

	<p style="text-align: right;">Date : 17/07/2007</p> <p>Bibliothèque 2.0 et contenu produit par l'utilisateur Que peuvent faire les usagers pour nous ?</p> <p>Patrick Danowski Staatsbibliothek zu Berlin, Berlin Allemagne</p> <p>Traduction : Pierre Tribhou Bibliothèque Inguimbertaine, Carpentras, France ptribhou@yahoo.fr</p>
Meeting:	113 National Libraries with ICABS and Information Technology
Simultaneous Interpretation:	Yes
<p style="text-align: center;">WORLD LIBRARY AND INFORMATION CONGRESS: 73RD IFLA GENERAL CONFERENCE AND COUNCIL 19-23 August 2007, Durban, South Africa http://www.ifla.org/iv/ifla73/index.htm</p>	

Résumé :

Bibliothèque 2.0 et contenu généré par l'utilisateur sont deux notions étroitement liées. Au cours de l'exposé je définirai brièvement ces deux notions. Je donnerai deux exemples de projets où contenu généré par l'utilisateur et bibliothèques interagissent. Le fichier d'autorité des noms de personnes coopératif allemand, fruit de la coopération entre Wikipedia et le *Personennamendatei*, sera le premier. Le second sera Wikisource où les usagers mettent en ligne des textes électroniques. Un autre domaine important de contenu généré par l'utilisateur est le partage d'étiquettes (*social tagging*) grâce auquel les usagers indexent différentes ressources. Et si les usagers faisaient tellement dans l'avenir, y aurait-il encore de la place pour les bibliothécaires ? Mais dans l'avenir, l'utilisateur et le bibliothécaire deviendront partenaires et la bibliothèque fournira l'infrastructure : la bibliothèque 2.0.

Chers collègues, je suis très heureux d'être invité en tant que jeune professionnel. Je suis bibliothécaire à la Staatsbibliothek zu Berlin, Bibliothèque nationale allemande, et je fais partie du conseil d'administration de la section germanophone de Wikimedia. J'ai étudié l'informatique et j'ai un master de sciences de l'information et des bibliothèques. Je ne vous exposerai pas un projet concret, je voudrais plutôt vous présenter deux notions intéressantes, qui sont étroitement liées.

Mon exposé sera composé de cinq parties. En premier lieu je voudrais examiner brièvement les notions de bibliothèque 2.0 et de contenu généré par l'utilisateur ; que désigne-t-on par là ? Je souhaiterais ensuite vous donner deux exemples de projets qui montrent que le contenu généré par l'utilisateur peut-être intéressant pour les bibliothèques et comment on peut s'en servir. Mon premier exemple est celui de la coopération entre la Deutsche Nationalbibliothek et le projet Wikipedia en langue allemande, où les fichiers d'autorité des noms de personnes sont utilisés. Mon second exemple est le projet Wikisource. Dans le Wikisource de langue allemande, livres et manuscrits numérisés sont retranscrits par la communauté alors que les bibliothèques n'ont souvent pas l'argent pour faire cela. Je montrerai comment le projet fonctionne et comment les bibliothèques pourraient offrir de meilleures conditions, requises pour ce type de projet, que ce qu'elles sont actuellement. Il nous faudra certainement aussi nous pencher sur le partage d'étiquettes (*social tagging*). Enfin, je voudrais conclure sur ce que les usagers peuvent faire pour nous et comment nous devrions peut-être repenser diverses choses. Je voudrais discuter de la question de savoir s'il y aura de nouvelles tâches pour les bibliothèques dans le contexte du contenu généré par l'utilisateur.

Mais d'abord je commencerai par la notion de « bibliothèque 2.0 ». Le concept de la « bibliothèque 2.0 » a été forgé par Michael Casey¹. Il n'y a pas de définition exacte de la notion de « bibliothèque 2.0 » et les discussions qui eurent lieu à ce sujet se prolongent encore. Mais il est admis qu'elle n'emprunte pas que son nom à la notion à la mode de « Web 2.0 »² ; elle lui emprunte aussi ses principes et techniques. Je peux juste vous en donner ici un très petit premier aperçu.

