

ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

Mémoire de Master II : Histoire visuelle

Indexation visuelle et recherche d'images
sur le Web :
Enjeux et problèmes

Amélie Segonds

Directeur de recherche

André Gunthert

2009

Sommaire

Introduction

Chapitre 1 L'indexation sur les moteurs de recherche du Web.....p. 9

- I. Internet : un nouvel espace documentaire
- II. L'indexation des documents visuels par les moteurs de recherche

Chapitre 2 L'indexation sémantique et la Folksonomie.....p. 34

- I. L'apparition d'une indexation sociale
- II. Vers les nouveaux usages de la folksonomie
- III. Les limites de la folksonomie

Chapitre 3 Le géotagging comme nouveau système d'indexation.....p. 74

- I. Les présupposés d'une indexation par coordonnées géographiques
- II. Les mises en application sur le Web
- III. Vers un déplacement de la sphère documentaire

Chapitre 4 L'indexation et la recherche par le contenup. 108

- I. La recherche de l'image par l'image
- II. Une observation des sites en ligne
- III. Les enjeux de la recherche par le contenu

Conclusion générale.....p.147

Bibliographie.....p.151

Introduction

L'une des évolutions majeures du Web aujourd'hui concerne le changement de nature du document à l'échelle de son déploiement sur les réseaux. Comme le souligne le collectif de chercheurs RTP-DOC « Le Web savant initial était un Web de documents rédigés. Il ne représente qu'une faible minorité du trafic d'Internet désormais. Ce sont des documents multimédia qui s'échangent massivement (musique, image, vidéo). Plus largement, c'est la place de l'image qui change le niveau de participation aux réseaux et à la production documentaire.¹ ». Une renégociation documentaire est à l'œuvre sur le Web, signant la fin de la suprématie du texte et ratifiant l'importance croissante portée au partage et à la circulation des documents visuels. De cette transformation de nature des documents du Web découlent nécessairement de nouvelles interrogations d'une part quant au traitement des masses documentaires inédites dans leur forme. Elle repose d'autre part la question de la recherche et de l'accès pertinent à ces mêmes documents. Et l'on suivra volontiers RTP-DOC quand ils considèrent que cette « transformation des propriétés sémiotiques des documents les plus échangés constitue une révolution dont on n'a sans doute pas encore perçu tous les effets [...] L'analyse du document, de ses enjeux en termes d'indexation ne peut ignorer la concomitance de ce développement des machines à produire des images numériques peu coûteuses et de la participation étendue à des milieux sociaux beaucoup plus divers que les premières époques d'internet.² ».

La généralisation des supports numériques, l'apparition de formats vidéo et photo compacts, l'émergence du Web 2.0, de ses espaces de partage de données multimédias et sites collaboratifs ont engendré une augmentation vertigineuse de la quantité de documents visuels sur le Web. Les chiffres récents du nombre d'images circulant sur Internet recensent des milliards d'items au moment même où plus d'un million de photographies numériques sont mises en ligne par jour. La question n'est plus aujourd'hui de produire de nouvelles images mais de trouver l'image déjà existante dans le stock disponible. Le problème crucial à prendre en compte face à ce flux ininterrompu de production d'images est celui de leur indexation. Les techniques dites d'indexation, dans le champ du visuel comme du textuel, se proposent d'attacher à une image ou à une vidéo ou tout type de

¹ Pédauque, Roger T. *Documents et modernité* [en ligne]. Paris : CNRS, 2006 (consulté le 25/03/09). Disponible sur : <http://rtp-doc.enssib.fr/IMG/pdf/Pedaauque3-V4.pdf>

² Ibid.

document, un ensemble de descripteurs de leur contenu, dans le but de mesurer la ressemblance avec les descripteurs correspondant à la requête. Or comme le souligne Bruno Bachimont « si le texte « *est sa propre indexation* », la mémoire audiovisuelle reste muette sans traitement documentaire³ ». Il s'agit donc pour nous de comprendre ces traitements documentaires nécessaires sans lesquels aucune indexation de ressources visuelles n'est possible.

L'indexation est un besoin ancien mais une pratique récente thématifiée par des disciplines jeunes, la documentation et la bibliothéconomie. Élaborée ces deux derniers siècles (au XX^e siècle en ce qui concerne la documentation), l'indexation a concerné de manière privilégiée les contenus textuels à partir desquels ont été mis au point ses procédés et méthodes. En tant que notion de classification, l'indexation est un terme récurrent en science de l'information et appartient au domaine de la gestion documentaire, l'indexation relève avant tout d'un travail sur le contenu, où l'interprétation et le discernement sont la clé pour représenter un document de manière à pouvoir le retrouver. Il nous semble utile, à ce stade, de préciser la différence entre les notions de catalogage, de classification et d'indexation. Le catalogage consiste à décrire physiquement un document, quel que soit son format, permettant d'une part de l'identifier de façon unique et d'autre part de le repérer par le biais d'une caractéristique qui n'a pas rapport à son contenu (numéro ISBN, nom de l'auteur, etc.).

La classification permet de rattacher un document, après en avoir analysé le contenu de façon générale, à une catégorie, à un ensemble de documents qui traitent du même sujet. Le document est ici considéré comme une entité.

Quant à l'indexation, on ne considère plus le document comme une entité distincte mais on considère plutôt les éléments d'information qui s'y trouvent. Si le but de l'indexation est toujours de créer des regroupements de documents sur un même sujet, la description se fait plus précise. L'indexation nous permettra donc d'accéder par exemple à tous les documents qui fournissent de l'information sur un même sujet, peu importe que le document qui contient cette information ait été classé dans telle ou telle catégorie.

En effet, l'indexation est cette technique consistant à caractériser le contenu d'un document et l'information qu'il détient de manière à le retrouver quand on effectue des recherches sur l'un des sujets dont il traite. La difficulté est donc de savoir caractériser et

³ Bachimont, Bruno. Bibliothèques numériques audiovisuelles : des enjeux scientifiques et techniques. *Document numérique*, 1998, vol. 2, n° 3-4, p. 219-242

représenter l'information documentaire pour qu'il soit aisé de la mettre en rapport avec des sujets d'investigation. Mise en rapport d'une requête et d'un contenu représenté et synthétisé, l'indexation permet de s'orienter dans la masse des documents et d'organiser ses connaissances. Si l'indexation permet une recherche efficace d'un document, demeure la clé du fonctionnement des moteurs de recherche, cette étape constitue un problème majeur sur le web, en raison de la croissance constante du nombre de documents numériques bruts, non renseignés, qui y circulent.

Dès l'apparition de la technologie numérique de nombreux projets ont vu le jour en recherche d'information visant à plus d'efficacité et de simplicité dans le classement, l'indexation et l'accès aux données. Ces projets aboutissent à la création de logiciels de recherche plein texte ou de reconnaissance des signes performants, aujourd'hui employés sur Internet. Selon l'enquête de Kobayashi and Takeda⁴, 80% des usagers du Web utilisent des moteurs de recherche, de type Google ou Yahoo!, pour trouver une information. Créés à une époque où le Web est principalement un outil de recherche de document textuel fonctionnant selon le principe de double articulation du langage, ces moteurs ont hérité des différentes avancées dans le domaine de l'automatisation du traitement documentaire.

Au regard du succès des moteurs de recherche développés pour le texte, les moteurs de recherche d'image sont comparativement encore rares et peu puissants du fait même de la difficulté à indexer ces documents. Une photographie ne peut en effet se décomposer en unités facilement repérables comme des chaînes de caractères séparées par des blancs permettant la recherche plein texte. Sur le Web, la recherche de documents visuels demeure encore très largement dominée par l'analyse du texte entourant une image et la requête par mots-clés. Si ces dernières connaissent des limites au niveau de la pertinence des résultats, elles sont aussi critiquées pour ne prendre en compte que les éléments textuels et non les caractéristiques propres aux images. Dès les années 1990, de nombreux laboratoires de recherche en France comme à l'étranger ont conduit des projets qui visent à une automatisation de l'indexation des documents iconographiques à partir de descripteurs visuels.

Deux approches coexistent donc aujourd'hui quant à l'indexation des documents visuels, une méthode sémantique qui, pour l'instant et comme le rappelle Antoine Manzarena est manuelle. Une étape selon laquelle, « l'opérateur d'indexation attache au document des

⁴ Kobayashi, M. and Takeda, K, *Information Retrieval on the Web*, ACM Computing Surveys, 2000, Vol.32, No.2, 144-173.

données de haut niveau relatives à la signification du contenu de l'objet. Les requêtes associées sont en général des mots, désignant un objet, une action, le nom d'un personnage ou d'un événement.⁵». Par opposition, des projets tendent à une indexation automatique comprise comme « essentiellement descriptive ou visuelle. L'algorithme d'indexation attache des données de bas niveau sémantique, relatifs aux contenus géométrique, spectral, de l'image, à un niveau local ou global. Les requêtes associées se font en général par l'exemple, ou par modèle. ⁶ ». À ces deux méthodes largement commentées par divers travaux de recherche nous rajoutons celle du géotagging autre système d'indexation, de classement et de navigation qui connaît sur le Web un succès croissant.

Les premières motivations des divers laboratoires de recherche pouvaient en creux faire apparaître le désir d'émanciper l'image de sa dépendance au mot avec pour toile de fond l'idée d'une navigation uniquement visuelle ou pour le dire autrement comment rechercher l'image par l'image. Il fallait y voir une défiance quant à la subjectivité présidant au choix des termes indexeurs à une époque où tous les espoirs se tournaient vers l'intelligence artificielle et l'informatisation du savoir. Cependant, bon nombre de ces projets de recherche d'images sont demeurés dans l'intimité des laboratoires de recherche. On constate depuis quelques années une recrudescence de ces questions d'automatisation de l'indexation dont les résultats ont pour l'instant abouti à quelques fonctionnalités encore souvent à l'état de test tel que la reconnaissance faciale, ou à la création de sites en version beta pour la plupart.

Si l'accès à la connaissance et aux savoirs est évidemment un enjeu motivant les diverses recherches dans le champ de l'indexation visuelle, l'enjeu économique est également un facteur décisif dans cette recherche d'automatisation d'indexation et de précision dans les descriptions de contenus visuels. Les moteurs de référencement et de recherche de documents tels Google trouvent leur viabilité et rentabilité par l'apport de revenus en particulier générés par la publicité contextuelle, des annonces ciblées en fonction d'une sélection de mots-clés. Si les systèmes de régie publicitaire sont des plus performants en ce qui concerne les documents textuels, ces applications connaissent de vraies lacunes quant à la liaison des publicités avec des documents visuels. D'une part parce que peu d'images sont renseignées par mots-clés et d'autres part parce qu'il est pour l'instant impossible d'analyser automatiquement le contenu d'une image fixe ou animée. Il

⁵ Antoine Manzarena, *l'indexation d'images*, [en ligne], disponible sur : <http://www.abhatoo.net.ma/index.php/fr/content/download/2563/28254/file/Indexation%20d'images.pdf>

⁶ Ibid.

va s'en dire qu'au vu du nombre de nouvelles images circulant sur le web chaque jour cette opération ne peut se faire manuellement. Les technologies d'indexation et de descriptions des images par leur contenu apparaissent donc pour Google et les autres moteurs comme une voie salubre dans la rentabilité de leurs services à une époque où le site Youtube, par exemple, et pour les raisons précédemment invoquées, s'avère déficitaire.

Le Web a depuis sa création connu de nombreuses évolutions et est devenu aujourd'hui un espace d'échange de documents polymorphes. Si le texte reste très présent sur le réseau, de nouvelles formes de documents ont fait leur apparition, des documents multimédias, photos, vidéos et sons, qui en plus de la variété du médium ont en commun la diversité de leur origine de production. Une dichotomie s'installe peu à peu entre le mode de fonctionnement originel du web, comme lieu de circulation de documents textuels et la place de plus en plus importante qu'occupent des objets visuels souvent peu légendés, documentés, sources brutes dans leur médium. Comment ces objets pour beaucoup issus de pratiques non officielles, documents privés (dans leur mode de création) mais publiés (rendus publics), peuvent ils être cherchés, retrouvés sur la toile dont le système d'indexation, d'organisation des documents, est voué au texte ? De cette documentarisation nouvelle de sources issues d'usages domestiques et privés, de nouvelles problématiques émergent, tenant à l'indexation et donc à l'utilisation même de ces sources. L'objet de ce travail est de rendre compte de ce que devient l'indexation quand on considère des contenus numériques non textuels comme les images fixes. Si l'indexation des documents non textuels sur des supports analogiques comme les photographies, les films, les dessins et autres schémas est pratiquée de longue date, l'introduction des techniques numériques a profondément modifié l'indexation qui est ainsi devenue « recherche d'information » et a permis d'élaborer un nouveau paradigme, « l'indexation fine des contenus » ou « l'indexation par le contenu ». Si la pratique de l'indexation est un système de classification ancien, il connaît aujourd'hui un renouveau en rapport avec l'introduction du numérique au sein de la sphère documentaire. La fonction documentaire de l'image au sein du web doit alors être comprise selon l'horizon de sa nature numérique en interrogeant l'apport de cette technologie aussi bien à la notion de document qu'aux méthodes de gestion documentaires.

La question transversale à tout ce travail est donc comment indexer des documents visuels, qu'est ce qui est indexable au sein d'une image ?

Au-delà d'une approche purement descriptive il s'agira de montrer comment la problématique de la structuration de la sphère documentaire est devenue l'enjeu de

pratiques sociales nouvelles, entraînant elles mêmes la création de nouveaux outils de gestion et de structuration des collections d'images sur Internet.

Pour ce faire nous avons pris comme objets d'étude le Web, ses moteurs de recherche et sites de partage d'images. Pour chaque chapitre de ce travail une sélection de sites et plateformes sera effectuée afin de conserver les plus spécifiques, mettant en jeu des problématiques cruciales en termes de technique d'indexation ou d'usages.

Nous divisons ce mémoire en quatre parties. La première est consacrée à une analyse du fonctionnement des moteurs de recherche du Web, de leur gestion des ressources visuelles, nous avons privilégié Google comme étude de cas.

À côté des moteurs possédant des fonctions de recherche d'images classiques comme Google Image, se sont développés, en parallèle de l'émergence du web dit 2.0, des plateformes de stockage et de partages d'images type de Flickr ou encore Panoramio dont les bases peuvent compter des milliards de photographies. Nous considérons donc la façon dont ces sites ont géré l'indexation des collections, les systèmes qu'ils ont adoptés pour faciliter le classement et la recherche de photographies dans des bases de données toujours croissantes et rassemblant des documents souvent non renseignés, à destination d'un public d'utilisateurs pour la plupart non spécialistes de la recherche documentaire.

En ce sens notre deuxième chapitre se concentre sur les principes de l'indexation textuelle appliquée aux images et plus particulièrement sur la folksonomie dans ses avantages, les types d'appropriations qu'elle engendre mais aussi ses limites. Flickr constitue ici le corpus principal.

Le troisième moment est quant à lui dévolu à une définition et une observation du fonctionnement du géotagging comme nouvelle forme d'indexation, de gestion documentaire et de navigation au sein des collections.

Enfin dans le quatrième et dernier chapitre nous revenons sur les présupposés techniques des systèmes d'indexation et de recherche d'images par le contenu. Par l'analyse des différents sites proposant cette fonctionnalité nous nous efforcerons de cerner les enjeux documentaires, économiques et sociaux du procédé. Nous envisageons sa compatibilité avec les attentes du grand public pour que demain l'image, vecteur désormais majeur du contenu du Web trouve sa réelle émancipation vis-à-vis du texte sur lequel repose encore aujourd'hui pour l'essentiel l'indexation des documents visuels.

Chapitre 1

L'indexation par les moteurs de recherche du Web

Sommaire du chapitre

I. Internet : un nouvel espace documentaire.....	11
À. Moteur de recherche et évolution des contenus du Web	11
1. Du Web aux moteurs de recherche	11
2. Extension des espaces documentaires indexables	14
B. Indexation des contenus et publicités contextualisées.....	16
1. Double articulation et indexation textuelle	16
2. Une indexation à visée économique	18
II. L'indexation des documents visuels par les moteurs de recherche	21
À. Accroissement des contenus visuels par le Web 2.0	21
1. Nouvelle approche du document	21
2. L'explosion visuelle du Web	23
B. Les systèmes d'indexation de la firme Google.....	26
1. Chercher le texte avant l'image	26
2. Google Image et Google Image Labeler.....	28
3. Les solutions pour la vidéo	31

I. Internet : un nouvel espace documentaire

A. Moteur de recherche et évolution des contenus du Web

Au début des années quatre-vingt-dix apparaît l'internet tel que nous le connaissons aujourd'hui, un système informatique en réseau où le Web devient l'application phare. Depuis sa création, le World Wide Web n'a cessé de connaître des évolutions dans ses programmes, fonctionnements et contenus. Récemment encore, la Toile a subi une transformation touchant ses appellations mêmes. Il est ainsi question de l'hypernet, du Web social, autant de termes qui témoignent d'une transformation progressive du Web, touchant aussi bien la technique que les contenus, les deux étant irrémédiablement liés.

1. Du Web aux moteurs de recherche

Dans le vocabulaire courant, il est commun que l'Internet et le World Wide Web soient confondus. Internet est un réseau informatique mondial qui rend accessible certains services de communication et d'informations tels que les courriers électroniques, le Web, les messageries instantanées, etc. Le Web quant à lui naît officiellement en 1989 avec la publication d'un article de Tim Berners Lee « L'hypertexte et le CERN ⁷ ». Il s'agit d'un système d'hypertexte public désignant un service, une application spécifique de l'Internet. Il fonctionne selon plusieurs protocoles d'échanges et de localisation de l'information dont L'http, protocole de communication qui permet le transfert de données entre le réseau et un ordinateur particulier. Mais pour que le Web existe comme espace documentaire, il faut la conjonction de trois éléments supplémentaires : des adresses URL grâce auxquelles les informations peuvent être localisées, des navigateurs permettant d'y accéder et un format d'encodage l'HTML, langage informatique par lequel les données récupérées peuvent être affichées et qui permet l'insertion d'hyperliens au sein des pages du Web. Ces divers protocoles sont les éléments constitutifs du Web. Ce dernier apparaît alors comme une vaste base de données navigable par le biais des systèmes de transmission d'informations, des navigateurs, et des hyperliens.

Si le Web peut donc être envisagé comme un espace documentaire, sa réalité sociale et informationnelle nous incite plus à le considérer comme un « continent »

⁷Cf. le texte original de Tim Berners Lee <http://www.w3.org/History/1989/proposal.html>

documentaire. Les dernières enquêtes de la société Netcraft⁸, recensent en avril 2009 plus d'un milliard d'internautes. Au niveau des sites Web, les derniers chiffres font mention de l'existence de plus de 900 millions de noms de domaines pour environ 230 millions de sites actifs, c'est à dire ayant du contenu pouvant être indexé par les moteurs. Nous suivons dans leur raisonnement Olivier Ertzscheid, Gabriel Gallezot et Éric Boutin quand ils affirment « Au vu de ces chiffres qui confinent à l'incommensurable, on comprend mieux la nécessité de l'existence d'outils de repérage, de classification et d'accès.⁹ ». Selon eux toujours, si les premiers systèmes de navigation au sein des pages de Web reposaient sur de simples signets ou listes d'adresses, il fallut mettre au point, à mesure que le nombre de sites et de pages croissaient de façon exponentielle, de nouveaux modes de recherche et de consultation des données. C'est dans cette perspective que les moteurs de recherche se sont créés, afin d'améliorer la recherche d'informations et rendre la masse de documents toujours plus accessible, au travers d'outils simples d'utilisation et globaux. En ce sens le Web s'inscrit plus que jamais dans une problématique documentaire et la pousse à son paroxysme en cristallisant ses interrogations aussi bien en termes de conservation qu'en termes d'accès aux documents.

Les principaux moteurs de recherche existant aujourd'hui sont Google, Yahoo, ou encore Live search (le moteur de Microsoft) ; ils fonctionnent selon trois étapes fondamentales :

- L'exploration ou *crawl* : Un moteur de recherche est d'abord un outil d'indexation, c'est-à-dire qu'il dispose d'une technologie de collecte de documents à distance sur les sites Web, via un outil que l'on appelle robot. Le Web est alors exploré par ce robot d'indexation suivant tous les hyperliens qu'il rencontre et récupérant les ressources jugées intéressantes.

- L'indexation des ressources récupérées consiste à extraire les mots considérés comme significatifs du corpus exploré. Les mots extraits sont alors conservés dans une base de données organisée selon la forme d'un index terminologique. Les termes significatifs sont associés à une valeur dit « poids » qui correspond à une probabilité d'apparition, à la fréquence du mot dans un document.

- La recherche correspond à la partie requêtes du moteur, qui restitue les résultats. Un algorithme est appliqué pour identifier dans le corpus documentaire, et selon l'index, les

⁸ http://news.netcraft.com/archives/2009/04/06/april_2009_Web_server_survey.html

⁹ Olivier Ertzscheid, Gabriel Gallezot et Éric Boutin, *Perspectives documentaires sur les moteurs de recherche : entre sérendipité et logiques marchandes*, 2007, en ligne : <http://hal.archivesouvertes.fr/docs/00/17/21/69/PDF/ertzsgallbout.pdf>

documents qui correspondent le mieux aux mots contenus dans la requête. Les résultats sont présentés par ordre de pertinence supposée.

Les moteurs de recherche ont connu des évolutions dans leur mode de fonctionnement, ceux dits de première génération, reposent sur l'analyse du contenu des pages indexées et se basent exclusivement sur le croisement entre les mots-clés de la requête et ceux récupérés sur les pages indexées. Les moteurs de deuxième génération tentent de prendre mieux en compte la structuration du Web où la notion d'hyperliens et d'hypertexte est déterminante. Cette étape permet d'aboutir à la création du fameux algorithme du Page Rank, rendant possible une recherche d'information « augmentée », que nous détaillerons plus loin. Toutefois, les principaux moteurs de recherche continuent d'offrir une vision limitée de l'organisation informationnelle du Web, ne retenant pour l'affichage des résultats que la forme de la liste et non leur catégorisation ou ordonnancement « raisonné ».

Les moteurs de dernière génération reposent sur l'idée que l'expression du besoin d'informations peut être affinée par l'internaute à travers l'expression de dimensions complémentaires au sujet de la recherche. Ainsi une page Web peut être décrite par son thème mais aussi selon son niveau plus ou moins commercial, sa fonction de page de contenu ou de pages de liens. De nouveaux projets de moteurs de recherche dits moteurs « sociaux » de type Rollyo¹⁰, Eurekster¹¹, Criteo¹², etc., apparaissent. Leur principe est d'ajouter un filtrage en amont des sources comme un modèle de pertinence différent, issu des usagers eux mêmes. Selon une logique similaire d'autres offres se développent aujourd'hui et proposent des types de recherche plus ciblés basées par exemple sur la recommandation dont Amazon fut le premier exemple en ligne, ou de catégorisation comme Exalead¹³, permettant une recherche générale sur le Web, ou directement sur Wikipedia.

