

Résumé

Mise en œuvre d'une plate-forme sociale permettant à un groupe d'experts d'indexer un fonds de photographies de la Bataille de Normandie (Seconde Guerre mondiale) selon les standards de métadonnées dans le domaine des images numériques fixes : origine et conditions d'utilisation des photos, description de la plate-forme sociale (Flickr), techniques de métadonnées utilisées, description du processus éditorial et documentaire, comparaison avec d'autres projets similaires et analyse de l'activité de redocumentarisation.

Mots-clés

indexation sociale, métadonnées, IPTC, Flickr, redocumentarisation, photos, Seconde Guerre Mondiale, retour d'expérience

Une plate-forme sociale pour la redocumentarisation d'un fonds iconographique

Preprint

« Document numérique et Société »

<http://doc-soc.fr/>

les 17 et 18 novembre 2008

Publié dans

Traitements et pratiques documentaires : vers un changement de paradigme ?

Actes de la deuxième conférence Document numérique et Société, 2008

Sous la direction d'Evelyne Broudoux et Ghislaine Chartron

ADBS, Collection : Sciences et techniques de l'information, 2008

Auteur

Patrick Peccatte

Soft Experience

peccatte@softexperience.com

www.softexperience.com

blog : blog.tuquoque.com

Les usages de partage collaboratif du Web 2.0 ont permis l'apparition de nouveaux modes de traitement de l'information. Dans le domaine de la redocumentarisation de fonds photographiques, différentes pratiques sont apparues dont la plupart font appel aux usagers et témoignent du rapide renouvellement des acteurs documentaires ainsi que de la mutation des outils de traitement. L'article décrit, dans ce contexte de transformation profonde du champ documentaire, le retour d'expérience du projet *PhotosNormandie*.

1. Présentation du projet PhotosNormandie

Le projet *PhotosNormandie* a pour but d'améliorer l'indexation d'un fonds de photos historiques sur la Bataille de Normandie qui s'est déroulée du 6 juin à fin août 1944 durant la Seconde Guerre mondiale. Il s'agit d'un travail collaboratif à finalité patrimoniale. Le projet est actif depuis janvier 2007 et utilise la plate-forme de partage de photos et vidéos *Flickr* ⁽¹⁾.

1 - Description du projet : www.flickr.com/people/photosnormandie/

Accès direct aux photos : www.flickr.com/photos/photosnormandie/

Tutoriel *Comment débiter sur PhotosNormandie ?* : www.ranes1944.org/PhotosNormandie/Aide.html

Je remercie tous les intervenants sur *PhotosNormandie* pour leur participation compétente et désintéressée, en particulier Michel Le Querrec, co-responsable du projet, qui réalise un travail considérable et sans quoi rien n'aurait été possible.

Origine des photos

Le site *Archives Normandie 1939-1945* (www.archivesnormandie39-45.org) est un service public du Conseil Régional de Basse-Normandie mis en place en 2004 à l'occasion de la commémoration du soixantième anniversaire de la libération de la Normandie. Ce site propose près de 13000 photos de l'Occupation, de la Libération et de la Reconstruction.

Chaque photo possède une référence de la forme *pxxxxxx* (où *xxxxxx* est un numéro à six chiffres) permettant un accès direct. Il est également possible d'effectuer une recherche sur un sujet, un lieu, un personnage, etc.

Ces photos sont aussi disponibles sur le site commercial *Archives de Guerres* à l'adresse www.archives-de-guerres.fr ⁽²⁾.

Conditions d'utilisation des photos

Les photos du site *Archives Normandie 1939-1945* qui proviennent des Archives Nationales des États-Unis et du Canada sont déclarées *libres de droits* selon les termes suivants :

Toutes les images dont les descriptifs comportent la mention de copyright "libre de droits" et qui ont été numérisées, peuvent être téléchargées en haute définition (fichiers numériques TIFF en 300 DPI) et utilisées librement, avec pour seule contrainte l'obligation d'inscrire, lors de toute utilisation, la mention : "Conseil Régional de Basse-Normandie / National Archives USA" pour les images dont l'origine indiquée sur le descriptif est "NA-USA"

"Conseil Régional de Basse-Normandie / National Archives Canada" pour les images dont l'origine indiquée sur le descriptif est "NA-CA".