Les technologies Web 2.0 potentiellement utiles à la bibliothèque peuvent être les weblogs³, les fils RSS⁴, les podcasts et les wikis. Toutes ces techniques ont en commun que les usagers peuvent très aisément et rapidement publier des textes, des sons ou, grâce à des services plus récents comme Youtube, des vidéos. Un autre type de service très intéressant pour les bibliothèques est le partage d'étiquettes⁵ grâce auquel les ressources peuvent être indexées par l'utilisateur avec des mots-clés libres. Certains premiers OPAC utilisent déjà cette technique. Un autre aspect très important du Web 2.0 est qu'il ne relie pas seulement des ressources : il relie aussi les gens, un aspect dont l'illustration la plus connue est MySpace⁶. Peut-être serait-il intéressant d'utiliser aussi certaines de ces nouvelles techniques pour « iflanet »⁷.

Les principes du Web 2.0 qui peuvent aussi s'appliquer aux bibliothèques sont l'interactivité et la possibilité pour l'utilisateur de collaborer. Le retour d'information de l'utilisateur devrait en outre jouer un rôle dans la création de nouveaux services et la réforme des anciens. Les services devraient s'appuyer sur des standards ouverts. Lawrence Lessig parle de culture du remix⁸ et les bibliothèques doivent y prendre part. Les bibliothèques doivent donc faire en

¹. Sur son blog Library Crunch <http://www.librarycrunch.com>

². O'Reilly, Tim (2005) : What is Web 2.0 ? Design Patterns and Business Models for the Next Generation of Software. En ligne : <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/What-is-web-20.html>.

³. Voir l'article « Blog » sur Wikipedia pour une définition : <http://en.wikipedia.org/wiki/Blog> Dernière visite 28.04.07 [article en français : <http://fr.wikipedia.org/wiki/Blog>]

⁴. Voir l'article « RSS » sur Wikipedia pour une définition : <http://en.wikipedia.org/wiki/RSS> Dernière visite 28.04.07 [article en français : http://fr.wikipedia.org/wiki/Really_Simple_Syndication]

⁵. Voir l'article « Collaborative tagging » sur Wikipedia pour une définition : http://en.wikipedia.org/wiki/Collaborative_tagging Dernière visite le 28.04.07

⁶. MySpace <http://www.myspace.com> Dernière visite 28.04.07

⁷. IFLA net <http://www.ifla.net> Dernière visite 28.04.07

⁸. Lessig, Lawrence (2006): Free(ing) Culture for Remix. In: Lutterbeck, Bernd; Gehring, Robert; Bärwolf, Matthias (Hg.): Opensource Jahrbuch 2006. Berlin: Lehmanns Media. En ligne:

sorte que leur contenu (généré par la bibliothèque) ou leurs services puissent être remixés en applications composites (*mash-ups*).

Pour la bibliothèque Web 2.0 cela signifie qu'il nous faut repenser les services des bibliothèques numériques. Nous n'avons pas à tout inventer par nous-mêmes. Nous pouvons aller là où se trouvent déjà les usagers. Ceci signifie que les bibliothèques doivent considérer quels services Web 2.0 sont utilisés par les usagers et penser comment la bibliothèque peut utiliser ces services pour aider les usagers ou les rapprocher de la bibliothèque. Les bibliothèques doivent créer des outils (widgets, c'est-à-dire gadgets windows) qui peuvent être utilisés sur des portails comme la page de démarrage de Google « iGoogle »⁹. Un second exemple en est l'usage de Second Life¹⁰ où des bibliothèques américaines ont créé Info Island¹¹, la zone dans Second Life propre aux bibliothèques et fournisseurs d'information. Les bibliothèques peuvent aussi jouer un rôle dans des services comme del.icio.us¹². Elles peuvent importer des guides de ressources par sujet ou des listes de liens par exemple et les étiqueter avec les mots-clés qu'elles utiliseraient normalement. Un usager qui travaille sur ce sujet découvrira ces guides plus facilement grâce à leur environnement interactif et pourra s'inscrire à cette bibliothèque ressource. Une autre façon d'y prendre part est de partager leur connaissance sur Wikipedia. Le partage devrait être un aspect important de la bibliothèque 2.0. Le partage de contenus et de métadonnées devrait être la base pour de nouveaux services de bibliothèques pour faire face au défi du Web 2.0. Les services d'applications composites sont encore une tendance nouvelle du Web 2.0 qui a beaucoup de succès.