Google a depuis sa création déjà largement fait évoluer l'ergonomie de sa page de résultats, notamment avec le développement de la recherche universelle. Celle-ci offre une présentation verticale et permet de réunir des résultats de Google Images, Youtube, Google Actu, etc. et de les présenter au sein des pages de recherche générales du moteur. L'internaute est ainsi confronté sur une même page aux résultats disponibles dans les différents outils Google, ce n'est plus seulement le rendu des correspondances écrites qui

¹⁰ <http://rollyo.com/>

¹¹ <http://www.eurekster.com/>

¹² <http://www.criteo.com/>

¹³ <http://www.exalead.fr/search/>

est privilégié, mais l'ensemble des contenus multimédias qui est interrogé, dans la mesure où l'indexation le permet. Dans ce contexte de renouveau de la recherche, il est important pour ce moteur d'optimiser l'indexation des contenus multimédia des sites Internet, car la recherche universelle ouvre de nouvelles perspectives stratégiques en termes de visibilité notamment et par conséquent de nouvelles perspectives économiques pour la société.

2. Extension des espaces documentaires indexables

Vers la fin des années 90, le Web est un espace de programmation statique et complexe, le lieu de prédilection des Webmasters. Le Web dit public, indexé par les moteurs, connaît alors une variété assez large des contenus et rassemble différentes formes documentaires où les articles scientifiques côtoient les pages commerciales, les sites de presse et d'autres documents factuels.

Une première évolution du Web est notable avec l'apparition de pages pouvant être générées dynamiquement, c'est-à-dire à la demande, contrairement au site Web statique. Si ce système offre une plus grande souplesse dans l'intégration des contenus il se confronte dans un premier temps à des soucis d'indexation de ces pages dynamiques. Se développe alors peu à peu un Web dit « opaque », invisible aux moteurs et donc non indexé. Ces premières pages dynamiques sont pour la plupart basées sur des informations issues d'organismes privés, « qui déploient, à partir de bases de données antérieurement constituées, des pages Web dynamiques (...). Il s'agit dans ce dernier cas de documents générés à la volée, dynamiquement, à partir de requêtes déposées sur les sites par les utilisateurs ¹⁴ ». Ces contenus sont donc purement virtuels et n'ont pas d'inscription physique stable puisqu'ils sont pour la plupart des hyperliens de navigation générés à la demande, et pouvant différer d'une visite à l'autre. Si les problèmes de l'indexation des pages dynamiques se résolvent, de nouvelles limites à l'indexation des contenus se révèlent peu à peu sur le Web. Les documents en format PDF par exemple ont échappé pendant longtemps à toute indexation par les moteurs puis le problème se posa à nouveau pour les pages en Flash notamment. Ainsi semble-t-il que l'évolution technologique du Web intervient et interfère toujours avec l'accessibilité des documents engendrant une scission entre un Web visible et un Web opaque.

¹⁴ Olivier Ertzscheid et Gabriel Gallezot, « Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », *Document numérique et société*, ADBS Éditions, 2006, p 3, en ligne : http://affordance.typepad.com/mon_Weblog/files/sdndocsoc.pdf.

Une autre rupture est à prendre en compte en lien avec le statut même des documents sur le Web. Nous sommes confrontés à l'existence d'un Web public d'un côté, indexant des documents librement consultables via les index des moteurs de recherche, et à un Web dit privé de l'autre. Cette privatisation et donc restriction d'accessibilité d'une partie du Web, est due à la nature même des documents en jeu ou aux politiques de confidentialité qui y sont déployées. Les pratiques informationnelles consistant à échanger des courriers électroniques ou à stocker des documents de travail sur son disque dur personnel échappent notamment à la problématique d'indexation des contenus du Web. Pour Gallezot et Ertzscheid « La question qui permet alors de scinder la masse documentaire en visible / invisible, indexée / non indexée est alors encore celle de la nature des contenus informationnels : « Quoi » indexer ? ¹⁵».

Toutefois et depuis peu les divers moteurs de recherche proposent des outils de plus en plus précis et axés sur le développement de l'indexabilité de tout type de documents y compris de nos données personnelles. Ainsi les informations contenues dans l'ensemble de nos courriers électroniques personnels ainsi que les fichiers et documents stockés sur nos ordinateurs, sont désormais accessibles aux moteurs, lesquels les indexent aussitôt par le biais de ces outils dédiés que sont Google Mail, ou encore Google Desktop. Et les auteurs de nous mettre en garde contre les dérives possibles de cette multiplication d'offres, « Un seul et même outil – ce qui constitue un gain – mais surtout, une seule et même société commerciale – ce qui constitue un risque - garantit l'indexation et l'accès à cet ensemble. (...) en termes d'accès et de droit à l'information, l'extrême mouvement de concentration qui touche ici la médiasphère est à tout le moins problématique.¹⁶ ». C'est à une concentration de tous les outils et données documentaires entre les mains d'un nombre restreint de sociétés à laquelle nous assistons, nouvel état de fait pouvant remettre en cause la notion déterminante de libre circulation des informations sur le Web.

Au-delà de l'aspect économique, un autre changement notable intervient et touche la notion même de document. En effet ce ne sont plus seulement des contenus validés par un processus éditorial (scientifique ou commercial) qui sont indexés et accessibles sur le Web, mais l'ensemble des entités émises dans un format numérique qui peut se trouver sur la toile. Ces nouveaux documents se définissent stricto sensu par leur capacité à être indexés, une capacité qui les constitue alors comme autant d'unités documentaires.

¹⁵ Ibid. *loc cit.*

¹⁶ Ibid. p 4.

De manière générale de nouveaux usages informationnels apparaissent sur le Web, si l'extension des techniques d'indexation modifie la teneur des documents du Web, la large ouverture de celui-ci sur les contenus directement engendrés par les internautes entérine le changement de paradigme documentaire. Après le Web statique, le Web dynamique, apparaît le Web 2.0, où se sont, à travers le phénomène des blogs dans un premier temps, les internautes eux-mêmes qui génèrent le contenu du Net. Ainsi « Avec cette nouvelle configuration de nos espaces informationnels, émerge une nouvelle écologie cognitive, un nouvel habitus, qui fixe l'interpénétration de deux sphères : celle de la recherche d'information « publique » (Web) et celle de la recherche d'information sur des données personnelles (Web privé et Web intime).¹⁷ ». Le Web 2.0 engendre un nouveau paradigme dans la recherche d'informations, une recherche qui concerne l'accès à des contenus fondamentalement différents dans leur conception, dans leur usage, dans leur nature et de fait dans leur possibilités d'indexation.

B. Indexation des contenus et publicités contextualisées

En France, le moteur Google représente à lui seul près de 85%¹⁸ de l'ensemble des recherches effectuées et 60%¹⁹ à l'échelle mondiale, les différentes études révèlent une claire domination du moteur. Ce monopole est non seulement économique mais également culturel, et se reflète largement au travers de l'adoption généralisée de l'expression « Googler quelque chose ou quelqu'un », de son passage dans le langage courant pour désigner l'ensemble du processus de recherche sur Internet. La question est alors de comprendre comment Google a réussi à asseoir sa suprématie, sur quels principes se distingue-t-il ?

1. Double articulation et indexation textuelle

Le fonctionnement de la majorité des moteurs de recherche grand public se base sur le principe d'indexation textuelle et de recherche plein texte. L'efficacité des technologies textuelles les plus élémentaires mises en œuvre par les moteurs de recherche, comme par exemple la recherche en texte intégral, est due à la double articulation du texte, qui est à

¹⁷ Olivier Ertzscheid, Gabriel Gallezot et Éric Boutin, *Perspectives documentaires sur les moteurs de recherche : entre sérendipité et logiques marchandes*, 2007, en ligne <http://hal.archivesouvertes.fr/docs/00/17/21/69/PDF/ertzsgallbout.pdf>

¹⁸Sources : [http://www.comscore.com/Press_Events/Press_Releases/2008/07/Top_French_Search_Engines/\(language\)/eng-US](http://www.comscore.com/Press_Events/Press_Releases/2008/07/Top_French_Search_Engines/(language)/eng-US)

¹⁹Sources : http://www.comscore.com/Press_Events/Press_Releases/2007/07/Top_Worldwide_Web_Properties.

envisager en écho de la double articulation de la langue déjà soulignée par les linguistes. Pour Bruno Bachimont l'efficacité aujourd'hui observable de la recherche textuelle sur la plupart des moteurs est due à la forme même du texte qui sert de structure matricielle à l'algorithme des moteurs. Ainsi il souligne comment « le texte, sous la forme que nous lui connaissons dans les langues occidentales, résulte d'une écriture alphabétique, constituée d'un répertoire fini de symboles permettant, par leur combinaison, d'exprimer tous les contenus de la langue. Par ailleurs, certains assemblages de ces symboles sont particulièrement stables et constituent ce qu'on peut appeler en première approximation des mots.²⁰ » Le texte obéit donc à une double articulation, alphabétique d'un côté et lexicale de l'autre.

Le répertoire fini de symboles, que constitue l'alphabet, assure que tout contenu textuel s'exprime selon une combinaison de ces symboles. Une requête visant un contenu textuel sera par conséquent exprimée par l'intermédiaire de ce même répertoire, de plus « le référentiel stabilisé permet de surmonter l'infinité des combinaisons possibles de ces symboles dans la mesure où ce répertoire est plus ou moins commun au producteur du document et à son lecteur, ou bien à l'utilisateur d'un moteur de recherche. Cette continuité entre l'auteur et le récepteur, fondée sur le partage d'un même référentiel, permet d'exprimer des requêtes possédant un écho dans les documents que l'on recherche.²¹ ». Les algorithmes des moteurs de recherche se basent donc sur les termes de la requête pour trouver leurs correspondances au sein de la base de données. Selon la combinaison de ces symboles est répertoriée la position exacte de chaque occurrence dans la base de données de documents.

Si la technique est au point, permettant à partir de quelques mots clés basiques d'accéder à une grande collection de documents, le problème principal des requêtes effectuées sur des moteurs de recherche textuels sont les retours non pertinents. À l'origine du problème se trouve l'ambiguïté lexicale, comme la synonymie, qui peut entraîner des résultats non utilisables par l'utilisateur. Des améliorations ont donc été ajoutées à la fonctionnalité de fouille intégrale de texte sur laquelle se base les moteurs. Parmi ces améliorations l'on trouve l'utilisation des mots clés. Il fut ainsi préconisé aux créateurs de documents ou aux indexeurs de fournir une liste de mots décrivant le sujet du texte, incluant des synonymes des mots qui définissent le sujet. Ce système améliore le retour

²⁰ Bruno Bachimont, « Image et audiovisuel : la documentation entre technique et interprétation. Critique et perspectives », *Documentaliste-Sciences de l'information*, 2005/6, Volume 42, p. 348-353.

²¹ Ibid.

des résultats, notamment lorsque la liste de mots clés inclut un mot de recherche qui n'est pas dans le document textuel mais qui permet néanmoins de catégoriser le contenu.

Les moteurs de recherche furent encore perfectionnés par la mise au point de requêtes booléennes. Les recherches se basent alors sur des opérateurs booléens (et, sauf, ou, etc.) qui permettent d'accroître considérablement la précision d'une recherche plein texte. Parmi ces termes descripteurs l'opérateur « et » permet d'associer les termes, l'opérateur « sauf » est un outil d'exclusion lexicale. Si les résultats se réfèrent à trop peu de documents, l'opérateur « ou » peut être utilisé pour accroître les retours en élargissant le sujet ou les thèmes possibles. Mais cette méthode, largement utilisée par les moteurs de première génération, et toujours opérante, atteint vite ses limites sur des corpus volumineux.

Les moteurs plus évoluées utilisent la formule TF-IDF²² (Term Frequency-Inverse Document Frequency), pour mettre en perspective le poids des mots dans une requête avec ceux contenus dans les documents. Il s'agit donc d'une formule de pondération qui calcule la fréquence des termes au sein du document, une variante à ce système s'appuie sur la fréquence maximale des termes afin de concilier l'aspect multi-thématiques d'un document avec des documents mono thématiques

Les moteurs de seconde génération comme Google, s'appuient non seulement sur la fréquence des termes pour indexer les pages Web mais aussi sur la popularité des sources en calculant un indice de notoriété des pages, le Page Rank. En naviguant de lien en lien, les robots indexent les termes utilisés par une page Web pour décrire une autre page Web. À chaque fois qu'un utilisateur suit ce lien, il "vote" la pertinence des descripteurs utilisés. Par exemple, l'algorithme PageRank de Google, qui grâce à son efficacité fait le succès du moteur, fournit plus d'importance aux documents qui renvoient, au travers de liens hypertextes, à d'autres pages Web.

2. Une indexation à visée économique

Si l'accès à la connaissance et aux savoirs est évidemment un enjeu motivant les diverses recherches dans le champ de l'indexation, l'enjeu économique est un facteur décisif dans la recherche d'automatisation de l'indexation. Les moteurs de référencement et de recherche de documents tels Google trouvent leur viabilité et rentabilité par l'apport des revenus générés par la vente de technologie mais surtout par la publicité, une publicité désormais de plus en plus contextuelle. Le développement de cette dernière est l'un des

²² Pour plus d'informations consulter <http://fr.wikipedia.org/wiki/TF-IDF>

fers de lance de Google, la société a en effet mis au point divers systèmes permettant à la fois à l'annonceur et au gérant de site de placer des publicités le plus en lien possible avec les contenus des pages.

*AdWords*²³ est le nom du système publicitaire du moteur de recherche *Google* qui permet aux annonceurs de placer leurs publicités de façon ciblée. *AdWords* provient de « Ad » pour *Advertising* (Publicité) et *Words* (mots), le principe repose sur l'achat de mots clés par l'annonceur en rapport avec son activité. Prenons en exemple une agence de voyage, celle-ci peut acheter des mots-clés comme « vacances », « hôtel » et « plage » ou des noms de villes ou de pays dont elle est spécialiste. Cet achat permet d'obtenir un référencement dit "référencement payant" à distinguer du référencement dit "référencement naturel" engendré notamment par le page ranking. Lorsque les internautes effectuent des recherches sur Google à l'aide de l'un des mots clés utilisé par le publicitaire, son annonce est alors susceptible d'apparaître à côté des résultats de recherche. Google met également à disposition des annonceurs un générateur de mots clés. Cet outil recense de façon automatique une liste de requêtes courantes effectuées sur le moteur de recherche de Google, en rapport avec le mot clé saisi par le publicitaire.

En parallèle de *AdWords*, Google a également développé un système cette fois tourné vers le gestionnaire de site Web, *AdSense*²⁴, autre régie publicitaire de Google qui utilise donc les sites Web comme support pour les annonces. *AdSense* propose aux entreprises ou aux particuliers possédant un site Web d'explorer automatiquement le contenu des pages et de proposer par la suite des annonces (textuelles et/ou illustrées) correspondant au public et au contenu du site. Pour accroître la pertinence des annonces, *AdSense* ne se base pas sur un seul mot, mais sur plusieurs et cherche à trouver le point commun entre eux pour déterminer le sens du contenu d'un texte.

Le moteur de recherche peut afficher la publicité de deux manières : en encart séparé ou en l'intégrant aux résultats de la recherche. Pour le visiteur, l'encart séparé se présente comme une publicité classique. L'intégration aux résultats se fait en revanche au détriment de leur pertinence et peut avoir des retombées négatives sur la qualité perçue du moteur. De ce fait, tous les moteurs ne vendent pas de placement dans les résultats.

²³<https://www.google.com/accounts/ServiceLogin?service=adwords&cd=FR&hl=fr-FR<mpl=adwords&passive=true&ifr=false&alwf=true&continue=https%3A%2F%2Fadwords.google.fr%2Fselect%2Fgaiaauth%3Fapt%3DNone%26ugl%3Dtrue>

²⁴https://www.google.com/adsense/login/fr/?sourceid=aso&subid=ww-fr_FR-et-awhome&medium=link&hl=fr_FR&gsessionid=sEQjZFR5RAoxfzoKTroK8Q

Pour Tim O'Reilly, la vente d'espace publicitaire est le cœur du système économique de Google. Le trait de génie de Google, qui permet à la firme de largement dépasser les autres moteurs de recherche sur le marché en matière de monétisation d'encarts, a été de comprendre que vendre les meilleurs espaces publicitaires aux plus gros enchérisseurs était en réalité une perte d'argent. Et Tim O'Reilly de préciser, « étant donné que les annonceurs ne paient qu'en fonction des clics, Google se rendit compte que s'ils pouvaient déterminer le nombre probable de clics sur une publicité, ils pourraient vendre l'espace le plus en vue à la meilleure combinaison de prix et de clics. Par l'instrumentalisation, la mesure et l'adaptation au taux de clics, Google rendit la vente de publicité plus intelligente, le tout grâce à l'implication implicite des internautes.²⁵». Parce que Google a su prendre en compte l'importance de la position d'une annonce sur une page Web, les usages et habitudes des usagers, parce qu'ils réussirent à traduire leurs intuitions en outils rentables, la firme se place aujourd'hui comme leader économique et technologique.

Les systèmes Adwords et Adsense, créés par Google, reposent pour le premier sur l'achat par les entreprises de mots-clés permettant l'apparition de leurs annonces sur la page de résultats Google et pour Adsense sur l'analyse du contenu des pages de sites permettant d'y loger les publicités les plus en lien. On comprend donc que si ce système est des plus performants en ce qui concerne les documents textuels, puisqu'ils se basent tous deux sur une analyse plein texte et sur la forme du mot clé, ces applications connaissent de vraies lacunes quant à leur application sur des documents visuels. D'une part parce que peu d'images sont renseignées par mots-clés et d'autre part parce qu'il est pour l'instant impossible d'analyser automatiquement le contenu d'une image fixe ou animée, il va s'en dire qu'au vu du nombre de nouvelles images circulant sur le Web chaque jour cette opération ne peut se faire manuellement. Les technologies d'indexation et de descriptions des images par leur contenu apparaissent donc pour Google comme une voie salutaire dans la rentabilité de leurs services à une époque où par exemple le site Youtube, pour les raisons précédemment invoquées, s'avère déficitaire.

Depuis 2008, Youtube a décidé d'utiliser AdSense pour monétiser ses contenus, en faisant s'afficher des publicités en bas des fenêtres vidéo, toutefois rien n'indique encore que la solution sera suffisante pour rentabiliser le trafic. Ce choix de régie publicitaire est loin d'être neutre, comme l'explique Rémi Douine « la valeur publicitaire d'un contenu sur Internet, même quand il s'agit d'une vidéo, ne peut être réalisée par l'exploitation de

²⁵Tim O'Reilly, "Why Dell.com (was) more enterprise 2.0 than Dell Ideastorm", *O'Reilly Radar*, 29/09/2008, en ligne <http://radar.oreilly.com/2008/09/why-dell-dot-com-is-more-enterprise.html>

l'information descendante du contenu (meta-tags des vidéos choisis par les internautes ou description du contenu ou sélection d'un contenu par un publicitaire) mais bien par l'exploitation et la mise en avant de l'information remontante des internautes (mots-clés que les internautes recherchent à un instant T)²⁶». Il faut donc prendre en compte, les utilisateurs, le contenu, mais aussi les pratiques développées par les usagers sur ce contenu et leur appréciation. Les possibilités de rentabilité des espaces documentaires visuels de Google semblent donc dépendre non seulement d'une possible contextualisation des annonces, et donc d'une analyse du contenu visuel, mais aussi et de façon déterminante de l'estimation de l'intérêt des internautes de ce même contenu.

II. L'indexation des documents visuels par les moteurs de recherche

A. Accroissement des contenus visuels par le Web 2.0

L'une des dernières évolutions majeures du Web est son ouverture à l'indexation de contenus divers issus des multiples sphères, publique, privée/professionnelle ou encore intime qui nous caractérisent. À la variété des sources de ces documents en constante circulation vient s'ajouter la diversité de leur nature dont la présence toujours accrue de l'image sur le Web constitue un des changements de régime documentaire les plus observables.

1. Nouvelle approche du document

Le Web opère un changement de la notion de document, l'indexabilité élargie des contenus semble favoriser une transgression de l'ordre documentaire ordinaire. Le Web «documentarise» des expressions qui relevaient autrefois de l'intime et de l'éphémère en les enregistrant, les traçant, les indexant, et la frontière se brouille alors entre privé et public, document et donnée factuelle. Toutefois le chamboulement dans la perception du documentaire trouve son origine avant l'essor du Web, medium ne faisant qu'accélérer le processus. En effet l'apparition du numérique ne fut pas sans apporter de changement vis-à-vis de la notion même de document et par conséquent au niveau également des méthodes de gestion et de traitement de l'information.

L'une des premières transformations apportée par le numérique à la notion de document, se distingue au niveau de sa forme tout d'abord, ce que le collectif RTP DOC

²⁶ Rémi Douine, «Extension du domaine de la recherche vidéo », *ARHV*,8/12/2008, en ligne <http://www.arhv.lhivic.org/index.php/2008/12/08/885-extension-du-domaine-de-la-recherche-video>

nomme la matérialité du document. À considérer le document à l'échelle des sciences de l'information, de la bibliothéconomie et de l'archivistique, il se définit comme une ressource matérielle manipulable, vecteur d'informations ou faisant office de preuves. Son aspect dominant est le support papier lui permettant une transmission d'info le plus souvent textuelle, « Le document a été construit comme un objet dont la concrétisation la plus banale est la feuille de papier, au cours d'un processus séculaire où se sont entrelacés outils, savoirs et statuts. Depuis quelques dizaines d'années avec le numérique, nous sommes passés dans une nouvelle période dont certaines caractéristiques sont en filiation directe avec la phase précédente tandis que d'autres marquent au contraire un changement radical et peut-être l'émergence d'une notion différente reprenant tout ou partie de l'utilité sociale de ce que nous appelions document.²⁷ ».

Jusqu'à l'arrivée du numérique nous pouvons ainsi stipuler une emprise du texte et du document papier sur la sphère documentaire. Or le document numérique qui déferle désormais au sein des réseaux ne saurait se caractériser par une quelconque matérialité mais implique au contraire, dans le cas d'une numérisation, une rupture du document avec son support originel. Le document devient une représentation d'informations, un enregistrement en langage binaire. L'un des premiers changements apporté par le numérique est donc d'ordre matériel, touchant à l'intégrité même de ce qui constituait, depuis l'apparition de l'écriture, l'apanage de la science documentaire. Aujourd'hui le document dépasse de bien loin la seule considération du texte sur support papier pour couvrir ce que l'on nomme déjà le multimédia. En considérant que, désormais, tout document existant est voué à la numérisation, que toute production matériellement tangible d'un document est forcément accompagnée de son double numérique, et enfin que certains documents ne connaissent qu'une existence digitale, le document ne doit plus être uniquement considéré du point de vue de sa matérialité. On assiste à un élargissement de la notion, le numérique dans l'abandon du support natif qu'il inflige au document, pour sa transcription en représentation binaire et unique, oblige à le considérer désormais en dehors de son lien au support originel.

De plus par la mixité des médias qu'il permet, le document prend aussi bien en compte les dimensions iconiques qu'auditives ou textuelles de la ressource, le texte n'étant plus le seul langage vecteur d'informations. L'image devient par conséquent dans sa forme numérique une unité documentaire au même titre qu'une source numérisée textuelle

²⁷ Roger T Pedauque, *Document : forme, signe et medium, les re-formulations du numérique*, Paris, ENSSIB, 2003, en ligne http://archivesic.ccsd.cnrs.fr/docs/00/06/21/99/PDF/sic_00000511.pdf

puisqu'ils partagent tous deux le même encodage binaire. De par sa nature numérique et sa fonction de document l'image fixe se trouve donc soumise au principe de la recherche et répond par conséquent à des nécessités de classification, d'indexation : seul système pouvant assurer son « utilisabilité » au sein de collections enrichies en permanence.