Source : www.archivesnormandie39-45.org/conditions.html

Il est important de comprendre qu'une image *libre de droit* n'est pas, en général, gratuite, et n'est pas non plus du domaine public. Le *libre de droit* (*royalty-free*) est une forme d'autorisation simplifiée de vente où le prix est forfaitaire. Les conditions d'utilisation sont fixées par l'éditeur et l'acheteur peut utiliser la photo autant de fois qu'il le souhaite tant que ces conditions sont respectées. Ce mode de licence se distingue du *droit géré* (*rights managed*) qui est une solution d'achat sur mesure en fonction de l'utilisation envisagée de la photo. Les photos en haute définition proposées par le site *Archives Normandie 1939-1945* sont donc à la fois gratuites et libres de droit, et comme telles soumises à la condition d'utilisation que nous avons rappelé (mention d'origine obligatoire).

Cette condition concerne 2763 photos disponibles en haute définition qui sont reprises dans le projet *PhotosNormandie* sur *Flickr* afin d'améliorer leur description ⁽³⁾.

Pourquoi ce projet ?

Il est apparu en effet que si la qualité des numérisations proposées sur le site *Archives Normandie 1939-1945* est correcte, il n'en est pas de même des légendes qui comportent de nombreuses inexactitudes et incohérences. Certaines erreurs sont très importantes du point de vue historique ou même tout simplement descriptif et diminuent grandement l'intérêt documentaire de cette collection accessible au grand public ⁽⁴⁾. À la fin de l'année 2006, Michel Le Querrec – un passionné de

2 - *Archives Normandie 1939-1945*, *Archives de Guerres* et *PhotosNormandie*, utilisent les mêmes références *pxxxxxx* pour identifier les photos ; le lecteur peut donc facilement comparer sur ces trois sites les descriptions des différentes photos mentionnées dans cet article.

3 - Les conditions sont stipulées sur la page de description du projet et les titulaires des droits sont mentionnés dans le champ *Copyright* de chaque photo haute définition. La licence d'utilisation transposée sur *Flickr* est de type *Creative Commons* : paternité (vous devez citer le nom du titulaire des droits, le Conseil Régional de Basse-Normandie et National Archives USA ou National Archives Canada), pas d'utilisation commerciale, pas de modification.

4 - Par exemple les photos de la série p012514 à p012524 prises à la pointe du Hoc aux premiers jours du Débarquement sont localisées au Berghof (*sic*), et la p012015 figure un bord de mer situé à Condé-sur-Noireau, à 60 km de la mer environ ! D'autres descriptions, imprécises ou minimales, témoignent aussi d'un manque de connaissance manifeste du sujet et d'une pratique documentaire déficiente - ainsi celle de la p012377 légendée « Sainte Mère Église pendant la

l'histoire de la Bataille de Normandie, co-responsable du projet – et moi-même avons décidé d'essayer d'améliorer les descriptions de ces photos.

Après quelques essais sur différents sites de partage de photos, la plate-forme *Flickr* s'est avérée la plus adaptée pour le travail que nous souhaitions réaliser, essentiellement par ses possibilités de discussion sur une photo et sa capacité à exploiter un sous ensemble significatif du standard de métadonnées IPTC/IIM (voir ci-dessous).

2. Description succincte de la plate-forme de partage de photos Flickr

Flickr est un site web de partage de photos et de vidéos. Les fonctionnalités de base, limitées en capacité de stockage, sont gratuites. Les fonctionnalités non limitées sont payantes (compte "Pro" à 24,95 dollars par an). Le site permet à la fois un stockage privé ou public ; les photos du projet *PhotosNormandie* sont bien entendu publiques.

Flickr a été développé par la société canadienne Ludicorp en 2002 et racheté par Yahoo! en 2005. Le site héberge actuellement plus de 2 milliards de photos.

Chaque photo peut être décrite par un titre, une description, des *tags*, des commentaires, des notes associées à des zones choisies de la photo. Le titre et la description sont modifiables uniquement par le gestionnaire du compte ayant téléchargé la photo tandis que les *tags*, les notes et les commentaires peuvent être ajoutés aux photos par tout visiteur si le gestionnaire du compte l'a autorisé. La recherche de photos s'effectue sur les titres, les descriptions et les *tags*, mais pas sur les commentaires ou les notes. Il est aussi possible de *géotagger* les photos, c'est-à-dire de leur attribuer des coordonnées géographiques, et d'effectuer des recherches cartographiques sur ces photos.

