

## Tous catalogueurs ! Exploiter les métadonnées « maison » pour les bibliothèques de recherche

*French translation of the original paper: “We are All Catalogers Now: Leveraging Do It Yourself Metadata for Research Collections”.*

*Translated by: Pauline Moirez, Bibliothèque nationale de France, Département de l'information bibliographique et numérique, Paris, France.*

*Le texte de ce document a été traduit en français et peut contenir des différences par rapport au texte original. Cette traduction est fournie à titre indicatif uniquement.*

### **Rebecca L. Lubas**

Claremont Colleges Library, Claremont, California, USA

E-mail address : [rebecca\\_lubas@cuc.claremont.edu](mailto:rebecca_lubas@cuc.claremont.edu)



This is a French translation of “*We are All Catalogers Now: Leveraging Do It Yourself Metadata for Research Collections*” copyright © 2014 by Pauline Moirez. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:

<http://creativecommons.org/licenses/by/3.0/>

---

### **Résumé :**

*Les professionnels du patrimoine culturel sont confrontés aujourd’hui à un foisonnement de richesses et de possibilités, qui ouvrent des questionnements inédits sur les collections et les métadonnées. Des ressources sont produites sur le terrain, 24 heures sur 24, 365 jours par an, grâce à l’évolution des appareils d’enregistrement numériques qui deviennent davantage portables et accessibles, ainsi qu’à la multiplication des sites de gestion de contenu. Les événements ne sont pas seulement enregistrés en temps réel, mais ils sont également archivés en temps réel. Les appareils d’enregistrement personnels et les méthodes d’archivage se perfectionnent, et enregistrent automatiquement des métadonnées.*

*Le présent article décrit les moyens dont disposent les institutions pour récupérer des ressources afin de soutenir leur activité de recherche. Il contient des exemples, des conseils pratiques et une analyse des implications sur le long terme. Dans cet article, nous examinons comment les bibliothèques et les autres institutions qui conservent des collections destinées à la recherche peuvent exploiter les ressources produites par tout un chacun à des fins historiques. Tout le monde produit des ressources, et nous, en tant que bibliothécaires, sommes bien placés pour en assurer la préservation et les inscrire dans un contexte culturel afin que les chercheurs les découvrent.*

**Mots-clefs:** métadonnées, mots-clefs, ressources produites par les usagers, collections numériques

---

## Tous catalogueurs ! Exploiter les métadonnées « maison » pour les bibliothèques de recherche

### Le contexte actuel

Avant l'avènement du Web, des médias sociaux et des smartphones, les bibliothèques et les services d'archives collectaient des documents sur l'histoire et les événements locaux du mieux qu'ils pouvaient, en ciblant une collection donnée. Coupures de presse, brochures, affiches et photographies dessinaient les contours d'événements qui avaient un jour fait l'actualité et qui appartenaient désormais à l'histoire. Le volume de ces documents était également beaucoup plus faible qu'aujourd'hui, car tout le monde ne disposait pas de matériel d'impression ou de photographie. Les ressources matérielles peuvent ne rejoindre les collections que plusieurs années après un événement, quand les souvenirs se sont estompés, et que les créateurs de ces ressources ne sont plus eux-mêmes que de l'histoire ancienne.

La généralisation de l'utilisation des équipements désormais disponibles a transformé et multiplié les ressources qui décrivent et documentent les événements. Chaque citoyen est un journaliste potentiel, et il n'est nul besoin de compétences technologiques spécialisées. Les créateurs de ressources se sont perfectionnés : il leur arrive de comprendre et d'utiliser les licences *Creative Commons*, d'ajouter des métadonnées descriptives aux ressources, et de sauvegarder leurs données, pour ne donner que quelques exemples. Comment les professionnels du patrimoine culturel peuvent-ils tirer parti de ces montagnes de ressources ? Doivent-ils même seulement s'y risquer ? Cela relève-t-il de la responsabilité des bibliothèques et services d'archives (Lampert 75) ? Il y a un véritable défi à trouver des méthodes pour faire face au volume de documents disponibles. Il y a cependant un risque à ne pas prendre en considération ces documents. Dans quelques décennies, les donateurs n'auront plus de boîtes de photographies à confier à un service d'archives : il leur arrivera parfois de télécharger les images et autres ressources dans le *cloud*, parfois de les traiter, parfois de les sauvegarder, mais pas toujours. Ce n'est pas parce que l'on trouve des outils facilement accessibles pour archiver ses propres ressources que les individus vont davantage s'appliquer à traiter leurs documents dans la durée qu'ils ne s'occupaient de leurs vieux albums de photographies, surtout s'ils doivent les migrer sur de nouveaux appareils ou plateformes.