La fragmentation de la notion de document ouvre la possibilité d'une « redocumentarisation massive²⁸ », laquelle mixe de manière inédite et avec une perméabilité nouvelle des espaces et des marqueurs sociaux jusque-là bien distincts. Toute trace, inscription, support, flux ou échange numérique est aujourd'hui à même de devenir document, la condition étant son apparition même fugitive ou volontairement éphémère sur le réseau. Une image, un texte, un extrait de page, un courrier électronique, quelques lignes de conversations sur messagerie instantanée, une vidéo, etc. sont aujourd'hui des entités documentaires par leur inscription au sein du Web.

Aujourd'hui le Web se comprend plus que jamais comme l'espace du multimédia, où les images, les vidéos et les sons ont une place privilégiée pouvant permettre un enrichissement des documents textuels. Il suffit pour s'en convaincre de considérer les formules en ligne des principaux quotidiens d'informations qui usent de ces mediums comme compléments essentiel aux articles. Les images ont envahi le Web dans leur forme animée comme mouvante, en support d'information ou de médiation. Elles accompagnent la plupart des sites officiels, comme forme d'expression personnelle et de communication globale. Leur nombre explose littéralement sur les blogs personnels et autres plateformes de partage grand public.

2. L'explosion visuelle du Web

Une transformation profonde du Web se dessine depuis peu avec l'introduction massive des documents visuels et multimédias issus des productions personnelles des internautes. Ce renouveau du paysage de la toile dépend majoritairement de deux facteurs, la démocratisation des Appareils Photos Numériques²⁹, dont les coûts proportionnellement à leurs capacités ne cessent de baisser ; et d'autre part l'accessibilité accrue des foyers au réseau Internet et au haut débit. Ces changements d'ordre sociaux-économiques influent non seulement sur le contenu du Web mais également sur l'essence même de la pratique photographique voire dans une perspective plus ontologique sur la notion d'image même.

²⁸Jean-Michel Salaün, *Le défi du numérique : redonner sa place à la fonction documentaire*, ADBS, 2008, en ligne <https://papyrus.bib.umontreal.ca/jspui/bitstream/1866/2225/1/salaun-jm-defi-documentaliste-mktg.pdf>

²⁹ Cf. les chiffres de vente des appareils année par année sur le site du SIPEC : <http://www.sipec.org/>.

Aujourd'hui l'image non seulement se partage mais semble plus que jamais être faite pour être partagée, et ainsi est-il permis d'observer un changement de paradigme au niveau de la production des images.

Les photos numériques ne sont plus seulement vouées à paraître sur papier mais se regardent directement sur l'écran (que se soit celui de l'appareil ou de l'ordinateur), il résulte de ce changement une baisse des coûts de la pratique photographique (dans l'hypothèse où le foyer est déjà équipé d'un ordinateur) et par conséquent une augmentation significative des prises de vues par famille alors stockées, classées sur le disque dur ou de façon plus récente mises en ligne sur des plateformes de partage.

Une autre mutation radicale des pratiques touche la circulation, le partage de ces images. Jusqu'à l'avènement du numérique la photographie issue des pratiques amateurs était vouée, dans le cas d'une pratique privée à rester dans le cadre familial, les photos rangées dans des albums se regardaient lors de réunions, de rituels familiaux. Objet de remémoration des événements passés, illustration de la généalogie familiale, supports de souvenirs des êtres chers et des lieux de son histoire personnelle, la photographie privée, n'avait pas vocation à transmettre un savoir visuel en dehors de l'univers familial³⁰. En outre, si l'on considère les pratiques éclairées du médium, les images circulaient alors au sein des photoclubs, lors de rencontres régionales, nationales des sociétés photographiques mais trouvaient difficilement une visibilité publique. Seules les meilleures d'entre elles pouvaient prétendre à une publication dans certaines revues spécialisées au public très restreint³¹. Quelle que soit la pratique envisagée la photographie amateur n'avait pas de reconnaissance publique.

L'image produite par un photographe amateur fait aujourd'hui pleinement partie du paysage d'Internet, la mise en ligne des photos offre aux particuliers ce qu'ils n'avaient jamais eu jusque là : un public et une visibilité. Si la création amateur se définissait jusqu'à présent par une pratique de loisir, ayant pour seule fin une diffusion de ces productions dans la sphère privée, ou dans un cercle restreint de connaisseurs, il faut donc convenir qu'une transformation du statut même de l'amateur est en train de s'opérer, au travers des pratiques de diffusion de ses travaux. Regarder les images des autres est perçu désormais comme une forme de plaisir et d'activité sociale. L'image partagée devient le lieu de sociabilité mais aussi de créations. En recherchant la visibilité pour ses propres travaux un

³⁰ Sur ce sujet voir notamment : Irene Jonas, « la photographie de famille au temps du numérique », *Enfances, familles, Générations*, n°7, automne 2007.

³¹ Sur ce sujet voir notamment Christian Joschke, « Aux origines des usages sociaux de la photographie », *Actes de la recherche en sciences sociales*, n°154, avril 2004, p 53-65.

amateur est souvent amené à nouer des liens avec d'autres internautes, la circulation d'images au sein des réseaux du Web entraîne ainsi souvent des pratiques nouvelles de recyclage ou remix, des participations nouvelles de la part des amateurs se constatent également au niveau des discussions, commentaires, annotations encadrant les images. Loin de ne demeurer qu'un utilisateur passif des techniques, l'amateur à l'ère du numérique est en instance de devenir l'un des acteurs majeur de la culture populaire s'exprimant sur Internet. S'inscrivant dans une perspective participative, la pratique vernaculaire de la photographie se révèle être un espace de création, dont le vecteur principal est l'échange, la circulation des images³².

Ce renouveau des pratiques, qui prend désormais en compte un travail de post production inhérent à la diffusion des images (transfert sur l'ordinateur, classement, gestion des dossiers, des comptes d'utilisateurs sur les plateformes, les photoblogs, réseaux sociaux, etc.), est dû non seulement à la technologie du numérique, mais également au développement d'Internet. La première voie de partage des images numériques a été le mail, dans un premier temps, la taille importante des fichiers images et les connections bas débit rendaient la tâche longue et fastidieuse. C'est finalement l'arrivée de l'internet rapide, les capacités de stockage et d'envoi accrues des boîtes mails, qui ont provoqué un premier renversement au sein des pratiques photographiques, que se soit au niveau du visionnage ou de l'échange autour de l'image. Mais c'est véritablement l'apparition des plateformes de partage, des logiciels de gestion des images en ligne et du phénomène des blogs qui provoquent une déferlante images sur le Web. Les différents sites existants aujourd'hui proposent aux internautes non seulement des espaces de stockage des images mais aussi des outils d'organisation des corpus voire de retouche et enfin des technologies de transfert des images gratuits et performants.

L'émergence de ces plateformes et autres photoblogs est concomitante à la phase de transformation du Web qui, dès 2005³³, se mue du Web classique en Web 2.0. Reposant sur une interactivité entre les sites et sur la participation des internautes à la formation de leurs contenus, ce nouveau Web a engendré et rendu possible ces nouvelles pratiques photographiques des amateurs. L'arrivée du numérique, la possibilité d'envoyer des fichiers images par le Web, l'émergence de plateformes de stockage des photographies sur

³² Sur le renouveau de l'amateur voir Jean-louis Weissberg « l'amateur : émergence d'une figure politique en milieu numérique », *Multitudes*, n°5, mai 2001.

³³ Nous nous référons pour la date à l'article de Tim O' Reilly, « What is Web 2.0 », en ligne <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-Web-20.html>.

Internet comme Flickr³⁴, ont profondément modifié le rapport des familles, des particuliers à l'image. Nous sommes passés d'un régime de production de l'image, à un régime de diffusion.

B. Les systèmes d'indexation de la firme Google

L'ouverture toujours croissante du Web aux documents non plus seulement textuels mais visuels, induit par le changement et surtout l'élargissement de la notion même de document de nouvelles problématiques en matière d'indexation. Comment appliquer les technologies des moteurs jusque là dévoués au texte sur des contenus iconographiques ?

1. Chercher le texte avant l'image

La plupart des moteurs de recherche généralistes proposent aujourd'hui des solutions de recherche d'images sur le Web. Toutefois, et en conformité avec leur propre système de fonctionnement, l'indexation et la recherche de ces images s'effectuent en se basant majoritairement sur des annotations textuelles ou des éléments de texte pouvant être rattachés aux images. La principale particularité de la recherche d'images est qu'il n'est pas possible d'effectuer une recherche plein texte avec des mots-clés sur les données des images présentes dans une base de données d'images. Pour pallier cet inconvénient, deux solutions existent. En premier lieu, la recherche peut s'effectuer d'après les données textuelles associées à l'image, par exemple les données de la page Web où est présentée l'image. La solution opposée, que nous examinerons dans le dernier chapitre de ce mémoire, repose sur un principe d'indexation et de recherche effectué à partir du "contenu" des images, par une analyse des données des images.

Dans le cas de la recherche d'images à partir de texte associé aux images, principe le plus répandu au sein des moteurs de recherche basiques, les méthodes d'indexation et de recherche s'apparentent à celles utilisées classiquement dans une base de données textuelle. Simplement, dans le cas de la recherche d'images, les données indexées par les outils de recherche servant à "retrouver" les images, sont les informations textuelles se rapportant aux images. Ces informations textuelles collectées et indexées par les outils de recherche d'images sur internet peuvent se situer à deux emplacements. Tout d'abord ces éléments peuvent être extraits du code HTML des balises et des attributs spécifiques aux images. Ainsi la balise sert à indiquer l'insertion d'une image dans une page

³⁴ <http://www.flickr.com/>, pour un aperçu des principes d'utilisation du site et de ses fonctionnalités on peut également se référer à l'interview d'André Gunthert sur le site Internet Actuel, « Flickr une des choses les plus importantes qui soit arrivée à la photographie », en ligne <http://www.internetactu.net/>.

HTML, la balise `` mentionne quant à elle la présence d'un lien hypertexte vers une image. Enfin l'attribut `ALT="..."` représente le texte alternatif d'une image, il présente l'inconvénient de ne pas toujours être renseigné par les créateurs de pages HTML.

Enfin les informations pouvant servir à l'indexation des images peuvent se trouver dans le texte de la page associée ou dans le titre de la page « support », ce sont alors des éléments lexicaux, mots entourant l'image. Les moteurs de recherche du type Google ont un système d'indexation basé sur la forme du texte, en ce qui concerne l'indexation des images fixes c'est le texte entourant l'image qui sert de critère, que ce soit celui contenu dans les métadonnées, ou l'adresse, le titre et le texte de la page. Il convient également de noter que ce qui permet de repérer des images présentes sur les réseaux par les différents outils disponibles et de les différencier des autres données passe par la recherche des extensions de fichiers qui correspondent à celles des images comme les fichiers Gif, Jpeg, etc.

Il apparaît avec ces méthodes d'indexation que l'image demeure toujours dépendante du texte qui l'entoure, la recherche ne peut s'effectuer que selon des critères qui lui sont extrinsèques. C'est un lieu commun que de penser nos sociétés industrialisées, où le cinéma, la télévision, la photographie s'inscrivent en pratiques culturelles de masse, comme étant exclusivement visuelles. Ce qu'il ressort d'une observation simple d'Internet et du Web est que le texte domine encore largement l'image malgré sa présence croissante sur la toile.

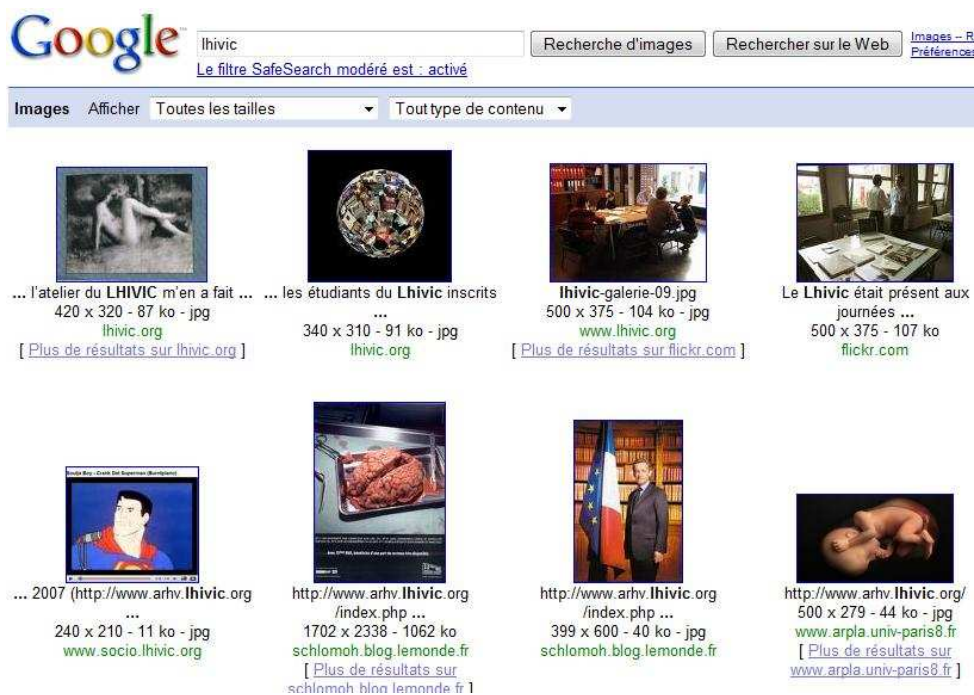
Pour exister physiquement sur un support numérique, chaque image porte un nom seul critère de son indexation, les moteurs de recherche, qui sont le moyen d'accès aux données numériques, étant essentiellement textuels. Et lorsqu'on cherche une image c'est par son titre que l'on y parvient. Cette domination interroge la relation complexe et historiquement stratifiée entre ces deux régimes de pensée. « Il y a toujours eu un entrelacs, ou si vous voulez un problème, une question, un *no man's land* entre l'alpha-numérique et l'iconographique³⁵ ». Là où le numérique a permis d'élaborer des outils automatisés de gestion des collections documentaires, d'aboutir à des moteurs de recherche performants pour des volumes de documents toujours croissants, son application à l'image, bien qu'ayant transformé les pratiques photographiques et vidéos, ne semblent pas pour l'instant avoir engendré un renouveau au niveau de la gestion des corpus iconographiques.

³⁵ Gregory Chatonsky, « La révolution a lieu à New-york », *Incident.net*, 03/07/2003, en ligne : <http://incident.net/users/gregory/wordpress/03-la-revolution-a-eu-lieu-a-new-york/>

Il perdure dans plusieurs cas, chaînes de télévision, journaux, Web un archivage et un traçage des images et des séquences vidéos ne pouvant se faire qu'au prix d'une étape d'annotation manuelle à l'aide de mots-clés. Cette indexation représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent aujourd'hui de plus en plus massives. En prenant en compte l'aspect fastidieux de ce type d'indexation, sa quasi incompatibilité avec les données visuelles du Web, mais aussi la subjectivité de cette tâche, dépendant de la culture, connaissance, et des sentiments de chaque personne, le besoin de méthodes d'indexation et de recherche automatique, directement basées sur le contenu de l'image n'est donc plus à démontrer.

2. Google Image et Google Image Labeler

L'outil de recherche d'images fixes de Google est aujourd'hui l'un des moteurs les plus utilisés. Google a choisi de dédier une page de son interface à cette fonctionnalité qui hérite donc de l'ergonomie simple et lisible du moteur. La recherche débute par une saisie de mots clés dans le champ imparti, les résultats sont retournés sous la forme d'une mosaïque de dix-huit miniatures, le nombre total d'images trouvées est quant à lui mentionné au dessus des résultats. Sous chaque image sont mentionnés le titre, la taille de l'image, le type de fichier et le nom du site hébergeur.



Ill. 1 Capture d'écran d'une recherche effectuée sur Google Image à partir du mot-clé LHIVIC.

La requête peut être affinée par la sélection de la taille des images (très grande taille, grande, moyenne, petite taille), par le type de contenu tel que contenu d'actualités, portrait, images clipart, dessin au trait, contenu photo. Depuis peu, le moteur permet également de sélectionner les images en fonction de la couleur souhaitée. En cliquant sur une miniature, une nouvelle fenêtre s'ouvre directement sur la page Web où se trouve l'image. Un cadre surplombe la page Web, où l'on retrouve la miniature et un résumé de ses propriétés, l'utilisateur peut alors suivre un lien conduisant à une nouvelle page où l'image est extraite de son contexte Web et présentée dans sa taille réelle. Il est possible d'effectuer une recherche avancée, plusieurs champs sont alors à remplir par l'internaute, dont des cadres dédiés à une requête booléenne. Il est aussi possible à ce stade de sélectionner le type de fichier désiré (JPG, GIF, PNG), ou de restreindre la recherche à un seul site en rentrant son adresse URL.

The screenshot shows the Google Images advanced search interface. At the top, the Google logo is on the left, followed by the text 'Images -- Recherche avancée' and a link 'À propos de Google'. Below this is a search bar containing the word 'flower' and a 'Recherche Google' button. The main interface is divided into several sections:

- Pages contenant:** Radio buttons for 'tous les mots', 'cette expression exacte', 'au moins un mot', and 'aucun de ces mots'. The search bar is positioned to the right of these options.
- Types de contenus:** A heading 'Envoyer des images représentant' followed by radio buttons for 'tous types de contenu', 'actualités', 'visages', 'contenu photo', 'images clipart', and 'dessins au trait'.
- Taille:** A dropdown menu currently set to 'Toutes'.
- Taille exacte:** Input fields for 'Largeur' and 'Hauteur', with a link 'Utiliser la taille de mon bureau' below them.
- Format:** A dropdown menu set to 'Tous formats'.
- Couleur:** A dropdown menu set to 'Toutes couleurs'.
- Domaines:** An empty input field.
- SafeSearch:** Radio buttons for 'Aucun filtrage', 'Filtrage modéré' (which is selected), and 'Filtrage strict'.

At the bottom center, there is a copyright notice: '©2009 Google'.

III. 2 Capture d'écran des options de recherche avancée sur Google Image

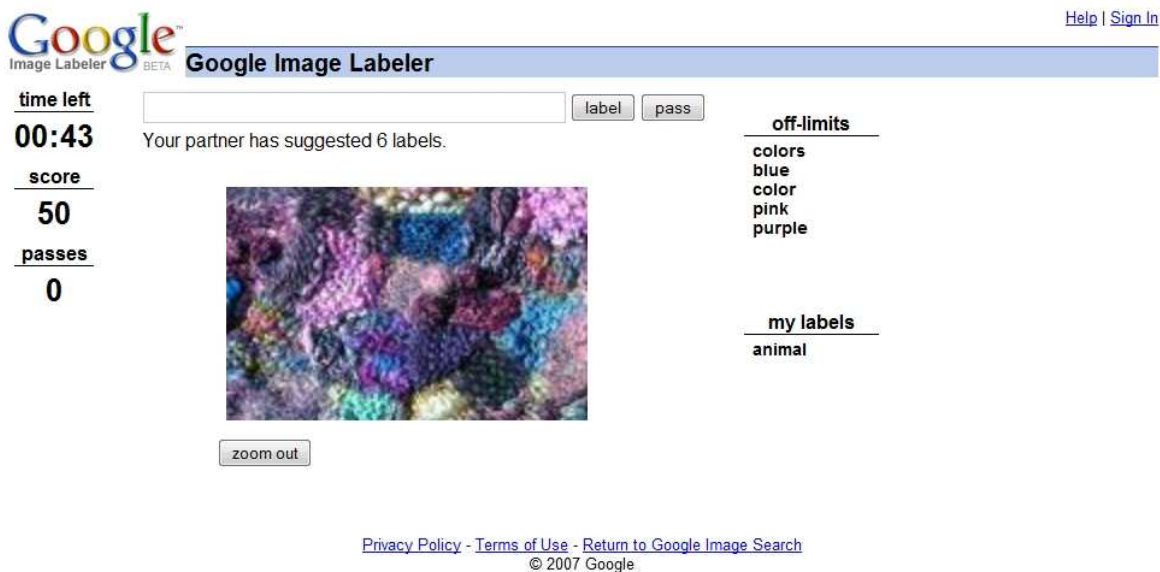
Si le site moteur propose une interface simple d'utilisation et efficace, il connaît toutefois quelques limites, notamment de pertinence, problème imparté au système d'indexation basé uniquement sur la recherche de termes associés aux images. Le principal manque de pertinence visible est celui dû au synonyme ou homonyme, des images représentant des objets différents mais référant au même mot seront présentées ensemble. L'expérience est simple à faire avec le terme « avocat », rassemblant alors des images aussi bien du fruit que de la profession. Cette limite n'est pas une faille spécifique au moteur mais se retrouve dans tous les systèmes utilisant une indexation en sémantique libre

et non hiérarchisée comme nous le verrons plus loin, de plus il est possible, en utilisant la recherche avancée de limiter sur cette interface ce genre de problème. Un autre type de retour non pertinent est quant à lui à incomber au fonctionnement du moteur. En se basant sur les termes de la requête croisés à ceux des pages Web avec images les comportant, il peut arriver que l'image retournée n'ait aucun rapport avec les termes soumis. Les mots se trouvaient bien sur la page mais ne se rapportaient pas directement à l'image, il convient ici aussi d'avoir recours aux outils de recherche experte pour amoindrir ces écueils.

Conscient des manques et limites de son système en termes d'indexation visuelle et dans une visée d'amélioration des services et de rentabilité accrue, Google Image se tourne en 2006 vers la pratique du taggage collaboratif. Ainsi est lancé Google Image Labeler³⁶, une application d'apparence ludique s'avérant même addictive, et qui suit le mot d'ordre de la firme "*Many hands make light work*". Le but est donc de faire participer les gens au (re)taggage des images, d'enrichir l'indexation visuelle de sa base de données par l'ajout de mots-clés choisis par les internautes et ce par l'intermédiaire d'un jeu.

Le principe est simple, le logiciel nous attribue un adversaire présent sur l'interface au même moment mais dont nous ne savons rien. Viennent ensuite les images, des photographies défilent de façon aléatoire sur l'écran, le jeu est de rentrer un maximum de mots clés pour définir chaque image. Si l'on entre le même mot-clé que son partenaire, celui-ci est donné " gagnant " et donc pris en compte par Google pour indexer à l'avenir cette image. De plus il sera désormais considéré comme mot tabou, les futurs joueurs tombant à nouveau sur l'image ne pourront pas l'employer. À la fin du jeu, qui dure 1 minute 30, un score correspondant au nombre d'images ainsi identifiées par les joueurs, et les meilleurs scores sont affichés en page d'accueil de ce "service". La folksonomie se met en place au sein de GI et permet une précision renforcée, une plus grande variété des mots-clés.

³⁶ <http://images.google.com/imagelabeler/>



III. 3 Capture d'écran d'une session de Google Image Labeler

3. Les solutions pour la vidéo

Si des solutions sont développées pour l'indexation et la recherche d'images fixes, le problème est tout aussi important du côté des contenus vidéos. Si bien sûr des vidéos aux fichiers bien documentés peuvent être retrouvées par leur titre, annotations, etc., la difficulté technique encore à résoudre concerne l'indexation des propos contenus dans ces vidéos. Pour l'instant les moteurs ne prennent pas en charge les contenus des vidéos, ce qui rend impossible de retrouver un passage particulier alors perdu au sein de la masse visuelle chaque jour mise sur le Web.