L'utilisateur peut organiser ses photos par albums et classeurs (ensembles d'albums). Il est aussi possible de créer ou de participer à des groupes d'utilisateurs réunis sur des thématiques très variées et d'ajouter des photos à ces groupes. *Flickr* maintient aussi des flux RSS et Atom divers permettant de surveiller l'activité d'un utilisateur, d'un album, d'un groupe, etc., et propose une API ⁽⁵⁾ permettant aux programmeurs de développer des services. On peut enfin choisir des contacts parmi les autres utilisateurs de *Flickr* et ajouter des photos quelconques à ses favoris.

Cet ensemble de possibilités que l'on vient de résumer rapidement permet de constituer un réseau de relations multiples entre les utilisateurs et leurs photos et justifie que *Flickr* soit souvent considéré comme l'un des sites exemplaires du Web social.

3. Les métadonnées utilisées dans le projet PhotosNormandie

Rappel sur les métadonnées IPTC/IIM et la technologie XMP

L'IPTC ⁽⁶⁾ est une organisation internationale créée en 1965 pour développer et promouvoir des standards d'échange de données à destination de la presse. En association avec la NAA ⁽⁷⁾, l'IPTC a défini un modèle global de données appelé *IPTC-NAA Information Interchange Model (IIM)*.

La société Adobe a utilisé ce modèle en 1994 pour définir dans son logiciel Photoshop les informations associées à une image et communément appelées métadonnées IPTC ; il s'agit de champs textuels (Titre, Description, Mots-clés, Ville, Pays, Copyright, etc.) qui sont stockés à l'intérieur de l'image.

seconde guerre mondiale » alors qu'il s'agit d'un portrait du général Théodore Roosevelt pourtant facilement reconnaissable.

5 - Application Programming Interface – bibliothèque de méthodes de base permettant à un programme d'accéder aux fonctionnalités d'un système ; l'API de *Flickr* est décrite sur <http://www.flickr.com/services/api/>
Nous avons développé pour *PhotosNormandie* deux petits programmes utilisant l'API Flickr. Le développement de ce genre d'outil n'est pas difficile pour un informaticien maîtrisant un quelconque langage supportant les requêtes REST (*Representational state transfer*).

6 - International Press and Telecommunications Council, site web : <http://www.iptc.org/>

7 - Newspaper Association of America, site web : <http://www.naa.org/>

XMP (*eXtensible Metadata Platform*) est un format créé par la société Adobe en septembre 2001. XMP repose sur une version simplifiée de RDF (*Resource Description Framework*) qui est un formalisme défini par le W3C ⁽⁸⁾ permettant d'encoder, échanger et réutiliser des métadonnées structurées. XMP utilise le schéma Dublin Core comme fondation et il est extensible, c'est-à-dire que l'utilisateur peut définir ses propres schémas de métadonnées (PECCATTE, 2007).

Le codage des métadonnées dans les photos du projet

Le modèle IIM est à présent considéré par l'IPTC comme un "standard obsolète" qui sera progressivement remplacé par la technologie XMP. Les métadonnées IPTC/IIM restent néanmoins très largement utilisées par les professionnels de l'image numérique.

Nous avons choisi d'encoder les métadonnées descriptives des photos de notre projet à la fois en IPTC/IIM et en XMP de façon à garantir la pérennité du travail réalisé.

Le projet s'appuie sur une fonctionnalité peu connue de *Flickr* : sa capacité à exploiter un sous-ensemble du standard IPTC/IIM ou de son "mapping" en XMP. Plus précisément, nous utilisons :

- le champ IPTC/IIM nommé *Object Name* (n° 5) pour la référence originale de la photo, de la forme *pxxxxxx* ; cette référence s'affiche automatiquement comme Titre en haut de la photo sur *Flickr* ;
- *Caption* (n° 120) pour la description qui s'affiche sous la photo sur *Flickr* ;
- *Copyright* (n° 116) qui contient soit "Conseil Régional de Basse-Normandie / National Archives USA" soit "Conseil Régional de Basse-Normandie / National Archives Canada" ;
- *Keywords* (n° 25) contient en général un seul mot-clé : "Seconde Guerre Mondiale" qui s'affiche comme *tag* à droite de la photo sur *Flickr* ;
- *City* (n° 90) contient le nom de la ville où la photo a été prise ;
Province/State (n° 95) contient le nom du département français où la photo a été prise ;
Country Name (n° 101) est en général vide excepté pour les rares photos qui ont été prises en Angleterre avant le Débarquement. Ces trois champs s'affichent également comme *tags* à droite de la photo sur *Flickr*.