Les bibliothèques sont particulièrement bien placées pour répondre à ce dilemme. Les bibliothécaires possèdent des compétences organisationnelles et des outils qui leur permettent d'assurer la collecte, le traitement et la diffusion des documents en produisant une bien plus grande valeur ajoutée que s'ils étaient restés « dans leur milieu naturel » sur le web. Nous sommes issus d'une culture de la normalisation, et il s'agit d'une compétence utile pour répondre à ce dilemme, qui est également une opportunité. En effet, si les usagers peuvent créer des ressources et interagir avec celles-ci plus facilement, cela ne doit pas seulement signifier que nous avons davantage de documents à collecter, mais également que de nouvelles possibilités d'impliquer nos usagers s'offrent à nous, car la technologie fait tomber les obstacles et donne aux usagers l'impression d'être associés d'une façon plus personnelle. Le travail ne s'achève plus avec la collecte et la conservation, les collections sont désormais dynamiques (Lampert 81). Elles ont leur vie propre, elles ne restent plus figées.

Ces possibilités provoquent une tension nouvelle chez les bibliothécaires et les archivistes. Les institutions patrimoniales ont deux missions : assurer la conservation des ressources pour le futur et les mettre à disposition des usagers du temps présent (Van Hooland 717). La communauté archivistique accepte de mieux en mieux d'utiliser des métadonnées produites par les usagers (Gorzalski 20). Les possibilités qui s'ouvrent à l'accomplissement de la première mission augmentent avec le volume des ressources, de même que le potentiel offert à l'accomplissement de la seconde augmente avec les techniques permettant la création de métadonnées par les utilisateurs.

### **Se préparer pour une récupération efficace**

La question la plus complexe est sans doute : par où commencer ? Quelles ressources votre bibliothèque collecte-t-elle d'ores et déjà, et quel est le périmètre de vos collections spécialisées ? Ensuite, avec Internet dans son ensemble comme territoire de chasse, comment affiner votre choix ?

Il est plus simple de gérer le problème en vous posant la question : « Qu'aurais-je voulu collecter s'il s'était agi de ressources imprimées » ? Ainsi, par exemple, si votre bibliothèque collecte des documents sur la politique locale (affiches de campagnes, photos de rassemblements politiques, etc.), c'est un bon point de départ.

On peut rechercher des ressources sur des réseaux existants afin d'identifier plus facilement d'une part les personnes qui possèdent des ressources et d'autre part des documents déjà regroupés par sujet. Le projet IMLS Collections et ressources numériques a testé cette solution en s'appuyant sur des groupes Flickr (Jett). Flickr, comme d'autres services similaires de partage et d'archivage de photographies, peut être un bon point de départ : les membres peuvent en effet utiliser les galeries de photographies, les albums et les groupes pour rassembler des ressources autour de lieux ou de sujets. En exploitant des réseaux existants, vous avez également la possibilité d'utiliser une compétence des bibliothécaires : s'appuyer sur la compréhension de l'organisation d'un système pour améliorer l'exploration des ressources.

Il peut être utile d'utiliser les médias sociaux comme si vous faisiez une recherche personnelle, surtout si vous cherchez à récupérer des ressources sur des événements actuels. Si vous connaissez les *hashtags* du moment sur Twitter, c'est une clef pour identifier les ressources et les communautés pertinentes pour le sujet que vous cherchez à traiter. Les usagers ont en effet tendance à décrire des collections (Stvilia 64). C'est pourquoi il est sans doute plus efficace de rechercher des ensembles de ressources pertinentes que d'essayer de découvrir une pépite, mais une seule à la fois.