La seconde notion présente dans mon titre est celle de « contenu généré par l'utilisateur ». Les nouveaux outils du Web 2.0 comme les wikis et les blogs permettent aux usagers de devenir très rapidement des éditeurs. Amazon et LibraryThing permettent aux usagers d'écrire des recensions de livres. Il y a beaucoup de nouveaux services où les usagers peuvent indexer différentes ressources comme des liens, des photos et des vidéos. Mais le projet le mieux connu et peut-être le plus grand projet de contenu généré par l'utilisateur est celui de Wikipedia, l'encyclopédie libre en ligne qui est disponible en plus de 100 langues. Wikipedia est libre, ce qui ne veut pas dire seulement en accès libre et gratuit, ou comme Richard Stallman le dit « free like beer »¹³, gratis comme la bière lors d'une fête ; Wikipedia est aussi libre d'être reprise, ou pour reprendre les mots de Stallman « free like speech », libre comme la parole. Chaque communauté peut avoir ses propres règles, son propre projet. Depuis 2005 Wikipedia en langue allemande a un projet en commun avec le Personennamendatei (PND) : le fichier d'autorité des noms de personnes coopératif allemand. Ce projet est mon premier exemple. Quand je parle de Wikipedia, je fais référence au projet germanophone.

Dans le Wikipedia germanophone se trouvent plus de 600.000 articles dont environ 20% d'articles consacrés à des personnes. Des métadonnées pour ces articles de personnes furent

<http://www.opensourcejahrbuch.de/portal/scripts/download?article=osjb2006-09-03-en-lessig.pdf> Dernière visite 28.04.07

⁹. Page d'accueil IGoogle <http://www.google.com/ig?hl=en> Dernière visite 28.04.07 [Page en français : <http://www.google.com/ig?hl=fr>]

¹⁰. Page d'accueil de Second Life <http://www.secondlife.com> Dernière visite 28.04.07

¹¹. Page d'accueil d'Info Island <http://www.infoisland.org> Dernière visite 28.04.07

¹². Heller, Lambert : Libraries that del.icio.us In netbib Weblog

<http://log.netbib.de/archives/2006/10/05/libraries-that-delicious> Dernière visite 28.04.07

¹³. Stallman, Richard : « Why « Open Source » Misses the Point of Free Software ». in: In: Lutterbeck, Bernd; Gehring, Robert; Bärwolf, Matthias (Hg.): Opensource Jahrbuch 2007. Berlin: Lehmanns Media. En ligne: <http://www.opensourcejahrbuch.de/portal/scripts/download?article=osjb2007-00-02-en-stallman.pdf> Dernière visite 28.04.07

aussi créées en 2005. La cause en est un autre projet. Un éditeur a produit une version hors-ligne de Wikipedia sur CD, mais il a eu un petit problème, parce qu'il souhaitait changer l'ordre des noms et prénoms pour des raisons d'extraction de données. Dans Wikipedia la forme « Prénom Nom » est habituellement utilisée. Mais pour le CD il souhaitait le modifier selon la forme usitée dans les encyclopédies : « Nom, Prénom ». L'éditeur a travaillé avec la communauté et nous avons créé les métadonnées pour les articles de personnes.

Ce fut le point de départ et certaines personnes dans la communauté Wikipédia pensèrent qu'il serait utile que les données puissent être reliées aux fichiers d'autorité des noms de personnes. Comme vous le voyez sur le tableau, nous avons un contenu comparable : nom, formes alternatives rejetées du nom, lieux et dates de naissance et de mort, enfin une courte description de la personne. La Deutsche Nationalbibliothek donna son accord pour mettre en place un projet reliant son fichier d'autorité des noms de personnes aux personnes dans Wikipedia. Ils durent faire quelques modifications dans leur base de donnée et un développeur de la communauté créa un outil pour comparer les entrées du fichier d'autorité et les métadonnées de Wikipedia.

Voici la procédure : l'outil-Persondata recherchait via une URL spécialement créée dans l'OPAC de la Bibliothèque Nationale Allemande où les fichiers d'autorité étaient localisés. Les métadonnées complètes pour une personne ou plus qui correspondait au nom étaient renvoyées depuis l'OPAC. L'outil présente à l'utilisateur les métadonnées de Wikipedia qui sont comparables aux entrées du PND. L'utilisateur décide s'il y a adéquation et si l'identifiant, le Numéro PND, doit être ajouté à l'article. Au moyen d'un « template » fut créé un lien entre l'article et le catalogue de la Bibliothèque Nationale Allemande. En suivant ce lien on peut retrouver de la documentation de et sur cette personne dans la catalogue de la Bibliothèque Nationale Allemande. Cela est possible parce que la recherche permet de n'extraire que les noms de personnes différenciées.