Attention : sur *Flickr*, seules les photos originales (accessibles par le menu *Toutes les tailles* puis *Originale*) possèdent des métadonnées IPTC/IIM et XMP.

Avantages et inconvénients de la méthode de codage des métadonnées dans les photos

Lors du chargement sur *Flickr* d'une photo contenant des métadonnées IPTC, celles-ci sont automatiquement décodées et utilisées pour renseigner le titre, la description et les *tags* selon la correspondance donnée ci-dessus. On retrouve ainsi les avantages bien connus qui ont fait le succès de cette technique d'encapsulation des métadonnées dans les images : la description textuelle de l'image est toujours disponible avec l'image et facilement réutilisable, il n'y a pas de risque de perdre des métadonnées, et l'utilisateur reste ainsi libre de la technologie de base de données utilisée pour l'exploitation de son fonds. Ce dernier point signifie que notre projet est totalement indépendant de *Flickr* que nous pourrions très facilement remplacer, le cas échéant, par une autre plate-forme de partage supportant les métadonnées IPTC. Le travail rédactionnel demeure contrôlé en local puisqu'il est stocké dans les photos. Nous ne sommes pas captifs de la plate-forme de partage.

La mise à jour d'une description est par contre lourde puisqu'il est alors nécessaire de recharger la photo contenant la nouvelle description. En outre, puisque chaque photo est identifiée sur *Flickr* par un numéro automatiquement attribué, l'URL *Flickr* pointant sur une photo de notre projet ne peut être considérée comme stable car le numéro en question change lors d'une mise à jour. La seule identification utilisable reste la référence d'origine sous la forme *pxxxxxx*.

8 - World Wide Web Consortium, site web : <http://www.w3.org/>

4. Le processus documentaire et rédactionnel

Mode de fonctionnement

Tout visiteur peut rechercher, afficher, télécharger les photos en haute définition. Pour commenter les photos et participer ainsi à l'amélioration de leurs descriptions, l'utilisateur doit ouvrir un compte gratuit sur *Flickr* et propose alors ses corrections dans le champ *Ajoutez votre commentaire*. Une discussion peut s'établir entre les divers participants du projet et se termine par la validation éditoriale des modifications proposées.

Pour faciliter le suivi des discussions, nous avons créé un groupe *Discussions sur PhotosNormandie* qui donne une vue d'ensemble sur les photos où une discussion est en cours. Ce groupe est alimenté automatiquement par un programme développé à l'aide de l'API *Flickr* à partir du flux RSS permettant de surveiller les commentaires postés par les utilisateurs ; la consultation d'un groupe *Flickr* est en effet plus simple pour les utilisateurs qui ne maîtrisent pas la technique des flux RSS. Quand la discussion est terminée, l'administrateur rédige une nouvelle description dans les métadonnées IPTC de la photo, sur sa base locale de photos. Puis un autre programme développé à l'aide de l'API efface l'ancienne photo sur *Flickr* en se basant sur la référence et envoie la même photo contenant la nouvelle description IPTC. La photo mise à jour avec sa nouvelle description apparaît alors automatiquement au début de l'ensemble *PhotosNormandie*.

Les photos sont classées dans différents albums : *Manche*, *Calvados*, *Orne*. Celles qui ne se rapportent pas directement à la Bataille de Normandie sont regroupées dans d'autres albums.

Bilan documentaire

[Chiffres au début septembre 2008 - le projet est toujours actif]

Depuis le début du projet sur *Flickr* en janvier 2007, les 2763 photos du corpus ont été vues globalement plus de 685 000 fois, ce qui correspond à plus de 1300 visites par jour.

Le groupe *Discussions sur PhotosNormandie* compte 39 membres dont une dizaine participe régulièrement au projet en postant des commentaires.

Nous avons complété, corrigé et mis à jour 3806 descriptions. Ce nombre plus grand que celui des photos s'explique parce que certaines légendes ont été corrigées plusieurs fois.

