

Métadonnées générées par les utilisateurs

Le cas des pratiques d'indexation sociale

Les pratiques liées aux usages des outils du web social ont contribué à inscrire progressivement la création et l'utilisation de métadonnées dans les habitudes de nombreux internautes, de façon consciente ou non. Les indispensables opérations de description des contenus créés, partagés et diffusés sur le web – sans lesquelles les ressources concernées resteraient difficilement accessibles – poussent les utilisateurs à devenir naturellement producteurs « d'informations sur l'information ». Qu'il s'agisse de décrire le contenu d'un billet posté sur un blog, de caractériser un document sur une plateforme de diffusion (Flickr, Picasa, YouTube, Dailymotion, etc.) ou encore de partager un ensemble de signets à l'aide d'un service dédié (Delicious, Diigo, etc.), l'internaute est amené à décrire et caractériser, de fait, un ensemble de ressources documentaires. Ce processus est aujourd'hui couramment qualifié d'indexation sociale en référence aux aspects collaboratifs et communautaires de l'exercice ainsi qu'à ses finalités. Ces pratiques présentent de nombreuses particularités tant dans leurs modalités de mises en œuvre que dans la nature des résultats obtenus. L'exploitation des corpus de métadonnées ainsi produits ouvre de nombreuses perspectives, parfois fort éloignées des raisons ayant conduit à leur création. Nous proposons, dans un premier temps, d'en rappeler les principales caractéristiques, notamment par leur mise en perspective du point de vue des pratiques documentaires plus classiques. Dans un second temps, nous présenterons plusieurs contextes d'utilisation de ces métadonnées afin de mettre en exergue la richesse et la diversité des applications envisageables.

Tags, folksonomies et pratiques documentaires

L'indexation sociale n'a pas vocation à se substituer aux pratiques professionnelles établies dans le domaine de l'information-documentation dont la rigueur reste gage d'efficacité. Elle en diffère sur de nombreux aspects. Les objectifs poursuivis ne sont pas nécessairement les mêmes. Si l'indexation a pour objet la description, au moyen d'outils documentaires¹, du contenu informationnel d'une ressource, l'indexation sociale s'inscrit dans une démarche plus vaste mais moins clairement définie. L'utilisation par l'internaute de descripteurs, généralement désignés par le terme de tags, énoncés sans aucune contrainte formelle ou sémantique ouvre un champ de possibilités beaucoup plus étendu qu'une approche purement descriptive. Différents travaux de recherche ont mis en évidence le fait que, outre ce travail de description, les utilisateurs de systèmes à base de folksonomies² ont recours aux tags afin de qualifier, de commenter ou d'annoter des ressources (en utilisant des tags comme « à lire », « à faire », « intéressant », etc.). Le travail d'indexation est traditionnellement le fruit de

l'activité de praticiens de l'information-documentation formés tant à l'accomplissement de cette tâche qu'au respect des normes et des bonnes pratiques propres à leur champ disciplinaire. Dans le cas de l'indexation sociale, les corpus de métadonnées générés par les utilisateurs résultent d'une accumulation de représentations individuelles exprimées librement par des individus qui ne sont pas nécessairement sensibilisés aux bénéfices qu'apporte une démarche rigoureuse. L'usage de tels outils entraîne le déplacement de certaines pratiques documentaires vers un public novice en la matière. La très forte hétérogénéité des résultats ainsi obtenus constitue bien entendu un obstacle à partir du moment où ils sont considérés uniquement à travers le prisme de la documentation. En revanche, ils tirent toute leur richesse de l'accumulation et de la mutualisation réalisées au sein d'espaces communautaires en ligne.

L'essor de ce type de systèmes collaboratifs d'indexation de masse sur Internet s'explique en raison de :

- leur facilité d'utilisation permettant au plus large public d'appréhender leurs fonctionnalités très rapidement et sans contraintes ;
- leur capacité à créer et nourrir des échanges d'information entre utilisateurs notamment par les mécanismes de suggestion de tags ;
- l'adaptabilité du vocabulaire utilisé pour décrire et rechercher les ressources ;
- l'ensemble de solutions qu'ils proposent pour répondre aux besoins individuels d'organisation et de stockage de l'information.

Le corollaire de ces propriétés est l'absence de règles de syntaxe, l'omission du traitement des ambiguïtés langagières ainsi que l'absence de prise en compte des flexions des différents descripteurs. Ces aspects sont particulièrement problématiques dans un environnement multilingue qui amplifie considérablement leur fréquence d'apparition. Autant d'obstacles très handicapants dans le cadre d'un travail d'indexation rigoureux que l'usage d'outils spécifiques, comme les langages documentaires, permettrait d'éviter ou du moins de limiter. Doit-on pour autant considérer que les métadonnées ainsi générées ne constituent qu'un ensemble d'informations désordonné, imprécis et au final peu signifiant ? En dehors de leurs finalités premières, il nous semble que ces sources d'information offrent des potentialités d'usage multiples tant dans le domaine de la veille informationnelle et stratégique que pour l'enrichissement de bases de données partagées. Elles ouvrent également d'intéressantes perspectives sur l'évolution des outils de recherche d'information qui accordent toujours plus de place à la dimension sémantique des liens existant entre les unités informationnelles repérées sur les réseaux.