Depuis peu ont été mis au point les premiers outils permettant d'extraire les bandes sonores des vidéos, de déchiffrer les propos et d'indexer les vidéos d'après cette analyse. Le but est de permettre aux usagers de chercher de l'information dans les propos tenus dans les vidéos, de sélectionner les commentaires vidéos selon des mots clés ou des termes de recherche. Les premiers logiciels prenant en charge ce type d'indexation sont Blinkx³⁷ ou EveryZing³⁸, mais Google a depuis peu lancé son propre outil d'indexation audio.

Google Audio Indexing³⁹, est le nom de la technologie mise au point par la firme pour permettre aux usagers d'effectuer une recherche par mots-clés directement à partir du

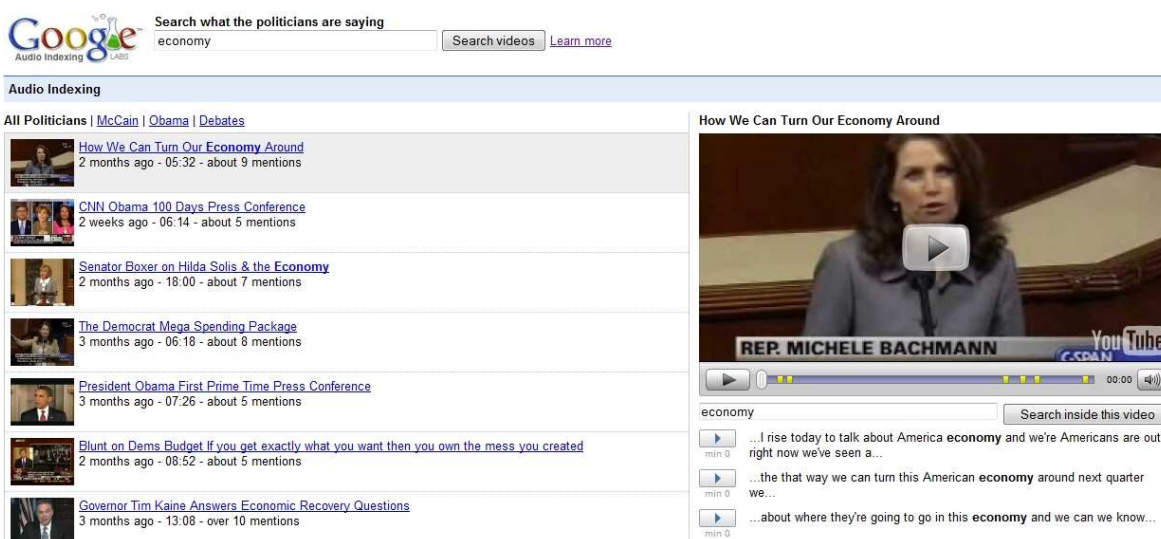
³⁷ <http://www.blinkx.com/>

³⁸ <http://search.everyzing.com/>

³⁹ <http://labs.google.com/gaudi>

contenu audio des vidéos. Pour l'instant encore en stade d'expérimentation, seules celles issues de YouTube Political channels ont été indexées.

Le principe repose sur la détection puis l'indexation de mots parlés à l'intérieur des images et leur transformation en texte indexable. L'utilisateur dispose d'une interface entièrement dédiée à cet outil pour effectuer ses recherches. Comme pour toute option de recherche Google, une barre de recherche est disponible dans laquelle l'utilisateur peut taper ses mots-clés. Une liste de miniatures correspondant aux vidéos comportant ces termes est alors affichée sur la gauche. Une fois une vidéo choisie, apparaît sur la droite un encart comportant un lecteur avec une vidéo chargée sous laquelle apparaissent sous forme textuelle, les passages où les termes sont mentionnés. Il est alors possible en cliquant sur le lien de visionner directement le passage désiré. La même opération peut aussi être effectuée en cliquant directement sur la ligne de défilement de la vidéo où les passages sont marqués en jaune.



III. 4 Capture d'écran de l'outil Google Audio Indexing

Si selon les propos même de Google leur mission est *"to organize the world's information and make it universally accessible and useful"*⁴⁰, le but de cette course à la technologie d'indexation est bien aussi certainement économique. Comme nous l'avons abordé précédemment, la plateforme Youtube connaît de vrais problèmes de rentabilité dus aux possibilités limitées d'indexation des vidéos et par conséquent au peu de retour publicitaire que les contenus suscitent. En accroissant les chances d'indexation et surtout

⁴⁰ <http://labs.google.com/gaudi/static/faq.html#how-it-works>

en effectuant cette étape directement à partir du contenu des vidéos, la firme pourrait enfin parvenir à lier pour chaque vidéo une annonce contextuelle pertinente.

Chapitre 2

L'indexation sémantique et la folksonomie

Sommaire du chapitre

I. L'apparition d'une indexation sociale	36
À. L'indexation des collections iconographiques	36
1. Une indexation sémantique surajoutée	36
2. Indexation, classification et thésaurus	38
3. Indexation et métadonnées.....	40
B. La folksonomie	43
1. L'Apparition de la folksonomie sur le Web	43
2. Les principes de la folksonomie	46
3. Une étude de cas : Flickr.....	48
II. Vers les nouveaux usages de la folksonomie	50
À. Les pratiques sociales et communautaires.....	51
1. Des tags aux groupes	51
2. Commenter par les tags.....	53
3. Tags et jugement de goût	55
B. Les détournements créatifs de l'indexation	57
1. Définition d'un mashup	57
2. Mashups et jeux de mots.....	58
3. Mashups artistiques.....	60
C. Flickr <i>Commons</i> et les nouvelles pratique documentaires	61
1. The <i>Commons</i> sur Flickr	62
2. Flickr comme moteur de recherche	66
III. Les limites de la folksonomie	68
À. Polysémie, synonymie et subjectivité.....	68
1. Problèmes de l'indexation sémantique des images.....	68
2. Polysémies des tags	69
B. Problème des multilinguismes et Restriction du champ sémantique	70
1. Régulation par les foules.....	70
2. Un vocabulaire limité, Folksonomie et multilinguisme.....	72

I. L'apparition d'une indexation sociale

Avec l'émergence du Web 2.0 se sont développés depuis moins de 10 ans de nombreux sites de partage de photographies ouverts au grand public. Ces sites d'hébergement de photographies ont connu un succès fulgurant et comprennent pour les plus importants des bases de données comptant souvent plusieurs centaines de millions d'images. Face à ce flux d'images quotidiennement téléchargées ces sites ont dû intégrer des systèmes d'indexation fiables et performants pour faciliter le classement et la recherche de photographies auprès d'un public souvent non spécialiste de recherche documentaire.

A. L'indexation des collections iconographiques

Nous considérons les sites de partage d'images au sein du Web à partir du modèle des collections de photographies, pour autant qu'une collection soit comprise « comme une réunion artificielle de documents en fonction de critères communs liés à leur contenu ou à leur support et dont la juxtaposition est le fruit de la volonté ou du hasard ⁴¹ ». Nous nous efforçons dans un premier temps de revenir sur les systèmes d'indexation d'images traditionnels, employés par les professionnels de la documentation, tout en nous interrogeant sur leur opérabilité dans le cas d'une application sur les corpus des sites Internet.

1. Une indexation sémantique surajoutée

Dans une perspective de rationalisation d'une collection, d'une base de données iconographique, la question essentielle est de savoir ce qui est indexable dans une photographie. De prime abord nous avons seulement les mots pour établir des passerelles entre les images et les utilisateurs de cette même banque d'images. L'image reste muette sans les mots qui non seulement lui donnent un sens mais permettent de la retrouver. Ainsi: « Nous parlons dans un monde, nous voyons dans l'autre. » mais nous sommes contraints de les relier car « ce qui fait d'un fonds une mémoire repose sur son indexation. ⁴² ». Bruno Bachimont le confirme : « Les images se caractérisent principalement par un aspect morphologique analogue à la réalité mondaine... Le signe qu'est l'image est un signe qui montre mais non un signe qui dit. Or, à l'instar de la réalité qu'elle prétend représenter,

⁴¹ Association des archivistes français, « Abrégé d'archivistique : principes et pratiques du métier d'archiviste », Association des archivistes français, Paris, 2004, p. 259-263.

⁴² Debray, Régis. *Vie et mort de l'image*, Paris, Gallimard, 1994. 526 p.

l'image ne peut dire par elle-même ce qu'elle signifie et doit reposer sur une paraphrase langagière pour gagner l'intelligibilité qui lui manque. C'est pourquoi tout contenu visuel doit comporter un commentaire ou une légende.⁴³».

Il convient de distinguer comme l'affirme l'auteur les « signes qui disent », signes renvoyant à une double articulation et qui forment un système fonctionnel, des « signes qui montrent », ceux qui ne font pas système. Le langage opère une transformation symbolique du réel, par l'intermédiaire du signe linguistique. À cette codification répond le signe qui montre qui tend à s'effacer devant ce qu'il montre dans la mesure où, ressemblant à son objet, il lui est homogène. C'est en effet l'ambiguïté inhérente de l'image qui se confond avec la réalité qu'elle montre, alors qu'une représentation n'est pas violente, mais seulement la réalité qu'elle montre, l'image, elle, pourrait l'être. Relevant davantage de la réalité perceptive que de la codification symbolique, les images et les sons possèdent une texture sémiotique différente et irréductible.

La prévalence d'un système d'indexation sémantique n'est pas étrangère au caractère indiciel définissant depuis plus de vingt ans aussi bien la photographie que le cinéma. Puisque les images selon cette conception renvoient de façon ontologique au réel, puisqu'elles ne sont souvent considérées que d'après la notion de signe alors leur catégorisation, leur classification au sein des systèmes bibliothéconomiques comme au sein des bases de données, s'opère d'après cet ordre sémiotique qui les place sous le joug du mot. En prônant que « toute photographie est le résultat d'une empreinte physique qui a été transférée sur une surface sensible par les réflexions de la lumière » et que « la photographie est donc le type d'icône ou de représentation visuelle qui a avec son objet une relation indicielle⁴⁴ », Rosalind Krauss est l'une des premières à théoriser cet attachement ontologique de l'image photographique à son référent.

Seulement, là où les moteurs de recherche sur le Web fonctionnent selon une indexation des mots entourant l'image, le principe d'indexation au sein de collections iconographiques se révèle différent de par la nature même des documents en jeu. En effet le moteur Google comprend les images comme un élément constitutif d'un plus vaste document qu'est la page Web. L'image est retrouvée par les balises html qui la signalent au sein du code de la page et par une analyse du texte qui l'accompagne. Au sein d'une banque d'image la photographie est le document même, celui devant être indexé, si dans le

⁴³ Bruno Bachimont. « Bibliothèques numériques audiovisuelles : des enjeux scientifiques et techniques. », *Document numérique*, 1998, vol. 2, n° 3-4, p. 219-242

⁴⁴ Rosalind Krauss, « Notes sur l'index », *L'Originalité de l'avant-garde et autres mythes modernistes*, Paris, Macula, 1993, p. 69.

flux de documents multimédias qui composent le Web une image est rarement isolée, dans une base iconographique la photographie peut se retrouver orpheline. L'utilisation des moteurs d'indexation textuelle est par conséquent inutilisable, l'indexation sémantique doit alors s'opérer à un autre niveau. Si « l'inscription écrite est donc sa propre indexation ⁴⁵ », il en va autrement pour les images pour lesquelles l'indexation n'est pas donnée avec le contenu et doit par conséquent être surajoutée.

2. Indexation, classification et thésaurus

Avec la masse croissante de documents circulant sur le Web, les domaines de la gestion documentaire, de l'organisation et de la classification des données au sein de bases iconographiques suscitent de nouvelles préoccupations, il ne s'agit pas toutefois de problématiques récentes. En l'absence de tout système permettant une automatisation efficace de l'indexation des documents visuels, les annotations d'images sont effectuées manuellement, ce qui permet d'obtenir le niveau d'abstraction désiré. Ces méthodes remontent aux années 1970 et font appel à des documentalistes experts en annotation manuelle d'images, qui vont traiter ces fonds selon des pratiques documentaires traditionnelles, leur appliquant les méthodes bibliothéconomiques développées pour le texte mais adaptées à l'image.

En pratique, et pour de nombreux domaines, les documentalistes ont l'habitude de produire des descriptions textuelles des images en s'appuyant sur des ressources terminologiques : schémas de classifications, listes d'autorité, dictionnaires de noms propres, thésaurus généraux ou spécialisés, etc. Le but de ces ressources est de contrôler les descriptions en atténuant leur variabilité et de remédier à la subjectivité du choix d'un terme descripteur quand la tâche incombe à une seule personne.

Indexer consiste donc à analyser le contenu d'un document et de le transcrire dans un langage documentaire. La double finalité du processus d'indexation est de normaliser la codification du contenu des documents et pour l'utilisateur de faciliter la recherche du document. Un langage documentaire est constitué par des termes choisis pouvant appartenir au langage naturel et par un système de relations entre ces termes, il se caractérise également par une syntaxe en reprenant les règles de construction du langage. Il existe plusieurs types de langage documentaire,

⁴⁵ Bruno Bachimont, « Image et audiovisuel : la documentation entre technique et interprétation. Critique et perspectives », *Documentaliste-Sciences de l'information*, Volume 42, juin 2005, p. 348-353.

- les classifications apparues à la charnière du 19^e et du 20^e siècle comme la Classification Décimale de Dewey, et qui ont essentiellement pour fonction de permettre le rangement physique des documents,
- les langages analytiques dont l'apparition est concomitante avec la montée en puissance des outils informatiques et leur capacité d'analyse.

Parmi ces langages analytiques on peut différencier :

- le thésaurus qui liste les descripteurs comportant des relations d'exclusion (élimination de la synonymie) et précise leurs relations sémantiques et hiérarchiques, il est généralement spécialisé.
- la liste d'autorités qui recense les termes décrivant un sujet de façon normalisée, appelés vedettes-matières, retenus pour l'indexation et ceux exclus reliés entre eux par des relations sémantiques élaborées et développées au fur et à mesure des besoins des bibliothèques qui l'utilisent. Elle est généralement encyclopédique, le système RAMEAU constitue l'exemple-type de la liste d'autorité.

En ce qui concerne les collections iconographiques des adaptations de ces systèmes documentaires sont créées. Comme le souligne Cécile Kattnig: « Les classifications ou les thesauri existent toujours et s'accompagnent souvent de listes complémentaires qui vont porter sur la forme, le genre ou la connotation qui recouvrent le niveau de lecture de l'image.», de même pour les listes d'autorités des listes complémentaires les accompagnent ainsi existe-t-il « Une autre liste d'autorité « Genre de l'Image » utilisée aux archives photographiques du Ministère de la Culture. Elle comprend 7 catégories, avec parfois un niveau hiérarchique : « photographie de nuit », « photographie d'évènementiel » (avec un niveau qui est interview), « portrait » (avec un niveau qui est portrait collectif), etc. ⁴⁶». Par l'intermédiaire de ces index la recherche aboutit à la présentation des notices descriptives correspondantes.

Le principe de l'indexation consiste donc à extraire des mots-clés en langage naturel et de les transposer dans le langage documentaire choisi. Les termes alors à la disposition des usagers sont issus d'un vocabulaire contrôlé, et donc restreint, dont le but est l'élimination des synonymes et le traitement de la polysémie afin de permettre une recherche des plus pertinentes. Toutefois l'indexation par vocabulaire contrôlé demeure problématique en ce qui concerne les documents iconographiques, comme l'explique

⁴⁶ Cécile Kattnig, *En quoi les nouvelles technologies renouvellent-elles le traitement intellectuel de l'image ?*, Actes de colloques tenu les 23 et 24 octobre 2003 à Lyon dans le cadre des rencontres nationales cdidoc-fr, 2003, (en ligne <http://www2.acrennes.fr/testscdi/rencontrelyon/kattnig/kattnig.pdf>)

Marie Didier « À la différence du texte, ou du moins bien plus que lui, l'image pose un problème de pertinence de l'indexation matière, car elle ne saurait être circonscrite, c'est-à-dire décrite de façon exhaustive et objective. Au-delà de sa description, une image est susceptible de renvoyer à d'autres images, d'évoquer d'autres lectures et interprétations, de provoquer d'autres émotions ⁴⁷ ».

D'un point de vue technique l'indexation peut porter soit sur du texte intégral, et se fait alors automatiquement par l'intermédiaire de moteur d'indexation plein texte, soit sur des termes ou expressions contrôlés qui constituent autant de points d'accès aux collections, mais suppose leur définition manuelle au sein des notices. Si des systèmes de traitement automatique de textes permettent l'extraction des termes les plus fréquents d'un corpus et dans une certaine mesure facilitent l'émergence de leurs relations sémantiques, ces outils sont inopérants dans le cas d'une indexation matière, et d'autre part quand le document à indexer est iconique. Une opération supplémentaire par rapport au texte est nécessaire et soumet bien plus l'indexation à la subjectivité de l'indexeur : l'interprétation du contenu visuel d'une image et sa transcription en mot-clé, puis vedettes-matières.

Ainsi la construction du thésaurus et de l'index en ce qui concerne l'image n'est possible que manuellement, l'automatisation des systèmes documentaires interviendra seulement dans le contexte de la recherche, par l'intégration et le traitement informatique des bases de données d'indexation et des notices descriptives.

3. Indexation et métadonnées

L'apparition du numérique, par la numérisation des documents et la constitution de bases aux sources nativement digitales, engendre un renouveau des pratiques d'indexation. On assiste au glissement d'une indexation par notices extérieurs au document à un renseignement compris dans le document par l'intermédiaire des métadonnées. Dans la plupart des systèmes informatisés, les informations relatives à un document et à son classement sont stockées dans une base de données spécifique. Cette séparation entre les documents et leurs métadonnées (au sens d'information sur les données) induit une opération de va et vient pendant le processus de consultation des documents, les métadonnées demeurant dans la base et devant être exporter séparément puis associer à

⁴⁷ Marie Didier, « Indexation, structuration et encodage des fonds iconographiques : le fonds Léon Lefebvre de la Bibliothèque Municipale de Lille », mémoire d'étude, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques, Paris, 2005, (en ligne : <http://www.enssib.fr/bibliotheque-numerique/document-809>)

nouveau avec la ressource. De plus un risque existe alors de perdre les informations associées aux images en cas de transfert ou de conversion de fichiers.

Dans le cas d'un document numérique consultable sur une base de données ou sur un réseau en ligne, les notices descriptives peuvent également être intégrées directement au document numérique en tant que métadonnées informatiques. Avec ces métadonnées, le travail des moteurs d'indexation et de recherche est grandement facilité car ils peuvent extraire automatiquement de nombreuses informations sur le document, ainsi que reconstituer les relations qui existent entre eux.

La question du classement, de la préservation et de l'archivage des documents du Web n'appartient pas aux seuls moteurs de recherche. Le monde des bibliothèques et de la documentation a mis au point des méthodologies, des normes permettant d'appliquer à ces contenus un ensemble de principes de classement de nature bibliothéconomiques au travers de la mise au point de métadonnées métiers, internes aux documents et permettant de renseigner les informations indispensables à la connaissance et compréhension du document. Pour que ces champs informationnels soient efficaces, il est néanmoins nécessaire que tous les acteurs (éditeurs, développeurs de logiciels clients, etc.) partagent une définition et une nomenclature communes des métadonnées, qu'il en existe une normalisation. Partant de ce constat, un certain nombre d'équipes impliquées dans la sémantique sur le Web organisèrent un atelier à Dublin (Irlande) en 1995. Ils définirent une liste précise de quelques métadonnées liées aux pages Web. Ce jeu de métadonnées fut surnommé « Dublin Core metadata ». Le Dublin Core définit une quinzaine de champs de métadonnées utilisables pour les pages Web et les documents numériques de façon générale. Ceux-ci concernent :

- Le contenu (métadonnées de contenu) : Title, Description, Subject, Source, Coverage, Type, Relation ;
- la Propriété intellectuelle (métadonnées de gestion) : Creator, Contributor, Publisher, Rights ;
- la Version (métadonnées de structure) : Date, Format, Identifier, Language.

Les éléments Dublin Core constituent ainsi des métadonnées internes, qui peuvent notamment être encodées dans des balises <meta> d'une page HTML, ils constituent ainsi des éléments pouvant être retrouvés par les moteurs de recherche grand public.

Pour ce qui est des images on peut considérer que les aspects techniques concernant les fichiers numériques sont insuffisamment transcritibles en Dublin Core. Cette limitation d'application a engendré l'apparition de formats de métadonnées spécifiques à la

photographie dont le plus connu est le modèle IPTC (International Press and Telecommunications Council), développé par les organes de presse, il est constitué de 33 métadonnées de type interne, c'est-à-dire stockées à l'intérieur des fichiers images JPEG. Appliqués à la sphère du Web ces systèmes d'indexation des documents en ligne nécessitent un savoir faire, une maîtrise des outils informatiques, ils demeurent par conséquent peu connus du grand public confiné à une pratique documentaire professionnelle, et comme le remarquent Olivier Ertzscheid et Gabriel Gallezot ils « ne s'appliquent, pour une très grande majorité, qu'à des contenus institutionnels ou scientifiques validés et ne concernent donc pas l'ensemble des documents constituant le Web ⁴⁸».

Si les méthodes d'indexation professionnelles sont valables pour des contenus normés, elles s'avèrent inopérantes dans le contexte qui nous intéresse. En effet comment les appliquer à des collections d'images formées à partir de la mise en ligne de corpus photographiques d'utilisateurs non spécialistes, corpus eux-mêmes hétérogènes issus d'une culture non officielle pouvant motiver des classifications vernaculaires. De plus face au nombre massif et croissant de documents en général et de photographies en particulier circulant sur la toile, des méthodes d'indexation s'appuyant sur une saisie manuelle des descripteurs est inenvisageable.

Une première solution apparaît par le biais de l'usage des métadonnées publiques. Il s'agit alors de mettre les opérations d'indexation entre les mains des utilisateurs des sites de partage de signets, d'images ; d'usage simplifié ces modèles de métadonnées ne renvoient pas, à l'inverse de celles du Dublin Core, à un mode d'organisation bibliothéconomique. Elles permettent ainsi à chacun d'insérer dans le code HTML des documents publiés un certain nombre d'informations annexes (titre, description, mots-clés). Comme l'explique toujours Olivier Ertzscheid et Gabriel Gallezot cette première tentative se solda pourtant par un échec ils invoquent ainsi deux raisons principales, comme pour toute utilisation de champs de métadonnées « les pré-requis techniques étaient bien trop lourds pour des utilisateurs n'ayant aucune connaissance du balisage HTML », de plus, des professionnels du référencement s'emparèrent de ces balises et les utilisèrent à des fins stratégiques d'amélioration du positionnement de leur site au sein des moteurs de recherche. Ces techniques aboutirent pour certains à l'utilisation de faux mots-clés pour

⁴⁸ Olivier Ertzscheid et Gabriel Gallezot, « Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », *Document numérique et société*, ADBS Éditions, 2006, en ligne : http://affordance.typepad.com/mon_Weblog/files/sdndocsoc.pdf.

indexer le site faussant du même coup les techniques de classement statistiques des moteurs qui en contre partie ne prirent plus en compte ces métadonnées ou de façon bien moindre.