La typologie des améliorations peut être résumée ainsi :

- identifications de localisations, ex. p011719 et p013151 (recherches sur des photos aériennes). Nous avons aussi identifié plusieurs photos qui n'ont pas été prises en Normandie comme la p013091 (Saint-Quentin dans l'Aisne) ou la p013035 (Rome) ; elles sont regroupées dans l'album *Autres Localisations*. Nous avons également créé un groupe intitulé *D'hier à aujourd'hui* qui met en relation des photos de la collection *PhotosNormandie* avec des photos récentes prises aux mêmes endroits ;
- identifications de personnages, ex. p013416 (le général Davis, premier général noir de l'US Army), p013400 (le secrétaire au trésor Morgenthau), p013391 (l'acteur Edward G. Robinson). Il ne s'agit pas toujours de personnalités et nous avons aussi tenu à citer les noms de soldats "de base" et de civils lorsque nous avons pu les connaître ;
- identifications d'unités militaires, ex. p013094 ;
- précisions de dates, ex. p011628 ;
- précisions descriptives sur l'image, ex. p013233 ;
- références : renvois à des livres, à des sites ;
- identifications des photos censurées, des photos en couleur, des photos en doubles et des séries. Il suffit ainsi de rechercher sur le mot "censure" pour retrouver la trentaine de photos que nous avons repérées comme "caviardées", c'est-à-dire présentant une ou plusieurs surfaces volontairement effacées ;
- contextualisations historiques : précisions sur un mouvement d'unité, une action, etc. en

- rapport avec l'image, ex. p013341 ;
- contextualisations iconographiques à l'aide d'autres sources :
 - ✓ séquences vidéos de la même scène : ex. p013292 (renvoi sur *youtube*), p000034 (renvoi sur *dailymotion*) ;
 - ✓ autres photos : ex. renvois à des photos prises par Robert Capa (rechercher sur son nom) ; en particulier la p013283 où nous avons "découvert" le célèbre photographe en train de changer la pellicule de son appareil.

Les descriptions proposées sont pratiquement toutes factuelles et objectivées. Pour quelques unes d'entre elles néanmoins nous avons obtenu et rapporté des informations plus subjectives (9).

Organisation du travail

L'un des intérêts du projet est bien sûr de faire appel à des spécialistes aux compétences complémentaires. C'est ainsi qu'un participant régulier est un expert sur les diverses unités combattantes de l'époque tandis qu'un autre possède une excellente connaissance de l'aviation durant la Seconde Guerre mondiale. Nous avons aussi obtenu de nombreuses informations précieuses de la part d'un spécialiste des vues aériennes, plusieurs intervenants nous ont aidé sur le Mur de l'Atlantique, et un collectionneur de cartes postales anciennes a identifié de nombreuses localisations. Il est à noter que le projet constitue un "espace virtuel" regroupant des intervenants qui ne se connaissent pas "dans le monde réel".

À côté des recherches et références classiques dans des ouvrages et magazines, nous avons aussi abondamment utilisé les recherches sur le Web et sollicité les habitués de certains forums spécialisés tel que <http://forum.6juin1944.com>. On peut donc distinguer à côté de la production directe une seconde strate dans l'organisation du travail. Elle est constituée par les contributions indirectes apportées par les utilisateurs des forums qui forment une sorte de réseau de second niveau qu'il est indispensable d'activer dans ce genre de projet. En effet, les passionnés ont souvent déjà mis en place des moyens de partage de leurs informations et peuvent être réticents à en utiliser un nouveau (nécessité de créer un compte, apprentissage, etc.).

Nous avons pu aussi tester un nouveau service de recherche d'images par similarité avec le logiciel TinEye de la société Idée (<http://tineye.com/>). L'application a identifié plusieurs utilisations d'images de la collection pour des couvertures de livres ou des jaquettes de DVD. Quelques photos retrouvées sur d'autres sites nous ont aussi permis d'améliorer certaines légendes, par exemple pour la p013390, où nous avons identifié une femme officier dont nous recherchions le nom depuis longtemps.

Il subsiste néanmoins de nombreuses photos que nous n'avons pas réussi à localiser (voir l'album *Localisations inconnues*), des personnalités non identifiées (album *Personnages inconnus*), ainsi que des dates, unités, événements ou situations qui demeurent encore énigmatiques malgré nos recherches.

