://motools.sourceforge.net/event/event.html> [Consulté 2 août 2012]

10 <http://code.google.com/apis/kml/documentation/kmlreference.html> [Consulté 2 août 2012]

- personnes : pour la gestion des informations concernant les personnes, le formalisme FOAF¹¹ s'impose lui aussi de plus en plus comme un standard de fait. Là aussi des outils gratuits permettent d'éditer des informations sans avoir à maîtriser une syntaxe complexe¹².
- concepts : concernant la dimension conceptuelle, il est beaucoup plus difficile de proposer une solution établie vue le nombre de chercheurs et de solutions qui existent dans le domaine de l'organisation des connaissances (Polity, 2005) des ontologies (Bachimont, 2007 ; Renaud, 2008) ou du web sémantique (Folch, 2004 ; Gerbé, 2008 ; Zaklad, 2011). Il émerge néanmoins un consensus autour des formalismes proposés par le W3C : OWL¹³, RDF¹⁴ et SKOS¹⁵ (Bachimont & Al., 2011). Toutefois, l'emploi du langage naturel dans ce type de format ne résout pas totalement les difficultés d'exploitation automatique de ces formats, ce qui laisse envisager l'utilisation d'autres types de langage comme IEMML (Lévy, 2011 ; Noyer, 2010).

3.2 Les solutions de partage ouvert

Si effectivement un effort est fait pour standardiser la description des ressources, il convient ensuite de mettre à disposition ces standards, les catalogues qui les utilisent et les ressources indexées. La problématique du partage ouvert va consister justement à mettre à disposition du public ces données et ces méta-données ouvertes. D'abord apparues sous l'impulsion de Tim Berners Lee dans le champ de l' « Open Government Data », les données ouvertes (Open Data) sont de plus en plus présentes dans les politiques publiques que ce soit au niveau mondial avec les groupes de travail du W3C, au niveau continental où des directives européennes favorisent l'ouverture des données publiques¹⁶ ou au niveau national, comme en France avec la création de l'APIE¹⁷.

On peut définir l'Open Data par un certain nombre de critères obligatoires concernant les données proposées : complète, primaire, opportune, accessible, exploitable, non discriminatoire, non propriétaire, libre de droit. Il peut paraître difficile quand on commence un chantier d'ouverture de données de respecter l'ensemble de ces critères mais quelques bonnes pratiques permettent facilement de répondre à ces exigences sans pour autant se lancer dans un chantier titanesque. Dans cette perspective, il faut signaler l'initiative de la FING (Fondation Internet Nouvelle Génération) qui a produit un guide pratique très utile pour avoir une vision globale des problématiques complexes liées à l'ouverture des données et faciliter leurs déploiements¹⁸. Quoi qu'il en soit des problèmes concrets qui devront être pris en compte, dans une bibliothèque, l'ouverture des données peut être une chance pour développer de nouveaux services comme par exemple :

11 <http://www.foaf-project.org/> [Consulté 2 août 2012]

12 <http://wiki.foaf-project.org/w/Tools> [Consulté 2 août 2012]

13 <http://www.w3.org/TR/owl-features/> [Consulté 2 août 2012]

14 <http://www.w3.org/RDF/> [Consulté 2 août 2012]

15 <http://www.w3.org/2004/02/skos/> [Consulté 2 août 2012]

16 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0098:FR:HTML>
[Consulté 2 août 2012]

17 <https://www.apiefrance.fr/> [Consulté 2 août 2012]

18 <http://fing.org/?Guide-pratique-de-l-ouverture-des> [Consulté 2 août 2012]

- mettre en place « des fonctionnalités de réseau littéraire »¹⁹,
- affiner l'indexation des contenus,
- suivre l'évolution des acquisitions et leurs adéquations avec la demande des utilisateurs...
- développer des applications de bibliothéconomie (Merceur, 2011)

En fait, l'ouverture des données permet de multiplier les initiatives autour des bibliothèques en permettant à de nouveaux acteurs comme les fournisseurs de technologies ou les associations de citoyens de participer à son développement suivant leurs intérêts propres qu'ils soient économiques ou publics.