Cette coopération fut un vrai succès. En très peu de temps (environ 2 semaines) plus de 22.000 articles ont eu leur lien. Je rappelle que ces liens furent faits par des usagers de Wikipedia. Durant ce projet une longue liste de souhaits de correction des fichiers d'autorité fut créée par la communauté, lorsque les membres trouvaient qu'il y avait des erreurs dans le fichier d'autorité. Les lecteurs de Wikipedia peuvent à présent trouver très aisément plus de documentation sur une personne. Le chemin d'une ressource électronique à une ressource hors-ligne est à présent très court et un usager de Wikipedia peut devenir un usager de bibliothèque.

Le second exemple dont je voudrais parler est celui du projet Wikisource. Je parlerai ici aussi du projet germanophone. Je le ferai parce que ce projet aussi est un peu spécial. Le but du projet Wikisource est de fournir des textes-sources originaux. Il est clair que ce sont seulement les textes qui ne sont pas sous copyright, mais du domaine public. La base technique pour ce service est aussi un wiki. Grâce à une petite extension la relecture est très facile parce que l'image se trouve sur la fenêtre de rédaction. Les bibliothèques peuvent bénéficier de ce service. Les usagers peuvent faire ce que les bibliothèques ne peuvent se permettre de faire. Le résultat sera certainement encore libre de telle sorte qu'il pourra être réutilisé par la bibliothèque. On débat souvent pour savoir si les bibliothèques devraient élaborer leurs propres services, comme ce dernier, sur leurs sites Web. Mais cela aurait un énorme désavantage. Chaque bibliothèque devrait constituer sa propre communauté, de telle sorte que la masse critique nécessaire de personnes pour faire une communauté qui fonctionne serait plus difficile à atteindre. Il vaudrait mieux permettre la réutilisation des matériaux.

Pour le moment les métadonnées de Wikisource ne sont pas parfaites. Elles sont en texte brut sur un tableau à côté de chaque page de titre. Mais pour le moment il n'y a pas moyen de télécharger seulement les métadonnées pour les réutiliser dans d'autres bases de données, par exemple des catalogues de bibliothèques. Nous avons déjà réfléchi à un moyen de transformer les métadonnées, mais nous n'avons pas le temps de mettre cela en pratique. Tous les gens qui travaillent sur ce projet sont bénévoles. Mais même si nous le mettions en œuvre nous aurions encore un problème. La forme d'entrée ne serait pas normalisée. Nous verrons comment résoudre ces problèmes. Si vous avez des idées, prenez contact avec moi. Je souhaiterais présenter une affiche à l'IFLA sur ce projet et j'aimerais vous inviter pour en parler.

Comme je l'expliquais au début, les usagers indexent différentes sortes de médias : des liens, des images, des vidéos et du texte. Toutes ces choses sont étiquetées.

Les étiquettes (*tags*) sont des mots-clés libres. Il peut donc y avoir beaucoup d'étiquettes qui signifient la même chose ou sont potentiellement synonymes. Et les étiquettes peuvent ne pas être seulement des mots-clés qui décrivent un objet. Elles peuvent aussi être des étiquettes fonctionnelles comme « alire » qui a simplement une signification pour son usager, ou une étiquette comme « ifla2007PD » qui décrit les ressources d'un exposé de sorte qu'elle a une signification pour un petit groupe.

De plus en plus apparaissent dans le Web2.0 des services qui utilisent l'étiquetage. Les exemples les plus connus en sont del.icio.us, LibraryThing¹⁴, connotea¹⁵, citeUlike¹⁶, flickr¹⁷, et YouTube¹⁸. En tant qu'utilisateur de ces services on peut étiqueter des ressources mais aussi voir comment les autres étiquettent une ressource et même qui a sauvegardé cette ressource. On peut chercher d'autres ressources qui ont été étiquetées avec un mot-clé qu'on choisit. Cela aide à trouver des gens qui ont les mêmes centres d'intérêt. C'est ce que je voulais dire au début en affirmant que le Web 2.0 ne relie pas seulement des ressources mais aussi des gens.