5. Usages de Flickr et redocumentarisation

Une pratique de redocumentarisation

Plusieurs études récentes ont montré que les usages de *Flickr* sont variés, et aucun de ces usages divers n'est majoritaire. De plus, les relations multiples permises par le système concernent en réalité peu de monde. Une très petite minorité d'utilisateurs produit la majorité du contenu public et organise ce contenu par leur activité sur les groupes et les photos (d'après PRIEUR et al., 2008). On peut d'autre part distinguer deux types de pratiques sociales selon que l'utilisateur est plus intéressé par les contacts ou par la socialisation des contenus (ibid.). L'usage de *Flickr* dans notre projet est donc très

9 - Ainsi pour la p000713 où l'on voit un couple de vieux Normands se recueillir sur la dépouille d'un parachutiste américain. Sur cette photo, la censure a voulu renforcer l'aspect émotionnel en effaçant un caméraman ; mais nous avons appris par un arrière petit-fils de ce couple que leur recueillement était bien spontané et que la photo n'est pas une mise en scène comme le seul examen de l'image pourrait le laisser penser.

minoritaire et certainement atypique. Il se caractérise certes par un faible nombre de contacts actifs, mais surtout par une activité de commentaires rapidement caducs et orientés vers un but unique, ainsi que par un grand nombre de mises à jour. *PhotosNormandie* correspond en fait à une démarche documentaire et rédactionnelle, à un travail collectif avec un objectif de production, et non à une folksonomie caractéristique du web social (ERTZSCHEID, O., 2008).

Le corpus présenté ne correspond d'ailleurs pas aux images habituelles sur *Flickr* qui apparaissent faiblement contextualisables pour la plupart des visiteurs, et sont habituellement décrites à l'aide de *tags* non contrôlés. De fait, nous n'utilisons pas le "tagging" si ce n'est de manière indirecte puisque les informations de localisation IPTC apparaissent comme des *tags* après chargement. On doit plutôt considérer que ce fonds est véritablement éditorial, fortement contextualisé pour une majorité de visiteurs, et présente des images d'actualité - même si cette actualité a plus de soixante ans - qui nécessitent des descriptions riches et soigneusement rédigées.

Par ailleurs, les commentaires écrits par les divers intervenants ne sont pas conservés puisque l'ensemble du processus conduit à effacer une photo pour la mettre à jour avec une nouvelle description, et donc à perdre la discussion ayant conduit à cette nouvelle description. Deux des principaux vecteurs de la popularité sur *Flickr*, les *tags* et les commentaires (COX, 2007 ; PRIEUR et al., 2008), ne sont donc absolument pas utilisés selon le mode cumulatif habituel sur les sites sociaux. Dans le même esprit, nous refusons de rejoindre des groupes qui ne sont pas en rapport avec le sujet ou la pratique ouvertement documentaire et rédactionnelle du projet, nous ne conservons pas de favoris, ne "taggions" pas les images de nos contacts, etc. En bref, la course à l'*interestingness* ⁽¹⁰⁾ comme on dit dans le jargon flickerien n'est pas notre but. *Flickr* apparaît donc dans ce projet comme un outil, et seulement un outil, intégré à d'autres usages du web (forums et autres sources d'information, recherche de documents, etc.) dans un processus de redocumentarisation, c'est-à-dire dans une entreprise collective qui vise à traiter à nouveau une collection de documents (cf. ZACKLAD, 2007 ; SALAUN, 2007); il s'agit bien de réindexer un corpus iconographique avec les possibilités ouvertes par les technologies numériques et Internet pris dans son ensemble (pas seulement le Web social).

L'activité régulière du projet durant plus d'une année a permis d'améliorer de façon substantielle la qualité des descriptions, et même, pensons-nous, de les enrichir d'informations probablement inédites. Ce fonctionnement qui s'est avéré efficace repose sur la simplicité et la souplesse de l'outil collaboratif qu'est *Flickr*, sur la qualité des intervenants, mais aussi sur un travail important de la part des co-responsables, et il est important de bien prendre en compte cette charge de travail quand on envisage un tel projet d'indexation sociale. Notre organisation a conduit à définir deux fonctions indispensables :

- un rédacteur en chef capable de synthétiser les informations recueillies dans une discussion ou sur d'autres sources et de rédiger une description précise et cohérente. Dans notre mode de fonctionnement, ce n'est pas un animateur ou un coordinateur proposant de nouvelles descriptions au groupe. Il s'agit d'une fonction analogue à celle rencontrée dans la presse puisqu'il décide seul de son texte final à partir des éléments rédactionnels recueillis ;
- un administrateur technique qui malgré les automatismes mis en place et décrits ci-dessus doit effectuer un travail régulier.

Les possibilités ouvertes par la numérisation généralisée des documents et l'usage collaboratif du Web engendrent un travail important. Si les deux fonctions évoquées – qui peuvent être assurées par une même personne ou non - ne sont pas définies avec précision dès le démarrage et assurées avec constance, toute tentative de redocumentarisation de ce genre nous semble impossible.