Les solutions de partage de données ouvertes rendent disponible un très grand nombre de données (Big Data) ce qui entraîne des problématiques propres d'un point de vue technique, cognitif et social (Boyd & Crawford, 2011 ; Rieder & Röhle, 2012). Une de ces problématiques concerne le design d'information. Ce qui est en jeu, c'est à la fois la modélisation technique d'un système informatique, l'implantation dans un écosystème d'information déjà existant, mais aussi la simplicité et l'efficacité des outils de gestion de l'information qui permettront en dernière instance de rendre accessible ou non ces données et leur enrichissement futur (Ertzscheid, 2002).

Comme nous l'avons vu plus haut, les questions des standards et des protocoles de communication sont relativement simples à régler. Il est en revanche beaucoup plus difficile de traiter les questions des interfaces graphiques qui permettront aux utilisateurs de s'approprier non seulement les informations disponibles mais aussi les méthodes qui leur permettront d'enrichir ces informations. Toutefois, l'avantage pour une bibliothèque d'utiliser des données ouvertes consiste justement à permettre l'externalisation de la conception des interfaces utilisateurs en utilisant des outils génériques (gadget, mashup...) déjà disponible pour la mise en forme de l'information ou en proposant des concours (cf. Apps4Ottawa²⁰) ou des ateliers créatifs sous la forme de « Barcamps », de forums créatifs, de « coding parties », de compétitions amicales, etc²¹...

C'est pourquoi il est important que la bibliothèque fournisse des données brutes qui respectent un format interopérable standard. Ainsi, en maîtrisant les sources d'information et leur format, les bibliothèques sont garantes de la qualité de l'information et de son interopérabilité (Borgman, 2003). Plus encore dans le cas d'une « petite » bibliothèque qui n'a pas les moyens de développer ou de faire développer ces propres outils et interfaces utilisateurs mais qui peut quand même participer à l'évolution globale de la bibliothèque :

« Le changement le plus profond pour les bibliothèques réside en effet dans les relations nouvelles de collaboration avec leurs usagers, de construction mutuelle d'outils de repérage et de dialogue avec des communautés. » (Carbone, 2011)

19 <http://www.bibliobsession.net/2010/05/05/bibliotheques-et-des-reseaux-sociaux-litteraires-ou-en-est-on/> [Consulté 2 août 2012]

20 <http://www.apps4ottawa.ca/> [Consulté 2 août 2012]

21 http://www.rescaufing.org/pg/blog/openid_82/read/52200/guide-pratique-de-louverture-des-donnees-publiques-territoriales [Consulté 2 août 2012]

4 Développer un écosystème d'information avec Zotero

Nous ne disposons pas aujourd'hui d'un état de l'art critique des nombreuses technologies et outils qui sont aujourd'hui disponibles sur le Web pour réaliser une indexation sociale²². Ce travail serait d'une très grande utilité mais dépasse largement le cadre de cet article. Nous présenterons donc un exemple qui se focalise sur l'outil que nous connaissons le mieux mais qui pourra sans doute être adaptée à d'autres outils. La solution qui nous apparaît aujourd'hui très simple pour mettre en place un système d'indexation sociale afin de créer une dynamique d'intelligence collective autour des « petites » bibliothèques, consiste à développer un écosystème d'information utilisant Zotero comme brique de base. Cette solution offre un double avantage, premièrement elle ne nécessite aucune compétence en développement informatique et deuxièmement elle est très économique puisque que le seul investissement nécessaire et celui du temps des personnes qui se consacreront au projet.

4.1 Qu'est-ce que Zotero ?

Zotero²³ est une application Open Source gratuite dont le développement a été initié par le centre Roy Rosenzweig pour l'histoire et les nouveaux médias. C'est un outil simple de gestion de bibliographies. Il permet de stocker, d'annoter ou de mettre en relation des documents. Il offre des fonctionnalités de création de communauté et de partage de ressources à l'intérieur de celle-ci. Zotero s'installe en quelques clics comme module complémentaire de Firefox²⁴. Un des grands avantages de cet outil est qu'il propose de nombreux formats d'exportation de données dont des standards : BibTex, Dublin Core, RDF. Ainsi, il est possible de rendre très facilement interopérables des bibliographies et les annotations qui leurs sont liées.