Potentialités d'usage des métadonnées générées par les utilisateurs

Les folksonomies sont caractérisées par une très faible inertie face à la nouveauté. L'utilisation de néologismes et l'expression libre de nouveaux concepts permettent une circulation et un partage rapides de l'information et l'émergence de nouvelles acceptions, généralement liées à l'usage. Le caractère labile des métadonnées mobilisées contribue à en faire de précieux indicateurs de tendances. L'étude approfondie de l'évolution temporelle d'un corpus folksonomique, notamment des variations de fréquences d'utilisation des descripteurs et de l'apparition de cooccurrences entre descripteurs, permet de définir des processus d'identification des « signaux faibles » essentiels au travail de veille.

Ces mêmes folksonomies peuvent également participer à l'enrichissement de données structurées préexistantes. Pour des organismes partageant des ressources en ligne, le fait d'associer les usagers à la production de métadonnées permet d'envisager une démarche de « redocumentarisation » particulièrement intéressante. Les fonds partagés et, par extension, les bases de données permettant de les gérer peuvent ainsi être enrichis d'une couche supplémentaire d'information ayant pour objet, au moyen d'une personnalisation accrue, l'augmentation des interactions entre usagers et contenus numériques.

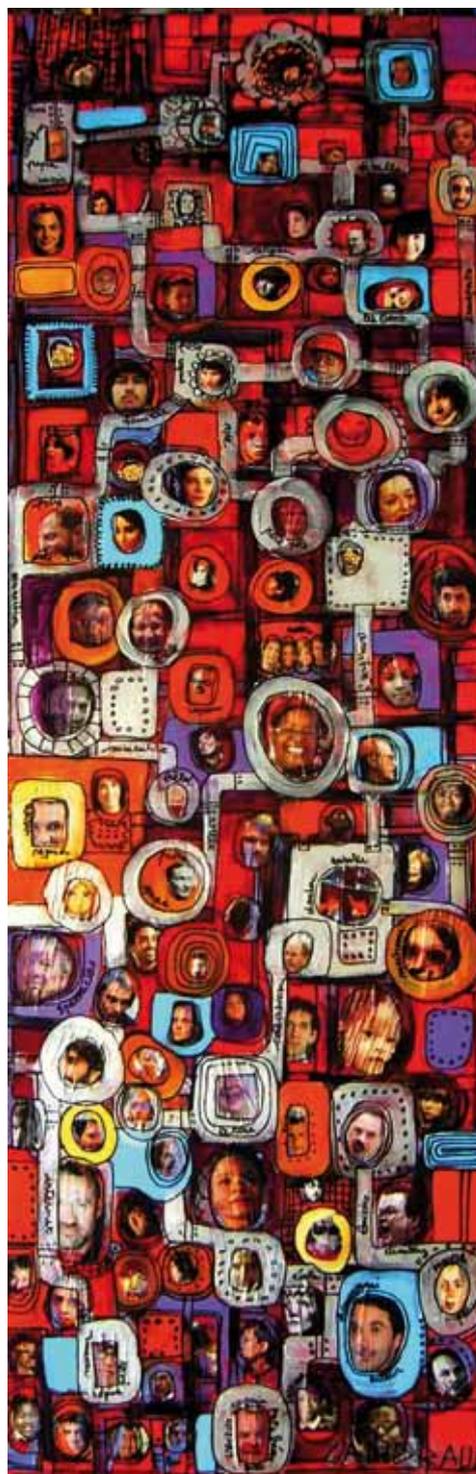
Enfin, si elles sont parfois présentées comme deux approches antagonistes, la conception d'un « web social » privilégiant les applications ouvertes, communautaires et partagées, et celle d'un « web sémantique » organisé, rigoureux et formalisé, ne semblent plus aujourd'hui s'exclure totalement. Elles peuvent même nourrir d'intéressants points de convergence. Les futurs outils de recherche et d'accès à l'information, en particulier pour leurs fonctionnalités de personnalisation et de filtrage des résultats proposés, auront tout intérêt à tirer le meilleur parti à la fois des schémas de métadonnées formalisés retenus comme éléments constitutifs d'une architecture raisonnée de l'information mais également de la richesse et de la diversité apportées par les métadonnées générées par les utilisateurs.

Fabrice Pirolli

¹ Lexiques, listes d'autorités ou thésaurus.

² Néologisme issu de la contraction des termes « folks » et « taxonomies » désignant le principe d'une indexation libre, de masse et partagée.

Fabrice Pirolli est maître de conférences en sciences de l'information et de la communication, Université de Bourgogne, et chercheur au laboratoire CIMEOS EA 4177
fabrice.pirolli@iut-dijon.u-bourgogne.fr



Quelques indications bibliographiques

- Casemajor Loustau, Nathalie, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », *Études de communication*, n° 36, 2011, p. 39-52.
- Filippi, Dominique, « Délicieuses folksonomies : comment gérer les signets de la bibliothèque dans Delicious ? », *Le web 2.0 en bibliothèque. Quels services ? Quels usages ?*, coll. Bibliothèques, éditions du Cercle de la librairie, 2009, p. 85-113.
- Francis, Élie, Quesnel, Odile, « Indexation collaborative et folksonomies », *Documentaliste-Sciences de l'Information*, n° 1/2007 (Vol. 44), ADBS, p. 58-63.
- Pirolli, Fabrice « Pratiques d'indexation sociale et démarches de veille informationnelle », *Études de communication*, n° 36, 2011, p. 53-66.