10 - Il s'agit de l'intérêt accordé par *Flickr* à une photo selon des critères variés liés à l'activité sociale autour de la photo, cf. <http://www.flickr.com/explore/interesting/>

Autres projets similaires sur Flickr

La *Library of Congress* a mis en ligne sur *Flickr* plus de 3000 photos en janvier 2008 ⁽¹¹⁾, soit un an après le démarrage de *PhotosNormandie*. Il s'agit de deux collections historiques, un ensemble de photos sur l'actualité des années 1910 d'une part, et des photos en couleur prises dans les années 1930-1950 d'autre part. La *Library of Congress* encourage les utilisateurs à ajouter des *tags*, des commentaires, des notes aux photos dans le but d'obtenir des informations ou des identifications inédites. La Bibliothèque a conclu un partenariat avec *Flickr* appelé *The Commons* ⁽¹²⁾ auquel participe aussi le *Powerhouse Museum* de Sydney, le *Brooklyn Museum* de New York, la *Smithsonian Institution* de Washington, la *Bibliothèque de Toulouse*, la *George Eastman House*, la *Biblioteca de Arte-Fundação Calouste Gulbenkian* de Lisbonne et le *National Media Museum* de Bradford (UK) [liste établie début septembre 2008].

Les photos publiques de la *Library of Congress* ne contiennent pas de métadonnées IPTC ou XMP ; un lien entre les photos sur *Flickr* et les notices au format Marc sur le site de la *Library of Congress* est réalisé grâce à un *machine tag* (tag de programmation), positionné lors du chargement à l'aide de l'API et visible en dessous des *tags* habituels à droite de la photo.

L'équipe en charge du projet compte 12 personnes. Ils reconnaissent que le travail d'exploitation et d'évaluation des commentaires laissés par les utilisateurs est lourd et prend beaucoup de temps. La *Library of Congress* a mis à jour une centaine de descriptions de photos après trois mois d'activité (BERMES, 2008).

6. Vers une plate-forme professionnelle pour la redocumentarisation des fonds iconographiques

L'expérience acquise au cours de ce projet est très positive et certainement applicable à de nombreux domaines qui nécessitent la participation de spécialistes disséminés dans le monde entier. Nous l'avons un moment envisagé avec un groupe d'experts en numismatique sur une collection de photos de pièces de monnaies anciennes. Ce dernier projet n'a pas abouti pour le moment parce qu'il n'a pas été possible de trouver une personne suffisamment disponible et motivée pour assurer la tâche de rédacteur en chef telle que nous l'avons définie.

Des initiatives comme *PhotosNormandie* et *The Commons* ouvrent de nouvelles perspectives dans la valorisation des fonds patrimoniaux – et probablement au delà de ce secteur. Elles s'inscrivent résolument dans un vaste mouvement d'émergence de nouveaux médias issus de la redocumentarisation du monde (PEDAUQUE, 2007) où les métadonnées descriptives ne sont plus figées mais deviennent de véritables contenus publics perfectibles. Dans un tel processus, les métadonnées ne sont définitivement plus considérées comme des fiches documentaires statiques. Il s'agit de contenus à part entière qui évoluent selon la dynamique collaborative mise en place.

L'expérience devrait néanmoins être approfondie dans différentes directions telles que la simplification des procédures actuelles, le traitement des descriptions multilingues, le support de schémas de métadonnées spécifiques ou orientés métiers, la géolocalisation des images, la sécurisation, l'utilisation de vocabulaires contrôlés pour les *tags*, le développement d'une ontologie sur le domaine couvert, le traitement de corpus plus importants et moins homogènes, etc.

Nous pensons que le développement de la technologie XMP permettrait d'améliorer considérablement l'ensemble du processus décrit en ce qui concerne les trois premiers points qui nous paraissent les plus importants : simplification, multilinguisme et schémas de métadonnées orientés métiers. XMP permet en effet de définir des schémas personnalisés de métadonnées, indispensables pour de nombreux secteurs, et plusieurs logiciels supportent d'ores et déjà des interfaces paramétrables pour travailler de manière efficace avec ces schémas sur mesure. XMP permet aussi de travailler avec certains champs en plusieurs langues, et là encore, même si le

11 - http://www.flickr.com/photos/library_of_congress/

12 - <http://www.flickr.com/commons/>

mouvement est moins rapide, les logiciels qui supportent les légendes multilingues selon ce format commencent à apparaître. Enfin, les métadonnées XMP peuvent non seulement être encapsulées dans une image au format habituel (TIFF, JPEG) mais elles peuvent aussi être véhiculées indépendamment pour faciliter leur ajout ou mise à jour.