4.2 Indexation sociale d'une bibliothèque avec Zotero

On peut tout a fait utiliser Zotero pour indexer collectivement une bibliothèque même si celle-ci ne possède aucun catalogue informatique. Il suffit pour cela de fixer des règles de classification que les participants devront respecter pour effectuer ce travail d'intelligence collective. Voici par exemple les différentes étapes d'un scénario que l'on peut facilement mettre en place :

- télécharger et installer le module Zotero,
- s'inscrire au groupe de la bibliothèque,
- ajouter un nouveau document à sa bibliographie personnelle,
- annoter le document avec des mots clefs,
- enregistrer une citation du document,
- synchroniser sa bibliographie personnelle avec le serveur,
- exploiter les données de la communauté.

22 Pour une liste non exhaustive cf. <http://www.mkbergman.com/sweet-tools-simple-list/>
[Consulté 2 août 2012]

23 <https://www.zotero.org/> [Consulté 2 août 2012]

24 il suffit pour cela de le télécharger à l'adresse suivante : <http://www.zotero.org/download/>
[Consulté 2 août 2012]

On le voit bien, rien n'est très difficile dans ce scénario, il suffit juste de convaincre les utilisateurs d'une bibliothèque de participer et de partager leur travail. Toutefois, pour plus d'efficacité, on peut définir des règles plus précises pour l'annotation des documents, par exemple utiliser les mots-clés « citation » lorsqu'il s'agit d'un extrait d'un ouvrage, ou encore joindre des fichiers respectant les formats que nous avons défini plus haut. De même, concernant l'exploitation des données, on peut aller plus loin que la simple consultation du site de Zotero et faire développer des applications à partir de la mise à disposition de ces données ouvertes.

5 Indexation sociale des bibliothèques numériques pour une société de la connaissance

Aujourd'hui les bibliothèques souffriraient d'un manque de fréquentation de la part des lecteurs du fait que le passage par la bibliothèque n'est plus obligatoire pour celui qui cherche des références dans la connaissance humaine. Au point où certains se posent la question de la disparition des bibliothèques :

« L'accès facile et instantané à toute l'information mondiale par le web et les formats numériques peut renforcer l'image vieillotte de la bibliothèque et du bibliothécaire. »²⁵

Cette vision caricaturale de la bibliothèque et des bibliothécaires ne reflète pas tout à fait l'esprit de ces lieux où d'autres y voient les clefs d'une société de la connaissance :

« Au lieu de considérer la bibliothèque comme une forteresse isolée ou comme un tigre de papier, je voudrais la peindre comme le nœud d'un vaste réseau où circulent non des signes, non des matières, mais des matières devenant signes. La bibliothèque ne se dresse pas comme le palais des vents isolé dans un paysage réel, trop réel, qui lui servirait de cadre. Elle courbe l'espace et le temps autour d'elle et sert de réceptacle provisoire, de *dispatcher*, de transformateur et d'aiguillage à des flux bien concrets qu'elle brasse en continu. » (Latour, 1996, p. 23)

Si effectivement les bibliothèques ont un rôle fondamental à jouer dans le développement d'une société de la connaissance, nous pensons qu'il consiste à proposer un espace pour la transmission des méthodes d'organisation rigoureuse de l'information. En effet, tout comme de nombreux chercheurs (Amiel, 2010 ; Boutang, 2008 ; Citton, 2008 ; Jeanneret, 2007 ; Juanals, 2010 ; Lévy, 2011 ; Stiegler, 2005), nous pensons que la société de la connaissance passe nécessairement par une éducation de ces membres aux enjeux de la catégorisation et de l'interprétation. Or, les bibliothèques sont le lieu adéquat pour cet apprentissage car elles possèdent trois avantages prépondérants : celui de l'expérience séculaire pour l'organisation de l'information, celui de posséder des ressources documentaires disponibles physiquement pour l'indexation sociale et celui d'une proximité territoriale qui en font des acteurs prépondérants du développement culturel. En ce sens, la bibliothèque est le terrain de jeu idéal pour le développement de l'intelligence collective territoriale par l'indexation sociale.