Lorsque vous mettez en place vos modalités de récupération, il convient de vous assurer qu'elles soient clairement structurées, mais qu'elles restent souples (Lampert 74). Comme les technologies évoluent rapidement, vous devez être capables de réagir face aux nouvelles tendances qui vont faciliter la création et la mise à disposition des ressources.

Après avoir établi des circuits pour identifier les ressources candidates, vous pouvez élaborer une stratégie proactive de dépôt de collections. La première étape est de découvrir où se trouvent les créateurs de ressources, puis de mettre en place des circuits leur permettant de venir d'eux-mêmes vers votre collection.

## **Les questions juridiques**

Lorsque vous identifiez un ensemble de documents que vous voudriez collecter, il faut examiner les aspects juridiques pour savoir si vous pouvez poursuivre ou non. Cela peut constituer une étape très chronophage dans le processus, et qui ne produit pas toujours des résultats exploitables. Les questions juridiques constituent à elles seules une bonne raison de commencer par un projet de faible envergure. Il peut être également préférable de cibler des réseaux qui collectent et mettent à disposition directement les données juridiques, surtout s'il s'agit d'une communauté qui encourage l'utilisation des licences *Creative Commons*. Vous pouvez ainsi souvent vous épargner d'avoir à contacter les créateurs des ressources. Il vous faudra également considérer si votre institution est à l'aise vis-à-vis de la réutilisation, et à quel point : par exemple, voulez-vous ne collecter que des documents dotés de la licence la plus ouverte (la licence Attribution, qui permet à chacun de réutiliser et modifier la ressource) ou bien êtes-vous prêts à intégrer des documents sous licence Attribution - Pas de modifications, qui impose de ne pas modifier la ressource ?

Il convient d'établir des règles qui encadrent ce que votre bibliothèque ou votre institution est prête à accepter ou, en d'autres termes, de mettre dans la balance les contrôles que vous êtes prêts à effectuer en regard de l'intérêt que vous avez à intégrer ces documents dans vos collections. Si vous pensez qu'il y aura probablement des documents qui méritent un investissement en temps afin de contacter l'auteur et de lui demander de vous ouvrir les droits, vous pouvez lui demander cette autorisation, mais prévoyez d'encadrer ces échanges. Votre institution peut se doter pour ce genre de contacts d'un modèle de lettre avec des formules type, et ensuite vous pouvez l'adapter autant que nécessaire. Vous devez garder à l'esprit que le temps que vous passerez à obtenir les permissions doit être proportionné à la priorité du sujet traité pour vos collections. Un événement populaire annuel ne nécessite sans doute pas le même degré d'investissement qu'un événement exceptionnel : si un propriétaire de ressources n'accorde pas son autorisation, il peut y avoir de nombreuses autres ressources au contenu similaire.

## **Le dépôt et la remise des documents**

Il vous faudra décider comment conserver les documents, en lien avec l'exploration de ces documents. Vous devrez définir les formats que vous voulez collecter et maintenir, et choisir un système en fonction des moyens techniques et des ressources dont vous disposez pour le gérer et l'entretenir.

Puisque nous connaissons l'importance des métadonnées pour l'exploration des ressources, vous devrez choisir un système qui collecte les métadonnées utiles.

Il convient de trouver un équilibre entre la qualité du système et les ressources nécessaires pour le gérer. Il ne faut pas oublier que vous pourrez sans doute prendre en charge un plus grand nombre de collections constituées par les usagers si vous en confiez le contrôle à leurs créateurs. Par exemple, la plateforme Omeka permet aux créateurs de ressources de créer facilement des expositions virtuelles de collections numériques, avec un soutien limité. Vous pourrez fournir aux communautés d'usagers aide et conseils pour qu'ils puissent créer par eux-mêmes des collections en ligne qui soient nettement plus élaborées que sur les médias sociaux.

## **Les métadonnées descriptives**

Vous récupérerez des ressources accompagnées de leurs propres métadonnées. Bien entendu, vous voudrez utiliser les normes que votre institution a choisies ; il est nécessaire d'appliquer nos valeurs et normes professionnelles dont la cohérence est la base de l'organisation et de l'exploration des ressources. Mais que faire des métadonnées qui accompagnent l'objet ?