B. La folksonomie

Les problèmes de ces systèmes de classification et d'indexation métiers sont la complexité des opérations de traduction du document en entité indexée normée. Cette opération ne peut être réalisée que par des professionnels de la documentation, maîtrisant les langages documentaires et leur syntaxe.

Si les sites de partage d'images ont une structure proche du modèle de la collection de photographies en ligne, leurs systèmes d'indexation ne sauraient correspondre à ceux envisagés jusqu'à présent. Celui-ci, selon des principes de simplicité, de non spécialisation développé par le Web 2.0 et surtout d'efficacité à gérer d'importants corpus, doit pouvoir être géré en partie du moins par les usagers, il doit s'inscrire dans un horizon social ou en tout cas ne pas freiner circulation et échanges. L'objectif est de proposer un modèle d'indexation des informations non plus fondé sur des langages documentaires au vocabulaire contrôlé et produits par des spécialistes, mais sur « des modes de catégorisation pragmatiques fondés sur des métadonnées indexées par les utilisateurs ⁴⁹ ».

C'est dans ce contexte que sont apparus les systèmes de tagging collaboratifs ou folksonomies, proposés par les sites de partage d'informations sur le Web, ces outils d'aide à la navigation apparaissent aujourd'hui comme un mode de catégorisation de l'information typique des sites du Web 2.0.

1. L'Apparition de la folksonomie sur le Web

Les folksonomies, qui désignent « un processus de classification collaborative par des mots-clés librement choisis, ou le résultat de cette classification ⁵⁰ », émergent dans un contexte où la structuration des données du Web et leur accessibilité à la fois par un public et des moteurs de recherche représentent des enjeux centraux de son développement. Cette méthode d'indexation est caractéristique du courant du Web 2.0 qui s'efforce de placer l'utilisateur au centre de la production des contenus (usergenerated- content) mais aussi de la production des métadonnées relatives à ces contenus. La structuration des données à

⁴⁹ Maxime Crepel, « Les folksonomies comme support émergent de navigation sociale et de structuration de l'information sur le Web », *Réseaux*, n° 152, juin 2008, p. 169-204, (en ligne : <http://www.cairn.info/revue-reseaux-2008-6-p-169.htm>).

⁵⁰ Définition extraite du site Wikipedia (<http://www.wikipedia.org>).

partir des folksonomies s'opère par une mise en partage des différentes classifications personnelles produites par les utilisateurs dans des logiques d'optimisation de recherche et de partage d'informations. Cette approche est liée à l'émergence d'outils techniques mis à la disposition des utilisateurs sur les sites Web 2.0.

Les folksonomies puisent leur origine dans l'élaboration de techniques efficaces sur Internet et concernant à la fois des problématiques de recherche et de partage de documents. Selon Olivier Ertzscheid et Gabriel Gallezot (2006, p.5) l'émergence des folksonomies sur la scène des systèmes d'indexation est à relier à deux tendances de fond à l'origine d'une renégociation documentaire sans précédent : la mise en ligne de documents appartenant jusque là à la sphère personnelle et l'explosion documentaire qui en découle nécessairement. On assiste ainsi en premier lieu à un processus de re-documentarisation à l'œuvre sur le Web avec la mise en ligne des données et systèmes de classement personnels. Parce que l'utilisation des outils informatiques se nomadise, le recours au réseau comme interface de centralisation et de partage des données devient indispensable, c'est ainsi qu'y bascule une masse considérable d'informations. Selon les deux auteurs ce processus qui marque un changement d'échelle documentaire engendre un changement de statut du document qui, en passant en ligne, s'autonomise, il en va donc de même pour les listes de signets créées par les internautes. Si ces dernières « se caractérisent d'abord par des fonctionnalités documentaires de nature technique », une fois sur le réseau, elles deviennent « autant d'unités documentaires stables ⁵¹», c'est à dire pouvant être recherchées et accédées sur les sites de partage ou directement par les moteurs de recherche.

C'est en ce sens que la première forme de taggage repérable dans l'histoire du net fut mise au point en 2003 sur le site Delicious⁵², une nouvelle plateforme d'échange de signets ('social bookmarking'). Delicious offre à l'époque un service novateur de partage de documents en ligne, l'utilisateur ouvre un compte sur lequel lui est permis de stocker et classer des marques pages renvoyant à autant d'articles, de pages Web ou tout type de documents en ligne dont il désire garder la trace. Dans l'esprit du Web 2.0, le site n'offre pas seulement des outils de classement personnel des données mais permet aux utilisateurs de partager le contenu de leur compte. Ainsi les informations sont classées et assorties de mots-clés librement choisis par le titulaire du compte, ce sont ces mêmes mots clés qui

⁵¹ Olivier Ertzscheid et Gabriel Gallezot, *op.cit.*, p.6.

⁵² <http://delicious.com/>

permettront de rendre les données visibles et accessibles. Le principe du taggage émerge dans un contexte de “social bookmarking”, où l’acte de tagguer une ressource procède de la même intention que de créer un marque page personnel. Selon ce schéma les tags permettent à l’usager de lister et classer des documents mais aussi de les retrouver en utilisant les mots-clés dédiés.

L’autre tendance soulignée par les auteurs est celle de la communautarisation des informations, « de l’explosion documentaire qui marque chaque évolution du Web » et dont découlent le développement et la diffusion du modèle folksonomique. La popularité constante de ces outils d’indexation s’explique alors par leur interaction avec des applications communautaires, par la liaison créée entre des activités de signalisation individuelle avec la trame dense des réseaux sociaux et du partage de ressources. Avec l’émergence du Web 2.0 on assiste à la création des communautés d’expertise se constituant pour faciliter l’accès à des contenus documentaires bien ciblés alors que sur le Web se multiplient des outils de recherche souvent experts, « bien trop complexes pour la majorité des utilisateurs ». Mais la révolution se révèle surtout dans la formation de communautés d’intérêt dont l’objectif est la production et le partage de liens, communautés qui « permettent à leur tour la multiplication d’unités documentaires en ligne ». Flickr⁵³ est ainsi un exemple de cette formation de communautés d’internautes se regroupant autour d’intérêts communs et mettant en œuvre des systèmes de classement spécifiques, des folksonomies adaptées aux usages. Les tags sont donc choisis, parce que permettant de catégoriser ses propres ressources sans nécessiter de savoir faire techniciste particulier. Comme le souligne les auteurs « Au regard de la difficulté de mise en œuvre de normes idoines sur des contenus et des corpus documentaires pourtant déjà homogènes, au regard ensuite de l’usage perverti des métadonnées grand public, une analyse sommaire pourrait porter à croire qu’un tel système non contraint et appliqué à des corpus documentaires totalement hétérogènes (dans leurs formats comme dans leurs finalités) s’avérerait rapidement être un échec. ⁵⁴ ». Aujourd’hui et contre toute attente les folksonomies demeurent le système d’indexation le plus évident pour les sites de partage de documents à visée sociale sur Internet.

⁵³ <http://www.flickr.com/>

⁵⁴ Olivier Ertzscheid et Gabriel Gallezot, *loc.cit.*

2. Les principes de la folksonomie

Face à la profusion des contenus multimédias sur le Web, les systèmes classificatoires traditionnels basés sur des systèmes hiérarchisés d'ontologie et autres vocabulaires normés produits par des professionnels s'avèrent inopérants, incapables de gérer le flux constant de documents à indexer sur Internet. Il est désormais nécessaire d'envisager l'indexation des documents du Web sous l'angle d'une pratique pouvant être ouverte au non spécialiste. C'est ainsi que les principaux modèles d'indexation des informations sont fondés sur des modes de catégorisation pragmatiques basés sur des métadonnées indexées par les utilisateurs eux-mêmes, seule voie possible à une indexation efficace des masses documentaires circulant sur la toile.

Les folksonomies, au contraire des systèmes classificatoires classiques, ne reposent sur aucun thésaurus, liste d'autorité, sur aucun vocabulaire contrôlé mais est basé sur le langage naturel. Les folksonomies sont entièrement centrées et tributaires de l'intervention de l'utilisateur et leur confère une totale liberté pour produire leur propre classification des informations en y associant des mots-clés, appelés « tags » qui vont permettre de décrire et de catégoriser les documents. Le mot tag signifiant en anglais étiquette de balisage, étiquetage, fléchage, marquage, voire traçage, il est possible de rencontrer des traductions comme taggage collaboratif. Voici la définition qu'en donnent Guy Marieke et Emma Tonkin:

« Que sont vraiment les tags ? Une définition simple serait de dire que les tags sont des mots-clés, des catégories de noms, ou des métadonnées. Essentiellement, un tag est simplement un jeu de mots-clés librement choisi. Cependant, du fait que les tags ne sont pas créés par des spécialistes de l'information, ils ne suivent aucune indication formelle. Cela signifie que ces items peuvent être catégorisés avec n'importe quel mot définissant une relation entre la ressource en ligne et un concept issu de l'esprit de l'utilisateur. Un nombre infini de mots peut être choisi, dont quelques-uns sont issus de représentations évidentes tandis que d'autres ont peu de signification en dehors du contexte de l'auteur du tag.⁵⁵ ».

L'utilisateur utilise les tags dans un but personnel, afin d'organiser son propre système d'information et de produire des classifications personnalisées de ses collections personnelles. Il s'organise en fonction de ses besoins propres d'information et sans être à

⁵⁵ Guy Marieke, Emma Tonkin , « Folk-sonomies. Tidying up Tags? », *Dlib Magazine* , N° 1, Volume 12, 2004, (en ligne : <http://www.dlib.org/dlib/january06/guy/01guy.html>), traduit par nos soins.

aucun moment soumis à une organisation restrictive établie par des professionnels. Tout peut se trouver indexé par l'utilisateur, depuis ses favoris jusqu'à ses photos, en passant par ses messages sur son blog.

Selon un esprit collaboratif qui définit les sites de partage de signets ou de documents visuels, les tags permettent également d'indexer les documents produits par les autres toujours selon ses besoins. À partir de cette activité individuelle de tagging, les sites vont agréger l'ensemble de ces métadonnées pour proposer à tous les utilisateurs un système de classification des ressources du site, ainsi l'utilisateur choisit d'attribuer un ou plusieurs tags à un document et en contre partie le site Web regroupe les mots-clés semblables en une liste ou « nuage de tags » duquel vont ressortir les thèmes les plus populaires. Les tags offrent ainsi la possibilité de rendre explicite les univers sémantiques relatifs à l'information et renouvellent les formes d'accès et de mise en valeur des ressources Web.

Les folksonomies loin d'être des systèmes fixes doivent être considérées comme des classifications dynamiques, parce que d'une part elles évoluent à travers le temps et d'autre part parce qu'elles sont le fruit d'une agrégation des classifications personnelles d'utilisateurs, dont les pratiques évoluent sans cesse selon leurs intérêts, leurs compétences et le développement de leur réseau de connaissances. Les folksonomies s'appuyant sur des réseaux sociaux ne font pas que classer de l'information et la partager, elles mettent en relation des usagers qui partagent les mêmes centres d'intérêts, se nourrissant de la diversité des représentations du monde et des catégories auxquelles les objets renvoient ; elles permettent de relier entre eux des individus ou des groupes sociaux distincts. Les tags postés par les utilisateurs sont produits de façon individuelle mais renvoient systématiquement à des concepts utilisés par des collectifs (plus ou moins étendus) qui partagent des expériences, des valeurs, des centres d'intérêt communs et pour lesquels s'établit un accord sur la manière de désigner les choses. Sans nécessairement prendre conscience de la dimension collective de cette activité, certains utilisateurs produisent une classification des ressources qui leur est propre à partir de catégories socialement marquées qui font sens pour d'autres. Ainsi en fonction de la diversité des groupes les sites vont offrir une diversité de tags pour définir un même objet.

Toutefois cette même diversité, qui pourrait vite s'avérer créer un frein au partage en diffusant de trop nombreux items isolés, va se trouver régulée par les pratiques des usagers eux-mêmes. Cette activité de tagging, à la différence d'une activité de catégorisation ou d'indexation dans un contexte documentaire professionnel, ne nécessite

qu'un très faible coût cognitif, c'est ainsi que Rashmi Sinha tend à expliquer l'engouement que suscite cette pratique : « le marquage ("tagging") élimine la phase de décision (choisir la bonne catégorie) et dissipe la phase de paralysie d'analyse ("the analysis-paralysis stage") pour la plupart des gens. (...) Il offre un retour social et sur soi-même immédiat. Chaque "tag" traduit un peu de vos centres d'intérêts et les ancre dans un contexte social immédiat. (...) La beauté du taggage est qu'il est inscrit dans un processus cognitif déjà existant sans lui ajouter de coût cognitif supplémentaire.⁵⁶ ».

3. Une étude de cas : Flickr

Si Flickr constitue désormais un exemple de site Web 2.0, il ne fut pas créé à l'origine dans la perspective d'en faire un outil de publication de photographies. Lancé en 2004 par Stewart Butterfield et Caterina Fake⁵⁷ le site était initialement destiné à devenir un jeu en réseau, une plateforme où les joueurs s'échangeaient des objets matérialisés sous la forme de photographies. Ce sont finalement les téléchargements de photographies privées qui prirent le dessus et devinrent ainsi la fonction principale du site. Racheté par Yahoo en 2005, des comptes d'utilisateurs sont alors créés permettant à l'utilisateur de gérer et classer ses images, les photographies peuvent alors être dotées de légendes, titres et descriptions à partir de l'interface *Organizr*. Le site permet aux usagers de télécharger leurs photos personnelles et de les héberger en ligne, mais à la différence des autres sites de stockage d'images existant avant lui, les photos mises sur Flickr sont par défaut publiques et visibles par les autres membres. Il existe plusieurs manières complémentaires de classer ses productions sur Flickr. Par défaut, le site classe les photos de façon chronologique suivant leur date de publication sur le compte de l'utilisateur. L'utilisateur peut également choisir de les classer dans des albums thématiques à l'intérieur desquels une photo peut être indexée, reliée à différents albums simultanément.

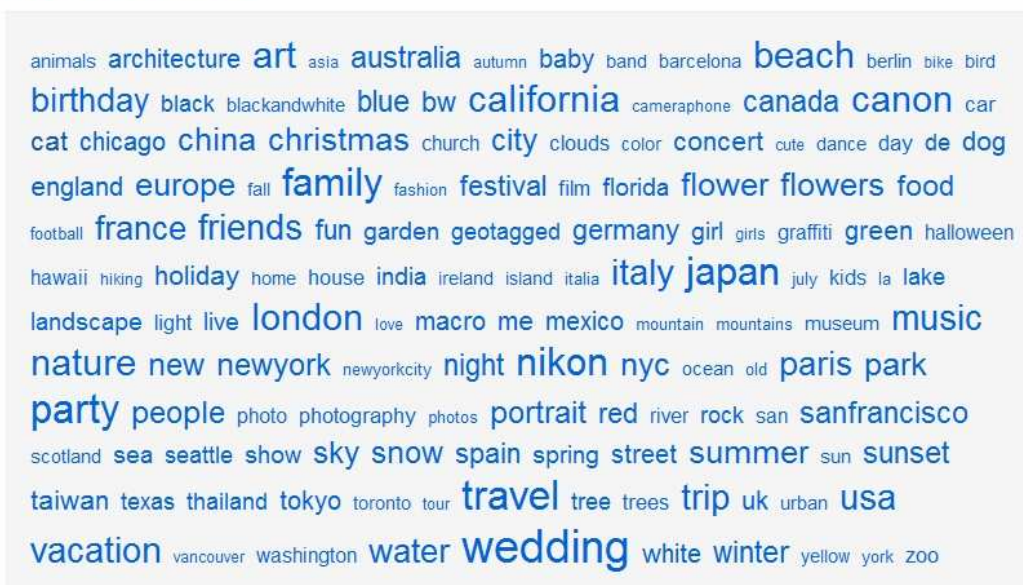
Flickr est aujourd'hui l'un des plus fameux sites de partage de photographies en ligne, dont les tags constituent les éléments essentiels au système d'échange et de communautarisation des images. S'ils représentent le troisième mode de classement des contenus, les tags permettent également de rendre ses images visibles et accessibles aux autres usagers alors à même de filtrer les innombrables items du site en effectuant une requête par mots-clés ou par l'interrogation du nuage de tags. Le succès de Flickr est en

⁵⁶ Rashmi Sinha, « À cognitive analysis of tagging. », *Rashmi's blog*, 27 Septembre 2005, (En ligne <http://rashmisinha.com/2005/09/27/a-cognitive-analysis-of-tagging/>), traduit par nos soins.

⁵⁷ Cf un résumé de l'histoire de la création de Flickr sur Wikipedia <http://en.wikipedia.org/wiki/Flickr>

partie dû aux importantes interactions sociales qu'il permet, en plus de télécharger des photographies les membres peuvent se créer des réseaux de contacts, rejoindre des groupes selon leurs centres d'intérêts, envoyer des messages, commenter, tagguer les images des autres, ou encore marquer leurs favorites. Cette abondance de moyens de communication place le site du côté des réseaux sociaux au sein duquel les tags jouent un rôle déterminant.

Tags les plus populaires



III. 5 Capture d'écran du nuage de tags de Flickr

Devant le flux massif d'images téléchargées chaque minute, le taggage des images apparaît comme une solution efficace pour prévenir de leur disparition, de l'incapacité dans laquelle nous serions de les retrouver au sein d'une collection sans cesse croissante. Le site Flickr compte ainsi plus de quatre millions de tags différents postés par les utilisateurs et s'il est possible d'utiliser jusqu'à 75 tags pour marquer sa photographie, il apparaît selon l'article de maxime Crepel que « 60 % des photos postées sur le site possèdent au moins un tag, pour une moyenne de 4 tags par photo. Sur l'ensemble des tags indexés par les utilisateurs 45 % des tags ne sont utilisés qu'une fois⁵⁸ ».

Le fait d'attribuer plusieurs tags à sa photographie permet non seulement de circonscrire au plus près le sujet de l'image mais aussi de relier son image à autant de catégories et par conséquent de groupes d'utilisateurs utilisant ces mêmes mots-clés. Sur la page dédiée à la recherche par tags, Flickr met systématiquement en avant les tags les plus populaires ou ceux, parmi les plus récents, qui enregistrent une forte activité, ce sont les

⁵⁸ Maxime Crepel, *op.cit.*, p.179.

« hots tags », des termes alors favoris et qui vont permettre de faire ressortir des groupements thématiques, des sujets communs ou d'actualités à partir de la masse d'images disponibles sur le site. Flickr offre également la possibilité de s'abonner au flux RSS d'un tag particulier.

Par les tags, les utilisateurs créent de la valeur pour les autres, et ce de façon le plus souvent inconsciente. Comme le dit Tim O'Reilly : « Le "tagging" se prête à des associations multiples et redondantes, plus semblables à celles que notre cerveau réalise que des catégories rigides. Exemple typique : la photo d'un chiot dans Flickr peut être taguée par les termes "chiot" et "mignon" - permettant donc de retrouver la photo suivant un cheminement naturel grâce, une fois de plus, à l'activité des utilisateurs. [...] L'attention n'est réciproque que de manière accidentelle⁵⁹ ». Toutefois comme le remarque Maxime Crepel toujours, « le système de tags relève le plus souvent des classifications personnelles qui mobilisent des catégories qui n'ont de sens que pour un utilisateur ou des collectifs très restreints d'utilisateurs qui partagent un ensemble de représentations et de catégories communes ».

Dans un contexte d'utilisation individuelle des tags, ceux-ci peuvent également jouer un rôle au niveau de la popularité et de la visibilité des images. Les images apparaissant les premières sur la page de résultat d'une recherche sont souvent celles possédant le plus de tags, tags grâce auxquels elles ont récolté de nombreux commentaires et vues. Les utilisateurs les plus avancés réfléchissent à leurs tags et aux manières dont ils peuvent les aider à sortir du lot ou à se relier à certaines communautés ; s'ils identifient un tag qui leur semble intéressant, ils cherchent à l'employer en prenant des photos qui pourraient entrer dans cette catégorie ou en associant ce mot-clef à leurs anciennes photos. Les tags peuvent ainsi devenir le moyen de se départir de l'anonymat, des outils pour une stratégie de promotion de son compte et de visibilité accrue de ses productions.

II. Vers les nouveaux usages de la folksonomie

La flexibilité de l'usage des tags donne lieu à de nombreuses et intéressantes formes d'interaction, pouvant se révéler, ludiques, créatives ou plus simplement pragmatiques, où le taggage « loin d'être une simple pratique individuelle de

⁵⁹ Tim O'Reilly, « What is Web 2.0 ? », *O'Reilly*, 09 septembre 2005, (en ligne <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-Web-20.html>).

catégorisation, amène les utilisateurs à entrer dans des formes multiples de coopération ⁶⁰». Nous prenons ici Flickr comme site de référence nous permettant de mener une analyse quant aux différents usages qui peuvent être faits de l'indexation par tags en dehors de ces applications basiques que sont notamment le classement pour soi et une aide à la valorisation de ses contenus.

A. Les pratiques sociales et communautaires

Si certains usagers inscrivent leurs pratiques des tags dans des logiques individuelles de classement ou de valorisation des contenus, d'autres obéissent à des usages plus sociaux, parfois conventionnels et pouvant prendre part à une politique spécifique de certains collectifs que sont les groupes thématiques de Flickr. Comme le précise Maxime Crépel, « Ces usages relèvent des stratégies différentes d'évaluation et de classement des contenus, mobilisant des catégories sémantiques hétérogènes qui permettent par anticipation de produire des marqueurs de l'information pour soi ou pour les autres membres du site ⁶¹ ».

1. Des tags aux groupes

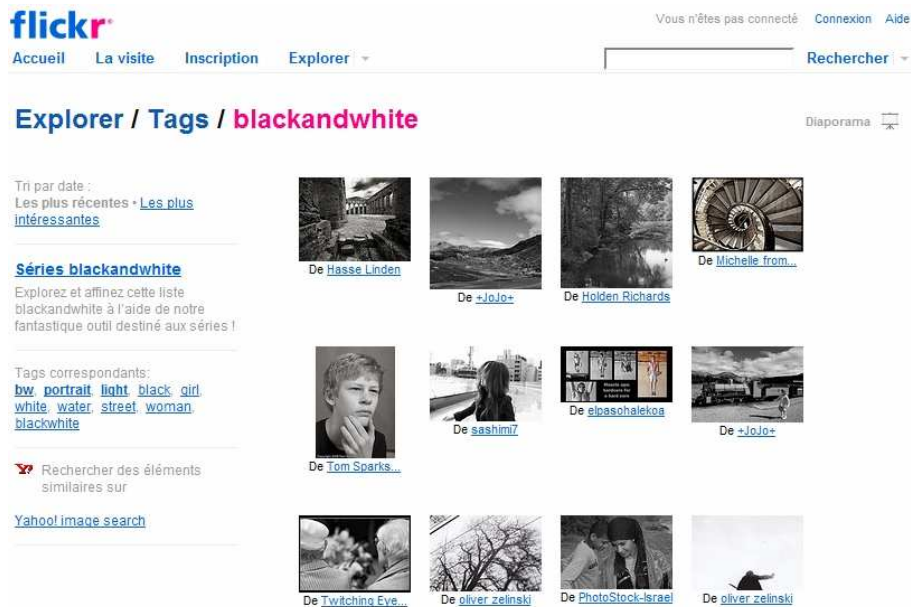
Nous l'avons précédemment signalé, Flickr permet aux utilisateurs de se regrouper en communautés de toutes sortes, les fameux groupes Flickr, dont le but est de rassembler les utilisateurs selon des points communs partagés. Le plus souvent se forment ainsi des communautés de goût, ou d'intérêt qui peuvent aller des marques de l'appareil photo utilisé pour prendre ses images, à des regroupements par sujet, par style ou esthétique. Ainsi existe-t-il des groupes centrés sur les photographies de coucher de soleil, sur le « street art » ou encore sur un lieu géographique déterminé.