J'aimerais donner un court exemple d'un service. CiteUlike est un service pour gérer de la documentation, particulièrement des articles de périodiques. On y trouve des fonctionnalités intéressantes comme la possibilité de consulter la table des matières d'un périodique imprimé et de télécharger les métadonnées d'articles intéressants dans un compte personnel. On peut étiqueter chaque article et voir combien de personnes ont eux aussi sauvegardé cet article. Je peux aussi rentrer un nouvel article manuellement en le tapant ou importer un fichier BibTex¹⁹.

Pour un meilleur aperçu les étiquettes peuvent être visualisées en un nuage d'étiquettes. Dans ces nuages les étiquettes les plus souvent utilisées et peut-être les plus pertinentes seront de plus grande taille. Un tel nuage se trouve sur le côté droit de l'exemple. Le KUG²⁰, l'OPAC de la bibliothèque universitaire de Cologne, utilise cette technique pour visualiser les mots-clés contrôlés qui sont utilisés.

¹⁴. Page d'accueil de Library Thing <http://www.librarything.com>

¹⁵. Page d'accueil de Connotea <http://www.connotea.com>

¹⁶. Page d'accueil de CiteUlike <http://www.citeulike.com>

¹⁷. Page d'accueil de Flickr <http://www.flickr.com>

¹⁸. Page d'accueil de YouTube <http://youtube.com>

¹⁹. Voir l'article Wikipedia pour plus d'information : <http://en.wikipedia.org/wiki/BibTex> [En français : <http://fr.wikipedia.org/wiki/BibTex>]

²⁰. OPAC KUG <http://kug.ub.uni-koeln.de>

Nous avons vu que les usagers peuvent faire beaucoup pour une bibliothèque, de l'apport d'informations nouvelles à l'amélioration de l'accès aux objets numérisés (à travers leur transcription), en passant par l'indexation des ressources. Se pose peut-être alors une question : que feront les bibliothécaires dans l'avenir ? Nous avons vu combien les documents présents sur le net sont variés. Les usagers peuvent les mettre en ligne très facilement. Cela signifie aussi qu'il y a des documents de grande qualité, qui peuvent être utilisés en environnements d'apprentissage et qu'il y a des documents comme le film ou les photos des dernières vacances. L'étiquetage partagé est très puissant mais a aussi ses problèmes. Des mots destinés à attirer l'attention (*buzzwords*) sont par exemple utilisées parfois un peu trop souvent. Les bibliothécaires peuvent créer des guides par sujet indiquant les meilleures ressources, en considération non seulement de leur contenu mais aussi de l'utilisation qu'on peut en faire. Ainsi le bibliothécaire peut filtrer et organiser les matériaux avec leurs mots-clés et systèmes de classification normalisés. Cette classification par le bibliothécaire n'est pas un point d'aboutissement, tout au contraire : elle peut être le fondement du Web 3.0, le web sémantique²¹. Ces systèmes seront plus largement acceptés et mieux reçus s'ils sont de format ouvert et libres d'usage.

Maintenant que nous sommes revenus à notre point de départ, et nous comprenons que bibliothèque 2.0 ne signifie pas seulement utilisation de la technologie du web 2.0 ; bibliothèque 2.0 signifie aussi que la bibliothèque a un rôle important à jouer dans le monde du Web 2.0. Le potentiel est vraiment énorme, alors soyez prêts, devenez un bibliothécaire 2.0, soyez ouverts aux contributions de vos usagers. J'espère vous avoir montré que les usagers peuvent soutenir l'effort des bibliothécaires. Il nous faut créer l'infrastructure pour cela, et cette infrastructure est une partie importante de la bibliothèque 2.0. Les usagers peuvent devenir nos nouveaux partenaires, ils peuvent nous aider à améliorer nos services.

J'espère que j'ai éveillé votre curiosité, parce que c'est toujours le point de départ vers quelque chose de nouveau.

Essayons.

Merci.

²¹. Voir l'article Wikipedia pour plus d'information : http://en.wikipedia.org/wiki/Semantic_web [En français : http://fr.wikipedia.org/wiki/Web_s%C3%A9mantique]