Il peut s'agir de métadonnées techniques fournies automatiquement (par exemple, par l'appareil photographique utilisé), mais aussi de métadonnées descriptives généralement beaucoup plus subjectives et personnelles.

Le créateur d'une ressource peut utiliser dans la description de l'image des mots qui relèvent d'un jugement de valeur ou d'un point de vue subjectif sur l'événement ou l'objet représenté. Dans quelle mesure ce point de vue subjectif fait-il partie intégrante de la ressource ? Le service d'archives ou la bibliothèque qui utilise cette ressource a-t-il la responsabilité intellectuelle de conserver ce point de vue, autant que celle de fournir des métadonnées et une analyse objectives ? Quelles considérations éthiques sont en jeu si une vue « objective » se trouve être opposée à celle du créateur de la ressource ? Le point de vue du créateur de la ressource peut apporter une véritable valeur pour les recherches portant sur cet objet informationnel.

Il est utile de permettre au créateur des ressources d'ajouter des mots-clefs, ou encore de ne pas se débarrasser des métadonnées produites par les usagers lorsque l'on récupère des ressources : l'utilisateur peut connaître la date, le lieu et les personnes représentées sur une image, par exemple. La situation se complique si vous avez des images qui ne représentent pas simplement « La Grand-Rue en 1910 » mais un événement politique actuel. Les annotations des usagers peuvent alors porter sur les opinions du créateur de la ressource. Dans les bibliothèques, nous sommes traditionnellement très attentifs à la neutralité de nos métadonnées, mais il peut y avoir une valeur historique à conserver également des commentaires subjectifs. Les commentaires en temps réel constituent une culture nouvelle susceptible d'enregistrer l'histoire en train de s'écrire. C'est pourquoi il est impératif de mettre en place une politique de gestion des métadonnées subjectives avant même de commencer la collecte, et cette politique joue un rôle dans la finalité même de la collection. Il ne faut cependant pas oublier que le créateur de la ressource peut en refuser la modification, et l'a peut-être indiqué dans sa licence d'utilisation.

Votre public potentiel peut vous apporter beaucoup, par exemple en identifiant le contenu précis d'une vidéo : les usagers qui ajoutent des mots-clefs peuvent créer des « liens profonds » vers des moments spécifiques d'une vidéo (Madden 699). On peut décrire les ressources avec un niveau de granularité que les bibliothécaires n'ont pas les moyens d'atteindre.

## **La pérennisation des ressources**

Vous avez identifié les ressources, résolu les problèmes juridiques, contacté le créateur si c'était nécessaire, et vous avez collecté et intégré les ressources dans votre système... et maintenant ? Il vous faut examiner si la conservation pérenne de ces ressources relève de votre responsabilité, ou si cette responsabilité repose sur les épaules des créateurs des ressources. Les questions habituelles de gestion des collections s'appliquent-elles de la même façon à ces ressources, ou doit-on envisager une nouvelle approche ?

Si vous définissez les formats que vous allez collecter avant même de commencer, cela vous aidera à répondre à cette question. Il ne faut pas oublier que si vous collectez une grande variété de formats, il se peut que vous ayez à vous soucier de mettre à disposition des logiciels pour lire ces formats, ce qui est encore une autre ressource dont il faut assurer la conservation. Mais vous pouvez également convertir les ressources sélectionnées dans un format normalisé pour étendre leur durée de vie, même si l'objet original est altéré.

C'est plus simple si l'on trouve des ressources sur des réseaux existants, car la plupart d'entre eux utilisent déjà des formats et tailles de fichiers normalisés. En d'autres termes, vous pouvez profiter d'une partie du travail déjà réalisé par ces réseaux et par les créateurs des ressources.

### **Aide-mémoire pour la collecte de ressources créées par les usagers**

Cet aide-mémoire est un outil destiné à vous aider à réfléchir aux règles que vous allez mettre en place pour poser les bases d'une collecte réussie de ressources produites par les usagers, avec des métadonnées produites par les usagers.