Les tags, s'ils peuvent déterminer le sujet d'un groupe, permettent également le rattachement de son image à une communauté, les administrateurs de groupes faisant régulièrement des recherches par mots-clés pour trouver de nouvelles photos pouvant venir compléter le pool de la communauté. Tagger ses images c'est donc, par elles, avoir la possibilité d'entrer en relation avec un groupe de membres partageant un intérêt ou un regard semblable. L'importance de bien tagger ses images, quand on est dans une optique de rencontre et d'échange social, est primordial, on ne rentre pas en contact avec les mêmes communautés si l'on marque ses images du tag « B&W » autour duquel se

⁶⁰ Maxime Crépel, *op.cit.*, p.171.

⁶¹ *Ibid.*, p.188.

concentre une communauté d'amateurs éclairés de la pratique du noir et blanc et qui propose le plus souvent des images travaillées et esthétiques, ou si l'on utilise un terme plus générique, comme « fleur » ou « coucher de soleil », ne se référant à aucune communauté particulière sinon à des dizaines.



III. 6 Capture d'écran du Pool du groupe Flickr BlackandWhite

Un tag peut ainsi se référer à une communauté et peu de tags en lien avec un sujet, un thème précis n'ont pas de groupes attachés. Si les groupes peuvent constituer un espace supplémentaire de visibilité de ses images, il est aussi et essentiellement un espace de sociabilité où les membres se retrouvent, discutent et échangent. Des pratiques collectives naissent ainsi sur ces groupes, pratiques que Caterina Fake, co-fondatrice de Flickr, désigne sous le terme de "culture de la générosité"⁶² et désignant les attentions et égards réciproques que les gens s'échangent via ces groupes. Il est aussi à noter que si les échanges communicationnels entre les membres d'un groupe passent bien sûr par des commentaires laissés sous la photographie, l'image elle-même peut devenir vectrice de communication. Ainsi à une image postée sur la page du groupe répondent d'autres images, un discours implicite sous-tendant alors les publications de photos sur les murs des groupes et la base des échanges discursifs de ces groupes passent par l'image.

⁶² Cf l'entretien avec Catherine Fake sur <http://ouatedepoque.blogspot.com/archive/2005/12/16/entrevue-avec-caterina-fake.html>

2. Commenter par les tags

Ce caractère discursif de l'image se retrouve largement sur les comptes personnels des utilisateurs et pages de présentation des photos, où nombre d'entre elles sont accompagnées de commentaires et réactions laissés par ses contacts voire par des inconnus. Certains même n'hésitant pas à recourir aux annotations appliquées directement à l'image, il s'agit là d'une application spécifique à Flickr, permettant de former un encadré sur l'image et d'y adjoindre un mot ou une remarque en liaison avec l'élément pointé. Cette abondance de moyens de communication entourant l'image et surtout le nombre croissant de personnes y ayant recours, témoigne du nouvel engouement que suscite l'exploration des collections des sites de photographies sur Internet. Les sites de partages d'images semblent avoir peu à peu suscité l'émergence d'un nouvel intérêt du public envers la photographie même amateur, intérêt transformant l'épreuve du visionnage des images d'autres usagers en activité sociale voire en véritable loisir. Regarder les photographies de famille, de vacances ou les œuvres plus artistiques d'un parfait étranger se conçoit désormais de façon ludique, pouvant aboutir à des pratiques de légendages et de commentaires voire de remix des images.

Au sein de cette conception de l'annotation des images comme pratique sociale, le tag est lui aussi amené à jouer un rôle et apparaît alors comme un moyen d'expression et un support d'inventivité. Selon la logique consistant à commenter et communiquer autour d'un contenu visuel et à partir de la possibilité offerte aux usagers du site de marquer les photos d'autrui de ses propres mots-clés, certains utilisateurs vont détourner et se servir de la forme du tag pour laisser des commentaires aux images. Les contenus ne sont alors plus indexés par de simples mots-clés mais par des groupes nominaux voire de véritables phrases. La liste de tag d'une image se retrouve ainsi affublée de phrases ayant permis aux auteurs, au-delà d'une fonction descriptive, d'exprimer une émotion, un avis ou une opinion en lien avec le contenu de la photo ou avec le moment, la date, le contexte de prise de vue.



Fofie aux fourneaux...

par **fofie57**

6 commentaires

Comportant le tag [food](#), [cooking](#), [de](#), [la](#) ...

Créé le 24 mars 2007, téléchargé le 24 mars 2007

Lieu : [Seine-St-Denis](#), [Île-de-France](#) (carte)



Voir plus de photos de fofie57 ou consulter son profil.

- la revanche du casse-croûte
- toi aussi joue au petit chimiste
- la
- vengeance
- de
- poupée
- vaudou
-
- jamais_plus_je_pique_la_casserole_de_ma_grand_mere
-
- comment_tu_perds_ta_caution_en_3_minutes_chrono
- mise en situation pour essai de la
- derniere eponge gratante
- bob l'éponge va avoir du boulot
- explosion
- sublimation de lentilles
- Flapa boum pouf
- miam
- hummy
- repas equilibre
- fratjapostruckbleurkjbf
- les lentilles c'est dég!!
- y'a meme pas de saucisses avec...
- lentilles bleues ?
- pffffplus que du jus
- c'est meilleur quand c'est bien cuit
- Fofie fon fon fon...
- heu finalement j'ai piscine le 25 avril, je ne pourrai malheureusement pas me joindre à toi
- J'appelle Monsieur Propre
- Pin...Pon... Pin... Pon...
- La prochaine fois : Salade de Lentilles
- Le kebab c'est bien aussi

Ill. 7 Capture d'écran de la liste de tags de la photo « Fofie aux fourneaux »⁶³

Comme le remarque Maxime Crepel « Une des spécificités du système de tagging par rapport aux systèmes de commentaires est qu'il ne permet pas d'identifier les auteurs des tags ajoutés sur une photo (à l'exception de celui qui a posté la photo et en a autorisé le tagging). C'est donc un espace qui peut servir pour commenter une photo, le plus souvent dans le cadre de jeux amicaux entre contacts.⁶⁴ ». En effet sans mention de l'origine du tag, les moyens d'entrer en relation avec l'auteur sont limités, si cette pratique ne se pose pas en créateur de lien social elle illustre toutefois les formes inventives et ludiques que peuvent prendre les communications amicales sur Flickr. Sur une image soumise à ce détournement des outils de taggage, les premiers tags de la liste conservent un usage classique d'outil d'indexation, par la suite apparaissent des tags-phrases pouvant ne plus avoir aucun rapport thématique avec l'image mais formant entre eux une sorte de dialogue dont le sens n'est alors souvent accessible qu'aux acteurs du jeu le plus souvent membres du réseau d'amis.

⁶³ <http://www.flickr.com/search/?q=fofie%20aux%20fourneaux&w=all>

⁶⁴ Ibid, p.198.

3. Tags et jugement de goût

Un autre détournement de l'usage des tags se constate dans leur utilisation comme indicateur de qualité. Cette pratique est souvent l'apanage de groupes ayant centré leur activité sur l'élaboration d'un mode d'évaluation collectif des contenus à partir du tag comme élément discriminant. Flickr recense différents groupes de ce genre, les plus célèbres étant certainement la communauté anglo-saxonne de *DeleteMe*⁶⁵ et son contraire *SaveMe*, groupes ayant entraîné la formation de leur homologue français *Vire-moi !* Ces groupes fonctionnent généralement selon un principe similaire, pour chaque publication d'une image sur le pool du groupe les membres sont tenus de juger ou commenter une ou plusieurs photos du groupe, certains administrateurs de site n'hésitant pas à effacer les photos des usagers ne respectant pas les règles. Les critiques et appréciations d'une image se font par l'utilisation de tags spécifiques venant marquer l'image. Ainsi sur le groupe *DeleteMe* les membres sont invités à tagguer les photographies avec soit « deleteme » ou « saveme ». Après dix tags « deleteme » ajoutés à une image celle-ci est supprimée du groupe, inversement une photo recevant dix « saveme » est invitée à rejoindre le groupe *TheSAFE* où elle retrouve les autres images sélectionnées durant la semaine, images alors soumises à une nouvelle épreuve de jugement. On notera l'existence de groupes équivalents dans la sphère francophone avec la communauté « Viremoi »⁶⁶ notamment. Les tags employés par ces groupes ne fonctionnent pas selon les principes traditionnels d'indexation, ils ne décrivent pas plus qu'ils ne caractérisent le sujet d'une image, leur fonction n'est pas de permettre à celle-ci d'être classée ou plus aisément repérable mais de lui accoler la trace d'un jugement opérant dans le temps du jeu, comme l'explique Maxime Crepel « Il s'agit donc de catégories « grégaires » n'ayant de pertinence que pour les membres qui participent à ce type de groupe. La coopération entre les membres est plus formelle, car elle repose sur le dispositif de tagging qui sert de support à l'évaluation des contenus par le collectif à partir d'une convention bien établie sur le type et le nombre de tags à mobiliser pour effectuer le filtrage, ainsi que l'application de règles de tagging formellement explicitées par les administrateurs qui en assurent également le contrôle.⁶⁷ ».

⁶⁵ <http://www.flickr.com/groups/deleteme/>

⁶⁶ <http://www.flickr.com/groups/viremoi/>

⁶⁷ Maxime Crepel, *loc.cit.*



cette tête me faisait trop de peine dans mon assiette...

Commentaires



[Stox - Ideas Playground](#) pro dit :

lol , tu nous fais du gore maintenant... C'est un coucours entre Blandine et toi ?

Après les cadavres de mouche , le poisson mutant brûlé ?

Somme toute , je suis sur que ça ne fa pas empêché de le manger...

Posté il y a 26 mois. ([permalien](#))



[fofie57](#) pro dit :



Cette photo fait aussi partie de :

- + Fofie elle-même (Album)
- + Petits plats (Album)
- + extraordinaire (Pool)
- + Duos doux dingues :-o) (Pool)
- + Graphistologie & Bigoode show (Pool)
- + Stop playing with your food! (Pool)

Tags

- fish
- virela
- virela2
- virela3
- virela4
- virela5
- virela6
- virela7
- virela8
- virela9
- vireal10
- virela11 finalement on va jusqu'à 20

III. 8 Capture d'écran d'une photographie ayant été taguée par le groupe « Viremoi »

Pour les membres de ces groupes le but est de permettre une sélection qualitative des images basée sur le vote des membres, ainsi est-il permis de lire sur la page d'accueil du groupe Vire-moi ! : « Donc le jeu, son essence, c'est quoi déjà ? Et bien ne sélectionner, ne garder, que les photos les plus éblouissantes, inouïes, fantastiques, celles qui ôtent les mots de la bouche et qui coupent les jambes. Vous n'avez juste qu'à oser poster une photo dans le groupe, voir comment réagiront les autres membres, et à quelle vitesse la photo se fera jeter du groupe ». Le groupe se veut ainsi un espace de libre expression au sein duquel les tags « virela » ou « gardela » font office de sentences mais apparaissent aussi comme autant d'avis donnés sur la valeur esthétique de l'image. Devenir membre de ces groupes c'est forcément en acceptant les règles et se placer aussi bien dans la position de juge que de jugé, c'est accepter voire rechercher un avis tranché sur ces images. Dans un article⁶⁸ Dominique Cardon et son équipe de chercheurs remarquent au sujet de ces groupes d'appréciation des images que leur fonctionnement redouble le fameux "interestingness", application phare de Flickr. Cet algorithme est censé prendre en compte les différents critères de popularité d'une image et permet ainsi d'extraire et de mettre en avant les plus intéressantes, originales, celles offrant des qualités esthétiques indéniables. Pour les

⁶⁸ Cf. Dominique Cardon et alt., "The Strength of Weak cooperation: A Case Study on Flickr", 2008, en ligne <http://hal.archives-ouvertes.fr/docs/00/25/66/49/PDF/weak-coop-flickr-v1.1.pdf>

auteurs cette redondance de fonction témoigne du désir de la part des usagers du site, ou de certains d'entre eux du moins, de définir eux-mêmes les critères d'une photo réussie, et de soumettre le jugement à une assemblée de personnes participant elles-mêmes à l'enrichissement visuel du site.

B. Les détournements créatifs de l'indexation

.Au-delà d'une simple étape de documentation de ses propres photos, de celles d'autrui et d'une méthode d'annotation collaborative, les tags se trouvent au centre d'usages détournés, souvent ludiques et créatifs parmi lesquels ont surgi une série de Mashups.

Nous considérons l'analyse de certains d'entre eux comme pouvant ouvrir sur des questionnements plus profonds en lien avec les principes fondamentaux du Web 2.0, comme l'utilisation des tags et le fonctionnement de la folksonomie, ou encore sur le passage du caractère personnel des données à celui de document réutilisable.

1. Définition d'un mashup

Si nous nous référons à la définition de Wikipedia, un *mashup* est une application composite «qui combine du contenu provenant de plusieurs applications plus ou moins hétérogènes. On parle de *mashup* artistique ou de *mashup* technologiques (...). On parle de *mashup* dans le cadre d'une superposition de deux images provenant de sources différentes, superposition de données visuelles et sonores différentes par exemple dans le but de créer une expérience nouvelle.⁶⁹». Au niveau du Web les *mashups* reposent sur des interfaces de programmation (API, Application Program Interface) ouvertes qui permettent de créer de nouveaux services "composites" en exploitant et en agencant de manière originale des programmes, des services et des contenus créés par d'autres.

L'une des plus célèbres sociétés au sein de l'univers des *Webs-mashups* est sans conteste Flickr. Ce site de partage de photographies amateurs a ouvert son API au public presque depuis sa création, permettant ainsi de développer des applications basées à la fois sur une recherche par tags, par nom d'utilisateurs, par groupes ou même par leur célèbre système de promotion des images les plus populaires «interestingness». L'API Flickr fut également un centre d'intérêt pour de nombreux développeurs Web l'ayant peu à peu rendue utilisable avec d'autres programmes tels que Flash, PHP, Java, etc. ce qui contribue

⁶⁹ http://fr.wikipedia.org/wiki/Application_composite

à enrichir d'autant plus les possibilités usuelles et esthétiques des *mashups*. Flickr en proposant une base de donnée bien organisée sur laquelle des millions d'utilisateurs mettent chaque jour en ligne de nouvelles images, offre des possibilités infinies et toujours renouvelées de mixage du contenu, apparaissant comme autant de propositions inédites de navigation dans ce flux ininterrompu d'images.

Un nombre toujours croissant de *mashups* s'est ainsi développé autour de Flickr, si beaucoup d'entre eux sont des jeux (à l'image de Flickr sudoku)⁷⁰, basés sur une approche ludique du Web et de ses potentialités, certains offrent des propositions intéressantes, originales permettant un rapprochement théorique avec des formes plus anciennes de jeu, montage, détournement d'images. Toutes ces applications fonctionnent comme autant de services renvoyant au contenu et visant à favoriser l'échange et la découverte. Leur succès est moindre que celui du site lui-même, mais ils contribuent à la visibilité du service en multipliant les chemins d'accès au catalogue.

2. Mashups et jeux de mots

Créé par Scott Michael Reynen, Web designer et programmeur de profession, Fastr⁷¹ est un *mashup* dont le principe consiste à retrouver, à partir d'une image prise sur Flickr, l'un des tags qui lui est associé. Plus précisément, deux systèmes sont disponibles, la recherche du mot-clé correspondant à une seule image, le « Mini Game », ou celui correspondant à une série de dix images pour la version chronométrée et dite officielle. Une pratique du site révèle rapidement la difficulté de trouver le bon tag à partir d'une seule image. L'exercice se trouve en revanche simplifié par l'apparition de plusieurs photographies ayant en commun le mot-clé ; la recherche s'effectue alors par une analyse déductive du point commun liant les images entre elles. Il est tout à fait hasardeux de réussir du premier coup le jeu de la photo unique, celle-ci pouvant être rattachée à de nombreux tags ayant un lien logique avec elle, soit au niveau des formes et objets représentés, de la couleur, de l'atmosphère, etc.

Ce jeu, aussi simple soit-il, met en lumière le fonctionnement même du taggage d'images, de la folksonomie. Si cet outil offre l'avantage de créer un système d'indexation communautaire, souple et infini c'est aussi cette infinité qui en fait sa limite, tout terme

⁷⁰ Cf. la sélection effectuée des dix meilleurs mashups de Flickr sur <http://www.Webmonkey.com/Webmonkey/06/08/index4a.html?tw=commentary>

⁷¹ <http://randomchaos.com/games/fastr/>

pouvant être employé à la caractérisation d'une image, retrouver selon le principe de Fastr un tag spécifique mais contingent est un exercice hasardeux.

fastr — a flickr game

Play the full game

- [Choose a player name](#)
- [Choose a game](#)
 - [Original \(English\) Game](#) (1)
 - [Group Game](#) (0)
 - [Deutsch \(German\) Game](#) (1)
 - [Español \(Spanish\) Game](#) (2)
 - [Français \(French\) Game](#) (1)

Recent fastr news

- [Fastr Bug Reporting](#)
- [Fastr Mini](#)
- [Fastr Groups](#)
- [AdBrite](#)
- [Daily High Score](#)

Mini game



pirate

Correct! [Try the full game](#)

[Add fastr mini to your website](#)

III. 9 Capture d'écran d'une partie du Mini Game Fastr

Les mots clés utilisés et choisis par les internautes se veulent la plupart du temps transparents, cherchant une adéquation parfaite entre le mot et l'image et ressortent de catégories largement usitées par les internautes. Mais comment réduire une image à un mot ? Si l'association tag/image se révèle efficace dans le cadre d'une recherche au sein même du site, le *mashup* Fastr révèle la complexité de cette pratique et surtout la quasi impossibilité d'une correspondance objective et absolue entre une photo et un mot. Les tags relèvent d'un choix avant tout subjectif, d'une interprétation, lecture personnelle de l'image. Ce que je vois dans la photographie et qui me semble essentiel ne sera pas forcément partagé par tous, d'où les listes de tags accompagnant bien souvent la mise en ligne d'une image. À la différence d'une légende qui offre une recontextualisation à l'image, le tag se base sur des critères descriptifs, sur un choix de lecture lui même inhérent au caractère polysémique des images.

Dans la série des mashups ludiques il serait également possible de citer l'application *Spell with Flickr*⁷². Ce jeu permet de créer des phrases où chaque lettre est une photographie du

⁷² <http://metaatem.net/words/>

signe. Pour ce faire il utilise la collection du groupe *One Letter*⁷³ présent sur Flickr, groupe centré sur la photographie de lettre isolée et tagguée par cette même lettre. Au fil du temps s'est constitué un véritable alphabet polymorphe et photographique, sorte d'encyclopédie visuelle des polices et lettrines existant dans l'espace public, sur les objets du quotidien. À partir de l'œuvre de cette communauté passionnée par les lettres et leurs esthétiques variées, a été développé un logiciel permettant d'associer à chaque signe son équivalent photographié et ce grâce à l'indexation des images dont se sert l'algorithme du site. Les mots, phrases entrées dans le moteur de recherche du mashup, se transforment en assemblages de styles graphiques, à partir de la base de données du groupe devenant police des polices.



III. 10 Capture d'écran de la galerie du groupe One Letter

3. Mashups artistiques

Le principe même du mashup se trouve également réinvesti dans les travaux de quelques artistes. On se réfère ici aux deux œuvres du flasheur Mario Kligenmann, *Flickeur*⁷⁴ et sa variante *Island of consciousness*⁷⁵. Associant un logiciel flash et l'API de Flickr, *Flickeur* extrait de façon aléatoire des images du site de partage pour créer un film infini aux frontières de l'esthétique documentaire, du clip vidéo, et de la vision surréaliste. De la même façon que la sélection des images, tous les fondus, mouvements, répétitions, effets de zooms ou d'accélération visibles sur l'image sont gérés par un algorithme basé

⁷³ <http://www.flickr.com/groups/oneletter/>

⁷⁴ <http://incubator.quasimondo.com/flash/flickeur.php>

⁷⁵ http://incubator.quasimondo.com/flash/islands_of_consciousness.php

sur le hasard. Dans *Island of consciousness* une bande son de Oleg Marakov est incorporée selon ces mêmes principes, des boucles musicales sans structures établies répondent alors au flux des images. Si *Flickeur* fonctionne sur le principe de la boucle vidéo virtuelle, sans fin et en perpétuel changement, *Island of consciousness* comporte pour sa part un algorithme modifié influant sur la sélection des images qui n'apparaissent plus de façon totalement aléatoire mais selon « an associative stream-of-consciousness rule - each new image shares at least one tag with the previous one ⁷⁶ ». Ce sont donc les tags des images de Flickr qui permettent à l'application de fonctionner, si *Flickeur* proposait un film comme une simple juxtaposition d'images, un enchevêtrement de photographies sans trame narrative, *Island of Consciousness* par le recours aux catégories des tags introduit une forme de suivi thématique dans le défilement des images.

Le résultat apparaît toujours comme un diaporama dynamique et anarchique mais la contiguïté sémantique entre deux images qui se laisse parfois saisir nous entraîne souvent à reconstituer un sens, une narrativité à ce flux ininterrompu d'images.

Ces deux œuvres offrent une approche inédite des photographies de Flickr, une nouvelle voie d'exploration de cette base de données aussi large et infinie que le film qui en est constitué. Nous assistons à la reversion en matériaux artistiques d'images personnelles, le potentiel de publicisation de ces images induit par leur mise en ligne et leur propension au libre partage est ici poussé à son paroxysme, ces images prennent part à un projet conceptuel qui les dépasse, elles feront partie à un moment donné, moment parfaitement imprévisible, peut-être unique, peut-être réitéré, de cet espace créatif voué à l'exposition publique. Et les scènes photographiées que nous voyons loin de demeurer des instants figés de la sphère privée, se transforment par cette entreprise artistique de décontextualisation de l'image en informations visuelles à décrypter. Les images amateurs deviennent moments de l'œuvre, matière filmique, et *Island of consciousness* leur restitue une autonomie en les faisant circuler, en parallèle de leur diffusion, dans un volet indépendant. En permettant de retrouver les sources originales de ces images sur Flickr, l'artiste joue d'autant plus sur l'ubiquité, la dualité de ces photographies à la fois documents personnels et documents mis en ligne, offerts au regard de tous.

C. Flickr Commons et les nouvelles pratiques documentaires

Ainsi les photos amateurs deviennent par leur mise en ligne sur Flickr des documents visuels appartenant à une base de données exhaustive, et non plus seulement à

⁷⁶ Citation extraite de la page d'accueil de *Island of consciousness*

un photographe. Leur rattachement à des catégories via les tags participe de cette ouverture de l'image dans l'espace public, « dissociée de son origine, identifiée seulement par des tags sémantiques et placée dans des groupes avec d'autres images qui partagent alors des métadonnées similaires, l'appréhension d'une image est dépendante de l'interface qui en médiate sa diffusion⁷⁷ ». Les images, par le site internet, se défont de leur privauté pour s'offrir comme document accessible à tous.