25 <http://owni.fr/2011/04/27/les-bibliotheques-vont-elles-survivre> [Consulté 27 février 2012]

6 Bibliographie

- Alvarez-Cavazos, F., Garza-Salazar, D.A. & Lavariega-Jarquín, J.C., (2005). PDLib: personal digital libraries with universal access. Dans *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*. JCDL '05. New York, NY, USA: ACM, p. 365–365. Available at: <http://doi.acm.org/10.1145/1065385.1065468> [Consulté 2 août 2012].
- Amiel, P.H.L., (2010). *Ethnométhodologie appliquée: éléments de sociologie praxéologique*, Presses du LEMA.
- Bachimont, B. et al., (2011). Enjeux et technologies : des données au sens. *Documentaliste-Sciences de l'Information*, 48(4), p.24.
- Bachimont, B., (2007). *Ingénierie des connaissances et des contenus: Le numérique entre ontologies et documents*, Paris: Hermès science publications.
- Borgman, C.L., (2003). Personal digital libraries: Creating individual spaces for innovation. Dans *NSF Workshop on Post-Digital Libraries Initiative Directions*. Available at: http://www.sis.pitt.edu/~dlwksshop/paper_borgman.pdf [Consulté 2 août 2012].
- Boyd, D. & Crawford, K., (2011). *Six Provocations for Big Data*. SSRN eLibrary. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431#122782 [Consulté octobre 6, 2011].
- Boutang, Y.M., (2008). Du design capitalisme au capitalisme cognitif. Dans *Le design de nos existences: A l'époque de l'innovation ascendante*. Mille et une Nuits.
- Carbone, P., (2011). Les bibliothèques ou la mémoire mobilisée. *Les Cahiers du numérique*, Vol. 6(3), p.39-47.
- Citton, Y., (2010). *L'avenir des humanités: Économie de la connaissance ou cultures de l'interprétation ?*, Éditions La Découverte.
- Drechsler, M., (2009). *Les pratiques du socialbookmarking dans le domaine de l'éducation, Affordance sémantiques, socio-cognitives et formatives - Volume 1*. Université Paul Verlaine – Metz – École Doctorale Perspectives interculturelles □: écrits, médias, espaces, sociétés - Centre de recherche sur les médiations (ÉA 3476).
- Ertzscheid, O., (2002). *Le lieu, le lien, le livre. Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle*. Université de Toulouse II – Le Mirail Faculté des lettres et sciences humaines.
- Folch, H. & Habert, B., (2004). Langages de métadonnées pour Web(s) sémantique(s). Dans *Méthodes avancées pour les systèmes de recherche d'informations*. Paris: Hermès science publ. □: Lavoisier.
- Fourie, I., (2011). Personal information and reference management: Librarians' increasing creativity. *Library Hi Tech*, 29(2), p.387–393.
- Gerbé, O. & Mineau, G., (2008). Métamodélisation pour le Web sémantique. Dans *Revue des nouvelles technologies de l'information*. Cepaduès.