Les technologies permettant d'indexer professionnellement en XMP existent pour le poste individuel ou la production de masse (agences), mais pas à notre connaissance dans un contexte collaboratif. Il nous semble ainsi possible (et souhaitable) d'aller plus loin que les possibilités d'indexation sociale permises par *Flickr* en définissant une véritable plate-forme collaborative facilitant le travail d'experts associés à la redocumentarisation de fonds iconographiques divers par la maîtrise intégrale de leur processus documentaire et rédactionnel. L'expérience de *PhotosNormandie* est suffisamment significative pour contribuer au cahier des charges d'un tel outil.

Bibliographie

BERMES, E. (2008), *LC+Flickr : Bilan d'une expérience 2.0*, sur le blog *Figoblog*, <document en ligne sur <<http://www.figoblog.org/node/1921>>>

COX, A. M., (2007), *Flickr : What is new in Web2.0?*. In Proc. Of Towards a social science of web2.0, Workshop "Towards as social science of Web2.0", University of York (document en ligne sur <<http://www.shef.ac.uk/content/1/c6/04/77/66/flickr%20paper.pdf>>)

DALBIN, S., (2007) , *Mariage entre ancien et nouveau, Web2 et mémoire*, sur le blog *Descripteurs* (document en ligne sur <http://dossierdoc.typepad.com/descripteurs/2007/02/mariage_entre_a.html>)

DELAGE C., GUIGUENO V., GUNTHERT A., *Le débarquement de 1944 par Capa (1944), Darryl Zanuck (1962) et Steven Spielberg (1998)*, in *La fabrique des images contemporaines*, Paris : Editions Cercles d'Art, 2007

ERTZSCHEID, O., (2008), *Indexation sociale et folksonomies : le monde comme catalogue*, Journées ABES, 20 et 21 mai 2008, sur le blog *Affordance*, (document en lignes sur <http://affordance.typepad.com/mon_weblog/2008/05/journes-abes-fo.html>)

MARLOW, C., NAAMAN, M., BOYD, D., DAVIS, M. (2006), *HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead*. Proceedings of the Seventeenth ACM Conference on Hypertext and Hypermedia, ACM Press, Odense, Denmark, August 2006 (document en ligne sur <<http://www.danah.org/papers/Hypertext2006.pdf>>)

PECCATTE, P., (2007), *Métadonnées : une initiation. Dublin Core, IPTC, Exif, RDF, XMP, etc.* (document en ligne sur <<http://peccatte.karefil.com/software/Metadata.htm>>)

PÉDAUQUE, R. T. (collectif), (2007), *La Redocumentarisation du Monde*, Paris : Éditions Cépadues

PRIEUR, C., CARDON, D., BEUSCART, J-S., PISSARD, N., PONS, P., (2008), *The Strenght of Weak cooperation : A case study on Flickr* (document en ligne sur <<http://www.scribd.com/doc/3008294/Case-study-for-Flickr>>)

SABOU, M., GRACIA, J., ANGELETOU, S., D'AQUIN, M., MOTTA, E., (2007), *Evaluating the Semantic Web : A Task-based Approach* (document en ligne sur <<http://kmi.open.ac.uk/people/marta/papers/iswc2007.pdf>>)

SALAUN, J.-M., (2007), *Éclairages sur la redocumentarisation*, Bloc-notes de Jean-Michel Salaün, (document en ligne sur <<http://blogues.ebsi.umontreal.ca/jms/index.php/2007/05/05/252-eclairages-sur-la-redocumentarisation>>)

SPECIA, L., MOTTA, M., (2007), *Integrating Folksonomies with the Semantic Web* (document en ligne sur <<http://www.eswc2007.org/pdf/eswc07-specia.pdf>>)

ZACKLAD, M., (2007), *Réseaux et communautés d'imaginaire documédiatisées*, in Skare, R., Lund,

W. L., Varheim, A., A Document (Re)turn, Peter Lang, Frankfurt am Main : 279-297
(document en ligne sur
<http://archivesic.ccsd.cnrs.fr/docs/00/18/01/85/PDF/imaginaire___document_Skare_Lund_Varheim.pdf>)