Selon un effet miroir, est apparue une logique de mise à disposition des collections de photos d'institutions publiques sur les plateformes ouvertes engendrant leur appropriation par les usagers du site au travers, notamment, des opérations d'indexation et d'annotation de ces images. C'est véritablement à une opération de documentarisation des images en ligne à laquelle Flickr se livre alors grâce aux outils de taggage ordinaires et à la force de la collaboration sociale.

Selon Michel Salaün, documentariser c'est « ni plus ni moins "traiter un document" comme le font, ou le faisaient, traditionnellement les professionnels de la documentation (bibliothécaires, archivistes, documentalistes) : le cataloguer, l'indexer, le résumer, le découper, éventuellement le renforcer, etc. On préfère "documentariser" à "documenter", qui renvoie plutôt à la création d'un ou de plusieurs documents pour expliquer un objet ou une action, mais dans nombre de cas les deux activités se recoupent. L'objectif de la documentarisation est d'optimiser l'usage du document en permettant un meilleur accès à son contenu et une meilleure mise en contexte.⁷⁸ ». Une pratique qui, aujourd'hui, est rendue accessible aux usagers du net et qui dépasse la seule sphère de leurs documents privés pour s'étendre aux collections publiques, corpus que se réapproprient le public dans une logique d'accroissement des bases de données et de rationalisation de l'accès aux images.

1. The Commons sur Flickr

En Janvier 2008 The Library of Congress⁷⁹ met en ligne sur Flickr une partie de ses collections de photographies, dans un premier temps et de façon à tester cette opération ce sont deux corpus historiques qui sont accessibles via Flickr⁸⁰, un ensemble de photos sur

⁷⁷Daniel Rubinstein et Katrina Sluis, "À life more photographic", *Photographies*, Vol n°1, septembre 2008, traduit par nos soins.

⁷⁸ Jean-Michel Salaün, « La redocumentarisation, un défi pour les sciences de l'information » *Études de Communication* n° 30, Entre information et communication, Les nouveaux espaces du document, Université de Lille3, décembre 2007.

⁷⁹ <http://www.loc.gov/index.html>

⁸⁰ http://www.flickr.com/photos/library_of_congress/

l'actualité des années 1910 d'une part, et des photos en couleur prises dans les années 1930-1950 d'autre part. Le but n'est pas simplement d'offrir une vitrine à ces images mais bien de profiter des aspects sociaux du site et ainsi faire participer les usagers de Flickr à l'indexation et à la documentarisation des fonds de l'institution : « La librairie vous invite à explorer l'histoire de façon visuelle en venant explorer les importantes photos de nos collections. Sentez-vous libre d'ajouter des tags et des commentaires également ! Nous avons besoin de toujours plus de mots pour aider le plus grand nombre à accéder et user de ces images.⁸¹ ». À la question : "pourquoi la Library of Congress met-elle ses photos sur Flickr ?", la réponse de la FAQ du projet est limpide :

- *"Pour partager ces collections photographiques avec ceux qui ne visitent pas notre site"*
- *"Pour mieux comprendre en quoi l'indexation sociale peut nous apporter des idées d'amélioration de nos services et de nos fonds"*
- *"Pour accroître notre expérience et notre implication dans des communautés Web partageant des centres d'intérêts communs avec les collections de bibliothèques."⁸²*

La Bibliothèque conclut un partenariat avec Flickr appelé *The Commons*, ce programme s'est fixé deux objectifs principaux, augmenter l'accès aux classeurs tenus par un établissement public et permettre au grand public d'apporter des informations et des connaissances.

Le succès de cette initiative convainc la Library of Congress de rendre accessible de nouveaux corpus de ses collections, de nouvelles Institutions vont également prendre part à l'expérience. On compte aujourd'hui pas moins de vingt-trois organismes partenaires, rassemblant aussi bien des bibliothèques que des musées, escomptant tous de cette collaboration avec Flickr qu'elle aboutisse à une documentarisation des corpus en ligne, mais aussi à la découverte par un public non habitué des fonds institutionnels de leurs collections. Selon Olivier Ertzscheid, « Mettre à disposition des fonds documentaires professionnellement construits et administrés, sur des services privés mais publiquement accessibles et massivement accédés, autoriser les publics, TOUS les publics, à indexer à leur tour ces fonds documentaires quelle que soit leur nature, laisser les publics se les approprier (l'un des objectifs du projet est de permettre aux usagers d'indexer, de tagguer ces images), c'est probablement le meilleur moyen de recréer le désir ... d'aller en

⁸¹ http://www.loc.gov/rr/print/flickr_pilot.html , traduit par nos soins.

⁸² http://www.loc.gov/rr/print/flickr_pilot_faq.html

bibliothèque.⁸³». Même sans aboutir à l'émergence d'un tel désir, ce genre de projets engendre nécessairement pour ceux qui y participent une prise de connaissance des images disponibles et des institutions. Mais surtout, et c'est là le premier constat effectué par la Library of Congress, cette collaboration est fructueuse en ce qui concerne les bénéfiques documentaires, puisque en octobre 2008 on pouvait compter 4929 photos mises en ligne, 7166 commentaires et plus de 70000 tags ont été déposés⁸⁴.



III. 11 Capture d'écran de la galerie de photographies de la Bibliothèque du Congrès sur Flickr

Ces résultats positifs s'expliquent, du côté des participants, par l'intérêt suscité de faire partie d'une opération de grande envergure d'indexation d'une collection, cette pratique ressort de la même logique que celle visant à documenter de façon systématique et sérieuse ses propres images avec pour toile de fond l'envie de favoriser l'accès des contenus au plus grand nombre. Comme le note Maxime Crepel, ces usages d'indexation sont associés à un discours altruiste, faisant référence au bien commun et à la notion de partage et qui « semble assez proche des motivations des membres actifs de Wikipedia qui justifient leur investissement au sein de l'encyclopédie en ligne par une volonté de participer à un projet qui rend la « connaissance accessible pour tous » et endossent la responsabilité d'assurer la

⁸³ Olivier Ertzscheid, « La bibliothèque du congrès et FlickrR : des accès desaxés », *Affordance.info*, janvier 2008, (en ligne http://affordance.typepad.com/mon_Weblog/2008/01/la-bibliotheque.html).

⁸⁴ Cf. le résumé du rapport, http://www.loc.gov/rr/print/flickr_report_final_summary.pdf

qualité et l'intégrité des connaissances qui y sont disponibles.⁸⁵ ». Toutefois, il convient également de voir en ces pratiques documentaires une volonté d'affirmation de son identité et d'intégration, de reconnaissance par la communauté des internautes. Une reconnaissance pouvant en l'occurrence se baser sur l'échange de savoirs sur l'histoire de la photographie ou de ses techniques, il s'agit donc de mettre à partie ses connaissances pour se retrouver entre amateurs, des communautés d'experts se tissent alors tous animés par un goût commun de l'image et de sa documentarisation.

Échantillon de trésors fournis par les Organismes publics [Lancez une recherche](#) | [Afficher par tag](#)



Egypt: Gizeh
from Brooklyn Museum



Study of a small girl with a prize Scottish terrier dog, c. 1935 / by...
from State Library of New South Wales collection



Mosman Bay Falls
from Powerhouse Museum Collection



Ill. 12 Capture d'écran de la galerie d'exemples de photographies fournies par les Institutions participant à The Commons sur Flickr

En citant Manuel Zacklad, il serait ici plus juste de parler de re-documentarisation. Selon l'auteur, « redocumentariser, c'est documentariser à nouveau un document ou une collection en permettant à un bénéficiaire de réarticuler les contenus sémiotiques selon son interprétation et ses usages à la fois selon la dimension interne (extraction de morceaux musicaux pour les ré-agencer avec d'autres, ou annotations en marge d'un livre suggérant des parcours de lecture différents...) ; ou externe (organisation d'une collection, d'une

⁸⁵ Maxime Crepel, « Les folksonomies comme support émergent de navigation sociale et de structuration de l'information sur le Web », *Réseaux*, n° 152, juin 2008, p. 169-204, (en ligne : <http://www.cairn.info/revue-reseaux-2008-6-p-169.htm>).

archive, d'un catalogue privé croisant les ressources de différents éditeurs selon une nouvelle logique d'association)⁸⁶».

2. Flickr comme moteur de recherche

Une des conséquences de l'apparition des *Commons* sur Flickr, au-delà des nouvelles collaborations sociales que le projet introduit et formations de nouvelles communautés expertes, s'entend dans l'augmentation considérable de la base de données du site. Si déjà cette plateforme est connue pour diffuser un nombre formidable d'images, pour la qualité de nombres d'entre elles, qualité nous l'avons vu favorisée par les échanges entre membres, Flickr se dote désormais de photographies historiques, pour certaines professionnelles, et toutes à teneur patrimoniale. Des collections qui de plus bénéficient des outils performants d'indexation du site, leur offrant une visibilité en constante augmentation. À côté de ces images anciennes connaissant une seconde vie sur le site et offrant par la même un choix accru d'images aux internautes, la politique de libre partage et circulation des images trouve un écho au niveau amateur avec les licences *Creative Commons* qui peuvent être appliquées aux images de tout un chacun sur Flickr. Créées en 2002 par un collectif pour qui la propriété intellectuelle ne saurait se résumer à la propriété physique sans engendrer un frein dans la diffusion de la culture, ces autorisations non exclusives, et pouvant s'appliquer sur tout type de contenu, permettent aux titulaires de droits d'autoriser le public à effectuer certaines utilisations, tout en ayant la possibilité de réserver les exploitations commerciales, les œuvres dérivées et de mentionner la paternité de l'œuvre.

Flickr a dépassé le cap des 100 millions de photos sous licences *Creative Commons* quelque part pendant la journée du samedi 21 mars 2009. À la date de lundi 23 mars 2009, 100 191 085 photos sous licence CC étaient recensées. L'intégration des licences CC dans Flickr est à la fois l'une des premières existant à cette échelle et l'une des meilleures, non seulement elle permet aux utilisateurs d'associer chaque photo à une licence différente, mais Flickr propose une page de découverte des images classées par type et association de licence CC. Il est par exemple possible de visualiser toutes les photos de New York sous licence CC Paternité, ou sans utilisation commerciale disponibles et triées par niveau d'intérêt.

⁸⁶ Manuel Zacklad, « Réseaux et communautés d'imaginaire documédiatisées », in Skare, R., Lund, W. L., Varheim, Å., *À Document (Re)turn*, Frankfurt, Peter Lang, 2007, 279-297, (en ligne http://hal.archives-ouvertes.fr/docs/00/18/01/85/PDF/imaginaire_document_Skare_Lund_Varheim.pdf)

Creative Commons

De nombreux utilisateurs de Flickr ont décidé de placer leurs œuvres sous licence Creative Commons. Vous pouvez parcourir ou rechercher des éléments en fonction de leur type de licence.

Voici quelques éléments ajoutés récemment :

Licence Paternité



De [dbdrobot](#)

De [dbdrobot](#)

De [parvbembridge](#)

De [dbdrobot](#)

» 13 407 436 photos [\(Plus\)](#)

Licence Paternité-Pas de modification



De [theSuperStar](#)

De [C.Y.R.L.L.](#)

De [bleuets9](#)

De [Toesmasha](#)

De [Toesmasha](#)

» 4 646 636 photos [\(Plus\)](#)



"Creative Commons est une organisation à but non lucratif proposant une alternative au copyright complet."

creativecommons.org

En bref...

Paternité signifie :

Vous autorisez la copie, la distribution, l'affichage et la communication de votre œuvre protégée. Vous autorisez également sa modification, à condition que votre nom soit cité.



Pas d'utilisation commerciale signifie :

Vous autorisez la copie, la distribution, l'affichage et la communication de votre création. Vous autorisez également sa modification, mais à des fins non commerciales uniquement.



III. 13 Capture d'écran de la page *Creative Commons* de Flickr

Ces photos ont été utilisées dans des centaines de milliers d'articles de Wikipédia, de messages postés sur des blogs, et même dans des articles de presse traditionnelle ; autant d'exemples de créations nouvelles qui n'auraient autrement pas vu le jour sans ces licences publiques standardisées. Cette possibilité non seulement d'utiliser mais pour certains de remixer les images, le choix de photographies disponibles sur le site font de Flickr une nouvelle base de données d'images dont les recherches documentaires peuvent s'avérer plus pertinentes que GI. En effet en alliant ses collections basiques, celles plus spécifiques *Creatives Commons* et les ensembles *the Commons* avec les systèmes d'indexation du site les requêtes et autres désirs de documentation visuelle sont largement fournies et permettent des résultats pertinents.

Toutefois en ce qui concerne l'utilisation, la copie ou le remix d'images des institutions publiques disponibles sur Flickr : les images mises en ligne au sein des *Commons* sont alors soumises à un copyright spécifique et doivent comporter la mention « Aucune restriction de copyright connue », les institutions culturelles stipulent ainsi qu'il n'existe pas de restriction de copyright pour une photo, celles-ci peuvent être légalement diffusées dans le cadre Flickr, mais cela ne garantit pas la possibilité pour l'internaute de pouvoir utiliser cette image en dehors du site.

Si l'internaute peut avoir accès aux images pour participer à l'indexation collaborative, ou pour simplement prendre connaissance des collections d'un musée ou bibliothèque présent

sur le site, toutes les images ne sont pas libres de droits et ne peuvent être librement diffusées. Ce statut des *Commons* complique l'utilisation des images puisqu'il s'agit de se renseigner au cas par cas de l'état juridique d'une image.

III. Les limites de la folksonomie

La folksonomie est devenue un formidable outil dont les possibilités semblent, d'ailleurs, encore largement sous-exploitées, toutefois elle connaît depuis son apparition de nombreuses critiques, un nombre constant de voix s'élevant pour dénigrer ce système d'indexation collaboratif en comparaison avec les indexations professionnelles jugées bien plus sérieuses, ou pour pointer ses manques et limites.

A. Polysémie, synonymie et subjectivité

1. Problèmes de l'indexation sémantique des images

L'indexation des images par des mots, bien qu'incontournable, reste une gageure tout d'abord en raison de la multiplicité des entités présentes à l'image. Dès les premières tentatives de description d'une photographie, nous nous trouvons confrontés à l'infinité des éléments qui la composent. Il semble que la charge sémantique d'un mot et la charge sémantique d'une photographie soient incomparables. D'où la multiplication du nombre de tags nécessaire pour circonscrire au plus près la signification, les réseaux de thèmes dont une image est vectrice. Comme nous l'avons vu avec le mashup Fastr, définir une image à partir d'un seul terme s'apparente à une véritable réduction de l'image. D'autant plus que le terme qui fait sens pour moi, ne sera pas forcément celui partagé par la majorité, une image est en effet le support d'une multitude d'interprétation possible.

Nous sommes tous portés à faire des images différentes lectures à travers nos propres schémas culturels, cognitifs, les émotions qu'elle suscite ou non en nous. Cette polysémie de l'image est magistralement affirmée par Régis Debray : « Une image est à jamais et définitivement énigmatique, sans « bonne leçon » possible. Elle a cinq milliards de versions potentielles (autant que d'êtres humains), dont aucune ne peut faire autorité (pas plus celle de l'auteur qu'une autre). Polysémie inépuisable. On ne peut faire dire à un texte

tout ce qu'on veut - à une image, oui. (...) Non, il n'y a pas de perception sans interprétation. Pas de degré zéro du regard⁸⁷».

Michel Foucault souligne également l'irréductibilité des images et des mots, parlant de la peinture, l'auteur remarque que : « le rapport du langage à la peinture est un rapport infini. Non pas que la parole soit imparfaite et en face du visible dans un déficit qu'elle s'efforcerait en vain de rattraper. Ils sont irréductibles l'un à l'autre : on a beau dire ce qu'on voit, ce qu'on voit ne loge jamais dans ce qu'on dit, et on a beau faire voir, par des images, des métaphores, des comparaisons, ce qu'on est en train de dire, le lieu où elles resplendissent n'est pas celui que déploient les yeux, mais celui que définissent les successions de la syntaxe⁸⁸.».

Barthes déjà dans son article « Rhétorique de l'image » observait « (...) toute image est polysémique, elle implique, sous-jacente à ses signifiants, une chaîne flottante de signifiés, dont le lecteur peut choisir certains et ignorer d'autres. ⁸⁹ » Et ce quelque'un n'est pas qu'un sujet psychologique, il est aussi sujet biologique, historique et social, autant de contexte influant sur la perception et l'interprétation qui est faite de l'image. Ce sont ces débordements mêmes de signifiés, ces réseaux de significations qui rendent impossible le fonctionnement monosémique de l'image. « La polysémie », dit encore Barthes, « produit une interrogation sur le sens ; or cette interrogation apparaît toujours comme une dysfonction (...) Aussi, se développent dans toutes les sociétés des techniques diverses destinées à fixer la chaîne flottante des signifiés de façon à combattre les signes incertains : le message linguistique est l'une de ces techniques. ». C'est ainsi à la langue et au message linguistique de préciser le sens de l'image, de le fixer, enfin d'arrêter la dérive sémantique à laquelle toute image convie.

2. Polysémies des tags

Si une description, une légende peut permettre de circonscrire les réseaux de signification d'une image, qu'en est-il pour les tags, simples mots-clés portant en eux-mêmes une signification limitée ?

La polysémie s'avère en ce sens limitée, dépendant entièrement du jugement d'un individu le tag n'est finalement qu'une chaîne de caractères dont le sens exact est connu de lui seul. L'indexation par mots-clés est infiniment subjective et ne peut surtout résoudre les

⁸⁷ Régis Debray, *Vie et mort de l'image*, Gallimard, Paris, 1994. 526 p.

⁸⁸ Michel Foucault, *Les mots et les choses*. Gallimard, Paris, 1966. 405 p.

⁸⁹ Roland Barthes, "Rhétorique de l'image", *Communication*, n°4, 1964.

problèmes de polysémie des images en particulier lorsqu'elle lui surajoute celle du mot isolé. Car c'est bien à l'impossibilité d'organiser les tags entre eux sous la forme d'une taxinomie ou d'un thésaurus à laquelle nous sommes confrontés dans une logique de folksonomie. Là où les langages contrôlés gèrent les liens sémantiques et syntaxiques du vocabulaire, les folksonomies dans leur organisation libre et non hiérarchisée ne le permettent pas. Deux tags différents peuvent être synonymes voire identiques ou presque sans qu'on puisse les relier, ou encore deux utilisateurs peuvent partager un tag identique, mais en avoir une conception différente. « Dès lors, c'est bel et bien une forte "non-organisation", voire "désorganisation", qui préside à l'évolution du contenu tagué⁹⁰ », les tags constituent dans leur globalité une simple liste de mots et d'expressions de même niveau ne permettant pas réellement de gérer la synonymie. Ainsi comment différencier les images tagguer avec « avocat », quand certaines se réfèrent au fruit et d'autres à la profession, de même pour les noms de villes semblables mais de pays différents. Effectuer une recherche multi-tags apparaît pour l'instant comme l'unique solution pour restreindre les champs de signification possible.

Pour certain la solution consisterait à rétablir un minimum de hiérarchie au sein du système, en affinant les catégories. Ainsi, Flickr a introduit dès 2006 la notion de groupes, séries (*clusters*) de tags, permettant de façon simple par le rapprochement de tags entre eux de préciser le sens du terme et de lier les images aux catégories ainsi soulignées. Par exemple, le mot "avocat" amène à une page proposant plusieurs "clusters", séries associant mots-clés liés et images thématiques, dont l'un regroupe des images du fruit, un autre des photographies de tribunal et de magistrats et un troisième, des photos de recettes réalisées à partir de l'aliment.

B. Problème des multilinguismes et Restriction du champ sémantique

1. Régulation par les foules

L'un des problèmes de la folksonomie largement soulevé par ses réfractaires est celui de la stabilité du système dans l'accès et la description des documents concernés. Si tout un chacun peut indexer par n'importe quel mot-clé considéré comme adéquat, quelle

⁹⁰ Olivier le Deuff, « Folksonomies : Les usagers indexent le web », *BBF*, 2006, n° 4, p. 66-70 en ligne <http://bbf.enssib.fr>.

est l'instance permettant une régulation des entrées sémantiques, de leur globale pertinence ?

L'étude de Scott Golder et Bernardo Huberman⁹¹ tente de répondre à cette question. Dans leurs travaux, ils analysent deux corpus de données issues du système d'indexation collaborative de Del.icio.us. Le premier analyse les tags correspondant à un ensemble de 212 URL elles-mêmes réparties dans 19422 signets. Le second ensemble comprend cette fois une liste aléatoire de 68 668 signets. Selon les auteurs, les faiblesses documentaires des systèmes de « tags », comme l'absence de prise en compte des relations sémantiques, ou de hiérarchie entre les termes, ne semblent pas discriminant. Les tags parviennent à engendrer des modèles récurrents de catégorisation des données, et il ressort ainsi que des usagers indexent à l'aide des mêmes « tags » des sites donnés. Les conclusions des deux auteurs attestent que des régularités apparaissent indépendamment du type de tag employé. Si les usagers semblent chacun employer une grande variété de tags, l'analyse de la répartition de ceux-ci dans le temps montre qu'il existe des phénomènes de saturation qui semblent indiquer l'émergence de consensus autour de la définition des ressources auxquelles ils sont indexés. La suggestion de tags déjà indexés entraînerait une paralysie qui amènerait les utilisateurs à adopter des tags similaires pour catégoriser un même contenu. C'est, semble-t-il, la mise en partage, la circulation des tags qui fait fonction de régulation (feedback). Le collectif permet d'atténuer les subjectivités, il est alors possible d'assister à l'élaboration de descriptions homogènes, de catégories reconnues de tous. Une indexation représentative des objets se met en place selon un principe d'auto-régulation de nos perceptions documentaires. Dans une interview Adam Bosworth, Vice-Président en charge de l'ingénierie de la firme Google confirme cette capacité de la folksonomie à engendrer des systèmes d'indexation stables : « La sagesse des foules fonctionne étonnamment bien. Les systèmes qui marchent sur le Web fonctionnent du bas vers le haut ('bottom-up'). (...) Par exemple Flickr ne dit pas à ses utilisateurs quel tag utiliser pour leurs photos. Loin de là. N'importe qui peut déposer n'importe quel tag sur n'importe quelle photo. Mais - et c'est la clé - Flickr offre un retour sur les tags les plus utilisés et les plus populaires, et les gens souhaitant attirer l'attention sur leurs photos (...) apprennent rapidement à utiliser ce lexique si celui-ci fait sens. Cela rend le système étonnamment stable. Del.icio.us fait la même chose. Le succès de Google pour rendre les recherches plus

⁹¹ Golder Scott À. & Huberman Bernardo À., « The Structure of Collaborative Tagging Systems », Information Dynamics Lab, HP Labs , 2005. (En ligne : <http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>)

pertinentes reposait sur la puissance de cette sagesse populaire (PageRank)...⁹²». Il ressort de ces propos que si le principe de communautarisation de l'indexation peut-être garant de son homogénéité et donc de son efficacité, les systèmes de gestion des tags mis en avant par les sites de partage eux-mêmes jouent également un rôle de modérateur dans l'utilisation de tel ou tel terme, dans sa mise en avant au sein du nuage et de son adoption par les usagers.