#### Récupération des ressources

- Sélectionnez les sujets ou les événements que vous allez surveiller et rechercher
- Ciblez les sites de gestion de contenu et les médias sociaux qui ont une politique d'ouverture des données
- Déterminez les formats que vous allez collecter

#### Dépôt des ressources

- Choisissez un système informatique
- Analysez les capacités et limites de votre système

#### Questions juridiques

- Déterminez quelles restrictions de droits vous êtes prêts à gérer
- Décidez si vous allez contacter les créateurs de ressources, ou si vous n'allez utiliser les ressources que si les droits sont déjà explicitement précisés
- Élaborez une convention que vous conclurez avec les créateurs de ressources

#### Métadonnées

- Définissez les métadonnées dont vous avez besoin pour toutes les ressources
- Évaluez vos besoins en ressources et méthodologie pour permettre la modification et l'ajout de métadonnées
- Déterminez quelles modifications seront possibles pour les métadonnées créées par les usagers

## Implication des usagers

- Décidez si vous proposerez un outil ou demanderez une inscription sur votre site pour permettre l'ajout de mots-clefs et de liens
- Déterminez si vous modifierez les annotations des usagers
- Rédigez une lettre type pour permettre aux annotateurs de céder les droits d'utilisation et de modification de leurs contributions
- Collectez les statistiques d'usage

## Pérennisation

- Déterminez pendant combien de temps vous offrirez aux usagers la possibilité de faire des commentaires
- Décidez si vous voulez archiver les annotations

## Conclusion

Les métadonnées recèlent un immense potentiel, elles constituent une porte d'entrée vers des collections plus actuelles, vivantes et complètes. Les bibliothèques et les services d'archives peuvent fournir un niveau inégalé de contextualisation de leurs ressources (Bullen 31). L'implication des usagers et les technologies se développent comme jamais auparavant et apportent de nombreux avantages : les usagers actuels peuvent avoir plus d'interactions avec les archives pour écrire l'histoire ; et les futurs chercheurs pourront accéder à des documents gardant la mémoire instantanée des événements. Nous pouvons désormais impliquer les usagers et créer des archives enrichies pour les chercheurs. Par le passé, les métadonnées nous appartenaient et n'appartenaient qu'à nous. Maintenant que le grand public est capable de créer des métadonnées, profitons-en pour garder de meilleures traces de l'histoire.

## Références bibliographiques

Bullen, Andrew. "The 'Long Tale': Using Web 2.0 Concepts to Enhance Digital Collections." *Computer in Libraries* [En ligne], 28.9 (2008): 31-35. Web. 27 Avril 2014

Gorzalski, Matt. "Examining User-Created Description in the Archival Profession." *Journal of Archival Organization* [En ligne] 11.1-2 (2013): 1-22. Web. 27 Avril 2014

Jett, Jacob, Megan Senseney, & Carole L. Palmer. "A Model for Providing Web 2.0 Services to Cultural Heritage Institutions: The IMLS DCC Flickr Feasibility Study." *D-Lib Magazine* [En ligne], 19.5/6 (2013). Web. 27 Avril 2014

Lampert, Cory, & Su Kim Chung. "Strategic Planning for Sustaining User-Generated Content in Digital Collections." *Journal of Library Innovation* [En ligne], 2.2 (2011): 74-93. Web. 27 Avril 2014

Madden, Amy, Ian Ruthven, & David McMenemy. "A classification scheme for content analyses of YouTube video comments." *Journal of Documentation* [En ligne], 69.5 (2013): 693 – 714. Web. 27 Avril 2014

Stvilia, Besiki & Corinne Jørgensen. "User-generated collection level metadata in an online photo-sharing system." *Library & Information Science Research* [En ligne], 31.1 (2009): 54-65. Web. 27 Avril 2014

Van Hooland, Seth, Eva Mendez Rodriguez & Isabelle Boydens. "Between Commodification and Engagement: On the Double-Edged Impact of user-Generated Metadata within the Cultural Heritage Sector." *Library Trends* [En ligne], 59.4 (2011): 707-720. Web. 27 Avril 2014