- Hudon, M., (2011). ISO 25964: vers une nouvelle norme pour l'organisation et l'accès à l'information et aux connaissances. Dans *8e Colloque international de l'ISKO France*. Stabilité et dynamisme dans l'organisation des connaissances. Université Charles-De-Gaulle Lille 3.
- Huynh-Kim-Bang, B., (2009). *Indexation de documents pédagogiques : fusionner les approches du Web Sémantique et du Web Participatif*. PhD thesis, Université Henri Poincaré-Nancy 1.
- Ihadjadene, M., (2005). Classification encyclopédiques et interface de recherche d'information. Dans *L'organisation des connaissances : approches conceptuelles*. Paris: L'Harmattan.
- Jeanneret, Y., (2007). *Y a-t-il (vraiment) des technologies de l'information ?*, Presses Universitaires du Septentrion.
- Juanals, B. & Noyer, J.-M., (2010). *Technologies de l'information et intelligences collectives*, Hermes Science Publications.
- Kim, H.N. et al., (2010). Social recommender systems on IEML semantic space. *Collective Intelligence Lab, Technical Report*.
- Deuff, O.L., 2011. Contrôle des métadonnées et contrôle de soi. *Etude de communication*, n° 36(1), p.23–38.
- Lévy, P., (2011). *La sphère sémantique : Tome 1, Computation, cognition, économie de l'information*, Hermes Science Publications.
- Latour, B., (1996). Ces réseaux que la raison ignore - laboratoires, bibliothèques, collections. Dans *Le pouvoir des bibliothèques. La mémoire des livres dans la culture occidentale*. p. pp. 23-46.
- Loustau, N.C., (2011). La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? *Etude de communication*, n° 36(1), p.39–52.
- Limpens, F., 2010. *Multi-points of view semantic enrichment of folksonomies*. Université de Nice - sophia Antipolis - Ecole doctorale STIC. <http://tel.archives-ouvertes.fr/tel-00530714/fr/> (Consulté 2 août 2012)
- Merceur, F., Le Gall, M. & Salaun, A., (2011). La bibliométrie: un nouveau cap pour une archive institutionnelle. Dans *Caught in the fishing net of information 14th Biennial EURASLIC Meeting 17-20 May, 2011, Lyon, France*.
- Monnin, A., (2009). Qu'est-ce qu'un tag? Entre accès et libellés, l'esquisse d'une caractérisation. Dans *Actes des 20es Journées Francophones d'Ingénierie des Connaissances*. Connaissance et communauté en ligne. Hammamet.
- Noyer, J.-M., (2010). Connaissance, pensée, réseaux à l'heure numérique. *Les Cahiers du numérique*, Vol. 6(3), p.187–209.
- Pédaque, R.T., (2007). *La redocumentarisation du monde*, Editions Cepaduès.
- Pirolli, F., (2009). Apports des folksonomies dans le cadre d'un processus de veille: vers la prise en compte des spécificités informationnelles. *7ème Colloque du chapitre français de l'ISKO Intelligence collective et organisation des connaissances*.

- Polity, Y., Henneron, G. & Palermi, R., (2005). *L'organisation des connaissances : approches conceptuelles*, Paris: L'Harmattan.
- Rastier, F., (1991). *Sémantique et recherches cognitives*, Paris: Presses universitaires de France.
- Renaud, S., (2008). Modélisation ontologique des connaissances□: Pour une lecture conceptuelle dans les manuels scolaires électroniques. Dans *Revue des nouvelles technologies de l'information*. Cépaduès.
- Rieder, B. & Röhle, T., (2010). Digital Methods: Five Challenges. Dans *The Computational Turn*. Swansea University.
- Stiegler, B., (2005). *De la misère symbolique : Tome 2, La catastrophe du sensible*, Editions Galilée.
- Vander Wal T., (2007), Site sur la folksonomie, <http://www.vanderwal.net/random/entrysel.php?blog=1635> (consulté le 11 février 2012)
- Wallis, J.C. et al., (2010). Digital libraries for scientific data discovery and reuse: from vision to practical reality. Dans *Proceedings of the 10th annual joint conference on Digital libraries*. JCDL '10. New York, NY, USA: ACM, p. 333–340. Available at: <http://doi.acm.org/10.1145/1816123.1816173> [Consulté 2 août 2012].
- Whittaker, S., (2011). Personal Information Management: from information consumption to curation. Dans *Annual review of information science and technology* (ARIST). p. 3–62.
- Zacklad, M., (2010). Évaluation des systèmes d'organisation des connaissances. *Les cahiers du numérique*, Vol. 6(3), p.133–166.