2. Un vocabulaire limité, Folksonomie et multilinguisme

Toutefois si les modes de gestion des tags adoptés par les sites comme Flickr s'avèrent efficaces pour la régulation de l'indexation, ne peuvent-ils pas également engendrer une uniformisation restrictive des termes utilisés pour caractériser une image ? Sous prétexte d'une visibilité accrue les internautes se tourneraient vers l'emploi récurrent d'un terme banalisé, unanimement adopté. Mais quelle pertinence est envisagée quand il s'agit de relier une image à une catégorie renfermant des milliers d'items, de même, est-ce réellement efficient d'employer un tag pour sa popularité plutôt que pour son adéquation relative à l'image. Cette pratique semble, pour reprendre l'expression d'Éric Sutter, être source d'*info-pollution*⁹³, vidant de leur sens de nombreux mots-clés, devenant indexeur vide parce que trop employé et selon des orientations pas assez circonscrites. Comme l'explique Olivier Le Deuff : « Les photos recevant les mots-clés de *vacances* sont tellement innombrables, que cela n'a plus grand intérêt si ce n'est pour celui qui les a indexées ainsi⁹⁴ », les singularités lexicales, l'unicité des termes, semblent s'effacer au profit de la récurrence et d'une uniformisation de la dénomination, de même qu'on assiste à des répétitions iconiques, que se multiplient les images d'un même contenu visuel. Comme le précise Daniel Rubinstein et Katrina Sluis, « Au travers des mécanismes sémantiques du tagging et des métadonnées ajoutées, c'est toute la spécificité de chaque photographie en ligne qui semble effacée et ce par la faculté d'un simple mot-clé capable de regrouper des milliers d'images disparates sous un hyperlien. Mais 4,150,058 photographies taguées avec 'party' peuvent-elles être incorrectes ?⁹⁵ »

⁹²Traduit par mes soins, <http://acmqueue.com/modules.php?name=Content&pa=showpage&pid=337>

⁹³ Eric Sutter, « Pour une écologie de l'information », *Documentaliste - Sciences de l'information*, no 2, vol. 35, 1998, p. 83-86

⁹⁴ Olivier Le Deuff, « Folksonomies : Les usagers indexent le Web », *BBF*, 2006, n° 4, p. 66-70 (en ligne <http://bbf.enssib.fr>)

⁹⁵ Daniel Rubinstein et Katrina Sluis, « À life more photographic », *Photographies*, Vol n°1, septembre 2008, traduit par nos soins.

Le système en visant la stabilité aboutit ainsi à une limitation sémantique. Si ce principe est central dans les systèmes d'indexation métiers, attaché à la multitude d'images présentes sur ces sites de partage il tend à confiner une masse de données sous un même terme indexeur. La recherche d'une image devient inopérante puisque renvoyant à une galerie de résultats pouvant se compter en milliers. Une autre critique adressée à l'indexation par tags est celle de la pauvreté des termes utilisés, la majorité de ceux employés sont bien souvent génériques, ils n'apportent alors que peu d'éléments sémantiques, aucune catégorisation du contenu et participe de cette globalisation neutralisante de la catégorie.

Il convient également de traiter ici du problème du non multilinguisme de la folksonomie, il n'existe pas encore de moyens pour relier un même tag selon qu'il soit exprimé en anglais, en français ou en tout autre langue. Le site Technorati⁹⁶ propose sur sa page de tags, de visionner tous les mots-clés enregistrés qu'elle que soit leur langue, le nuage de tags se transforme alors en une liste sur laquelle des caractères asiatiques ou arabes côtoient une multitude de termes anglais. Ce système est loin d'offrir une lisibilité à l'ensemble, de plus de nouveaux soucis sémantiques apparaissent inhérent à l'utilisation non spécifiée de divers langages, ainsi le mot-clé "pain" fait-il référence à l'aliment en français, ou plutôt au terme anglais se référant à la douleur ? Il convient toutefois de noter la prédominance des termes anglais dans les index, assurant à l'utilisateur employant cet idiome en langue de requête d'obtenir des résultats pertinents. Bien sûr et selon un principe circulaire ce quasi monopole de la langue anglaise, du moins en terme d'efficacité des requêtes, ne peut que s'accroître, les usagers étrangers pour assurer une bonne indexation et circulation de leurs images ayant alors recours à des tags anglais, augmentant ainsi leur nombre dans la base.

Selon de nombreux spécialistes de la documentation, l'indexation par les usagers peut sembler moins coûteuse en temps d'indexation, mais c'est le temps passé à retrouver l'information qui s'accroît. On se place ainsi dans le schéma exactement inverse des systèmes d'informations hiérarchisées, où c'est le coût d'indexation par les professionnels qui est plus élevé, mais la recherche d'informations facilitée pour l'utilisateur.

⁹⁶ <http://technorati.com/tag/?showall=1>

Chapitre 3

Le géotagging comme nouveau système d'indexation

Sommaire du chapitre

I. Les présupposés d'une indexation par coordonnées géographiques	76
À. L'indexation par les métadonnées EXIF	76
1. La remise en question de la photographie comme index	76
2. La structure informationnelle des images	78
3. Les métadonnées comme outil de classement	79
B. La géolocalisation.....	81
1. Les principes	81
2. Les photographies géolocalisées.....	82
II. Les mises en application sur le Web.....	84
À. La géolocalisation dans les sites de partage	84
1. Flickr	84
2. Zoomr.....	86
3. Picasa	87
B. Des exemples de sites dédiés.....	88
1. Panoramio	88
2. Locr.....	90
3. Loc.alize.us	91
III. Vers un déplacement de la sphère documentaire.....	92
À. Vers une recherche documentaire incarnée	92
1. Les groupes Flickr	93
2. Panoramio ou le nouvel ordre documentaire	96
3. Echanger son expérience en plus de ses images	98
B. Vers de nouveaux modes de recherche d'images.....	100
1. Les applications composites : géolocalisation et locative media.....	100
2. L'image comme interface de navigation.....	103

I. Les présupposés d'une indexation par coordonnées géographiques

Depuis le début des années 2000 de plus en plus de travaux de recherche visent à mettre au point des systèmes automatiques de classement des images basés sur la gestion de leur localisation géographiques. Nous concevons le géotagging comme une méthode d'indexation nouvelle et originale proposant une alternative à l'indexation purement textuelle des images. Divers sites disposent aujourd'hui de cet outil et proposent par ce biais une forme originale de navigation au sein des bases de données où l'approche visuelle prime.

A. L'indexation par les métadonnées EXIF

Nous l'avons vu précédemment les métadonnées peuvent renouveler les possibilités d'indexation, engendrer une nouvelle forme de base de données de notices et faire correspondre les informations documentaires directement à l'image, de façon interne. Si les champs de métadonnées descriptives et textuelles servent aujourd'hui les professionnels de la documentation, d'autres types de métadonnées ont été mises au point et permettent une indexation jusqu'ici inédite des images et ouverte à tous.

1. La remise en question de la photographie comme index

L'indexation textuelle est largement majoritaire au sein du traitement et de la gestion des bases de données d'images. Le recours à la sémantique, puis à la folksonomie est apparu comme une clé d'organisation des collections en ligne, par l'intermédiaire des mots et du langage naturel, les mots-clés d'indexation permettant alors de représenter le contenu des documents : thèmes, sujets, aspects, etc. Une transparence semble ainsi se former entre le mot et l'image, le terme choisit pour la représenter paraphrasant le contenu de l'image, l'image se réduisant au mot. La photographie en effet, parce qu'elle ressemble à ce qu'elle représente et qu'elle est donc proche de notre expérience perceptive, est réputée compréhensible immédiatement par tout un chacun. L'image donnerait à lire sa signification sans la médiation de conventions, par pure transparence : son sens lui serait donc immanent, aisément traduisible par l'interface langagière, l'image représentant son référent il suffirait de nommer celui-ci pour aboutir à la réduction verbale de l'image, à la clé de son indexation.

Comme le souligne Patrick Peccatte : « La prégnance de cette formulation indicielle dans les études visuelles, dont on peut s'étonner qu'elles ne semblent pas avoir proposé d'autres concepts majeurs depuis plus de vingt ans, montre qu'il reste difficile de s'affranchir de la forme ontologique du lien qu'entretiendraient un objet (photographié) avec un sujet (photographiant) produisant un autre objet (photographique). Et cela, même sans théorisation particulière.⁹⁷ ». Selon l'approche documentaire et textuelle, l'indexation est une représentation extérieure, surajoutée forcément réductrice du contenu.

Le développement de la photographie numérique a engendré une remise en cause de la notion établie d'index comme caractérisant le rapport de la photographie au photographié. Cette soudaine méfiance passe dans un premier temps par la prise en compte de la « dématérialisation » de l'image. L'absence de processus et de traces chimiques sur une surface sensible semble interférer avec la notion indicielle de la photographie, lui être antinomique et provoque le déclassement de l'image numérique comme objet photographique. Comme l'explique Tom Gunning : « L'indicialité de la photographie dépend d'une relation physique entre l'objet photographié et l'image ainsi créée. L'image se forme sur le négatif grâce à la transformation de l'émulsion photosensible sous l'action de la lumière réfléchi par l'objet photographié et passant à travers les objectifs et le diaphragme. Dans une image numérique, en revanche, au lieu d'une émulsion photosensible affectée par l'objet lumineux, la formation de l'image répond au codage de données sur la lumière en une matrice de chiffres.⁹⁸ ». Dans cette perspective beaucoup considèrent que la relation de contiguïté entre l'objet et l'image se trouve rompue avec le numérique. Depuis les interprétations ont rétabli la photographie numérique dans sa potentialité d'empreinte, offrant donc au numérique une nature photographique possible.

À posteriori donc la polémique a permis de revenir sur le monopole théorique de l'indicialité forgé par Krauss, et de faire émerger de nouvelles considérations quant au statut épistémologique de la photographie. Ainsi apparaissent des théories selon lesquelles : « L'assimilation de l'objet photographique à un simple fichier informatique n'est en soi qu'une approximation sans grand intérêt mais permet de comprendre très

⁹⁷ Patrick Peccatte, « Notes sur la structure informationnelle de la photographie », *Du bruit au signal (et inversement)*, 5 août 2008, en ligne <http://blog.tuquoque.com/post/2008/08/05/Notes-sur-la-structure-informationnelle-de-la-photographie>

⁹⁸ Tom Gunning, "La retouche numérique à l'index. Pour une phénoménologie de la photographie" (traduit de l'américain par Marc Phéline), *Études photographiques*, n° 19, décembre 2006, p. 96-119.

facilement en quoi il est nécessaire et naturel de convoquer d'autres cadres conceptuels que la sémiotique pour progresser⁹⁹ », et de la sémantique pour indexer.

Si le numérique offre bien une révolution elle survient non au niveau ontologique de la photographie mais à celui de son indexation, une indexation qui lui est consubstantielle. Ainsi n'importe quelle partie du support d'enregistrement d'un document électronique se trouve d'une part susceptible d'être accédée comme contenu documentaire, d'autre part de servir d'index pour ce contenu. On peut ainsi affirmer que tout système numérique est un système documentaire et mobilise une indexation. En fait, un système numérique est un système documentaire dès qu'il en existe une indexation.

2. La structure informationnelle des images

Nous proposons donc de considérer la photographie comme une structure non plus strictement indicielle, mais « informationnelle » selon l'expression de Patrick Peccatte. Si pour Barthes la sphère iconographique était bien définie par une structure informationnelle cette acception devait être comprise dans son sens sémiotique (informatif ou documentaire), l'image n'étant rien sans les mots, l'univers du texte. Nous envisageons la photographie non seulement pour ce qu'elle représente, mais pour les informations qu'elle contient dont l'image à proprement parler n'est qu'une strate, visuelle en l'occurrence, parmi d'autres. Il en ressort une tentative d'offrir à la photographie une nouvelle définition, affranchie de sa dépendance à la notion d'indice et prenant en compte les paramètres plus larges de ses composants, on suit Patrick Peccatte quand il propose de voir la photographie numérique comme « une structure d'informations typée dont une des composantes (l'image) peut être restituée visuellement à l'aide d'un dispositif banalisé. L'image et une partie des autres composantes de la structure enregistrent directement et automatiquement certains paramètres caractéristiques d'un événement tandis que les autres composantes générées indirectement décrivent l'événement ou l'image elle-même.¹⁰⁰ ». Une photographie qu'elle soit numérique ou non ne doit plus, dans cette optique, être considérée uniquement pour ce qu'elle représente, ce changement de regard porté sur l'objet entraîne un changement au niveau du principe de l'indexation des images fixes.

Ces nouvelles données intrinsèques aux images rappellent que la photographie est avant toute image révélée un processus technique s'exprimant par des choix de prise de vue, ces données permettent ainsi de spécifier une caractéristique fondamentale de la

⁹⁹ Patrick Peccatte, *loc.cit.*

¹⁰⁰ *Ibid.*, p.3.

photographie: le cadrage, la découpe de la réalité enregistrée comme une image. Comme le démontre André Gunthert, il convient d'en finir avec « le fétichisme de la continuité de matière entre les choses et les images sur quoi reposerait la photographie argentique. Même dans le cas d'un signal analogique, pas plus le dispositif optique que le support d'enregistrement ne constituent des intermédiaires neutres d'un flux émanant de l'objet. Au contraire, les photographes savent bien qu'en changeant d'objectif ou de film, ils disposent au moment de la prise de vue d'une marge de manœuvre importante, qui leur permet de modifier l'aspect, la géométrie ou les couleurs d'une scène. On ne saurait décrire un appareil photographique comme un médiateur transparent du réel: il doit plutôt être compris comme une machine à sélectionner des interprétations, selon un ensemble de paramètres aux interactions complexes, qui requièrent des choix précis. Un aiguillage plutôt qu'un miroir.¹⁰¹ ». Avoir accès à ces informations lève le voile jeté par la suffisance artificielle de l'image conçue comme double iconique du réel. Contre la transparence qui se veut en réalité opacité sur les conditions de production d'une image, il ressort que si une photographie est une image, l'image n'est pas toute la photographie. L'indexation des images peut alors être puisée dans les données intrinsèques aux images et les définissants.

3. Les métadonnées comme outil de classement

Le numérique facilite une approche renouvelée de l'indexation puisqu'il a permis de considérer en quoi une photographie n'est plus à proprement parler uniquement l'objet iconique mais se compose de données essentielles dans l'approche contextuelle ou technique de l'image. Si la photographie numérique est un fichier, elle contient comme nous l'avons vu dans la partie précédente, une série de métadonnées pouvant offrir des informations et être utilisées comme voie possible d'indexation. Ces données qui entourent l'image sont consultables au niveau du fichier EXIF de la photographie. EXIF est une abréviation de *EXchangeable Image File*. Ce format définit les informations d'ordre technique contenues dans les fichiers image. Comme les champs IPTC, ce sont donc des métadonnées de type interne. Ces informations renseignent le plus souvent sur les paramètres de prise de vue, distance focale et ouverture, mais les champs date et heure peuvent également être mentionnés, si le matériel le permet, le lieu ou plus précisément les coordonnées géographiques sont visibles. L'ensemble de ces éléments inscrivent ainsi l'image dans un contexte spatio-temporel.

¹⁰¹ André Gunthert, « L'empreinte digitale. Théorie et pratique de la photographie à l'ère numérique », *ARHV*, 3 octobre 2007, en ligne <http://www.arhv.lhivic.org/index.php/2007/10/03/506-l-empreinte-digitale> .

Ces paramètres peuvent servir de nouveaux moyens d'indexation des images, en offrant par exemple une organisation temporelle des collections. Ce cas s'avère surtout pertinent dans un contexte privé de classement personnel de ses images, dans un cadre d'échange, de mise en ligne, les dates spécifiques, à moins qu'elles ne soient reliées à un évènement mémorable collectivement n'offre pas une bonne entrée d'indexation. Certains sites de partage offrent des modes de classement et de recherche par moment, Zoomr¹⁰² propose une entrée dans ses collections à partir d'un choix temporel entre les téléchargements du dernier mois, semaine ou encore jour, ces types de segmentations temporelles étant les seules universellement valables et, faisant sens au plus grand nombre. Les métadonnées de lieu peuvent au contraire aboutir à une pertinence tant au niveau de l'indexation que de la recherche de documents. Se référant aux études de Nancy Van House¹⁰³ la localisation géographique apparaît, dans un contexte domestique et de pratique privée de la photographie, comme une information naturellement vouée au classement des photos. Il ressort des enquêtes de la chercheuse une certaine aisance de la part des particuliers avec ce mode « d'indexation » des images. Si l'on considère les sites de partage des collections et de recherche d'images l'indexation par lieu de prise de vue offre une belle ergonomie, notamment grâce à l'introduction d'une carte qui se fait index amenant la recherche sur les voies de la navigation.

Certes il ne s'agit pas là d'un procédé révolutionnaire mais d'un moyen de classement des images, existant déjà à l'époque de la photographie analogique et connaissant aujourd'hui dans son application au numérique et au Web de nettes évolutions. La première étant la possibilité d'un placement automatique des images sur la carte alors qu'un tel classement faisait autrefois forcément appel à la mémoire des sujets pour être valide, les données contenues dans l'image remplace la recherche d'informations, ou de remémoration. La photographie argentique classée par nom lieu est le plus souvent rattachée à une date particulière ou s'inscrit dans la durée, un évènement familial, etc. Il est rare de classer par le lieu pour le lieu dans le cas des collections argentiques personnelles, le géotagging au contraire agit par extraction de lieu indépendamment de toute date, ou évènement. Les images sont donc spatialisées parfois automatiquement mais le lien temporel ne suit pas. Enfin on note une nouvelle différence au niveau du sujet, du point géographique indexé, le lieu pris en compte dans le cas des métadonnées géographiques est

¹⁰² <http://www.zoomr.com/>

¹⁰³ Cf Nancy Van House et al, *From 'What?' to 'Why?': The Social Uses of Personal Photos*, School of Information Management and Systems, University of California, Berkeley, 2004, (en ligne http://people.ischool.berkeley.edu/~vanhouse/van%20house_et_al_2004a.pdf)

celui de la prise de vue, au contraire le sujet indexant manuellement ses images et les classant par ère géographique le fera en privilégiant le nom du paysage, du site visible.

B. La géolocalisation

1. Les principes

Le terme de Géolocalisation est récent et correspond à l'ensemble des mesures techniques et des coordonnées géographiques qui composent une information, un document numérique localisé. Il était auparavant question de Géomatique pour définir l'ensemble des procédures de traitement des données géographiques par ordinateur.

Géolocaliser c'est adjoindre des métadonnées géographiques aux ressources numériques (sites Web, photographies, vidéos, etc.), si cela consiste initialement à préciser les coordonnées de type longitude et latitude, il est également possible de préciser l'altitude ou les noms de lieux on parle dans ce dernier cas plus spécifiquement de Géotagging (il est à noter que cette opération accompagne presque toujours celle de géolocalisation). Cette pratique s'est largement popularisée par le développement et la diffusion grand public des technologies GPS et par les avancées de la cartographie satellitaire sur le Web. Aujourd'hui la géolocalisation englobe de larges usages permettant à la fois de situer (repérer, cartographier), d'accéder (donner, tracer des itinéraires) et de s'informer sur ce qui est à proximité. La géolocalisation devient effective par son couplage à une opération de cartographie, avec les nouvelles technologies GPS et d'Internet toutes sortes de cartes sont créées et intègrent des contenus géolocalisés.

Concernant les images fixes, les informations nécessaires à leur géolocalisation se trouvent dans les métadonnées EXIF (Exchangeable Image File Format) de la photographie. Le développement des appareils intégrant un système GPS facilite l'indexation automatique des images en fonction du lieu de prise de vue. Néanmoins les téléphones portables s'avèrent en avance sur les appareils photos pour ce qui est de la géolocalisation. À l'heure où les compacts numériques avec GPS intégré commencent à voir le jour, les téléphones portables de ce type sont pour leur part bien plus nombreux et aguerris à l'exercice consistant à inscrire en direct des informations GPS dans les Exif d'une image.

À défaut de pouvoir disposer d'un appareil photo numérique avec GPS intégré, il reste possible de se tourner vers un récepteur GPS externe. Cet appareil agit comme un véritable journal de bord de tous les déplacements, enregistrant les coordonnées GPS selon

un intervalle de temps souhaité, reste ensuite à synchroniser ces informations enregistrées avec les informations Exif (date et heure) contenues dans les images. C'est la confrontation de ces deux sources qui permet de récréer l'itinéraire sur une carte et d'y associer automatiquement des photos.

Si aucun récepteur GPS n'est utilisé dès la prise de vue, les données de type latitude et longitude ne seront pas renseignées dans les métadonnées des images. Le Géopositionnement demeure néanmoins possible mais s'effectue manuellement, directement sur les sites de partage d'images ou sur les applications dédiées comme nous le verrons plus loin.

2. Les photographies géolocalisées

Dès le début des années 2000, plusieurs laboratoires de recherche pour la plupart américains, ont pour but d'élaborer des plateformes de partages d'images où le lieu géographique serait à la fois un critère de classement des collections et une voie d'accès aux photographies. En 2003, le projet *World Wide Media eXchange* (WWMX) mené par Kentaro Toyama¹⁰⁴ pour Microsoft, est l'un des premiers à proposer une organisation des collections photographiques selon leurs coordonnées spatiales et à les placer sur une interface cartographique. Pointant du doigt les lacunes de la recherche d'images par mots-clés, les concepteurs arguent de la simplicité et des possibilités d'indexation automatique permises par le traitement des informations géographiques. Le navigateur du WWMX propose ainsi au public de naviguer parmi des collections d'images à la fois selon le lieu mais aussi le temps grâce à une interface cartographique et chronologique, il est ainsi prévu de permettre à l'utilisateur de retrouver des images d'un lieu précis mais à différents moments de la journée voire saisons. PhotoCompas¹⁰⁵, une interface élaborée par Mor Naaman, reprend ces mêmes caractéristiques et ajoute la possibilité de découvrir des images géographiquement proche grâce à des outils de sélection de la direction cardinale, en cliquant sur les icônes de Nord, Sud, Sud-Est, etc., l'utilisateur a donc la possibilité de naviguer parmi les images environnant celle sélectionnée de prime abord. Nous pourrions également citer des projets plus récents comme Photo Tourism (sur lequel nous reviendrons dans le dernier point du chapitre) ou encore *The Big Picture*, de Carlo Torniai, dont le but, s'il semble proche de WWMX, prétend également employer un nouveau genre

¹⁰⁴ Cf. Kentaro Toyama et alt., *Geographic Location Tags on Digital Images*, in Proc. ACM Conference on Multimedia, October, 2003.

¹⁰⁵ Cf. <http://infolab.stanford.edu/~mor/worklinks.html>

de métadonnées standardisé et spécifique au géotagging, permettant de conserver des informations géographiques plus précises que les seules coordonnées EXIF, le tout dans la perspective d'une application sur le Web. Tous ces différents travaux ont pour objectif de permettre aux images d'une base de données d'entrer en résonance les unes avec les autres grâce à leurs coordonnées géographiques, et ainsi créer une expérience enrichie de la recherche d'images et de leur visionnage.

Appliquée à la sphère du Web, la géolocalisation est devenue une méthode d'indexation originale et une pratique répandue dans le cadre toutefois précis des sites de partage d'images. Outre un moyen sûr de se rappeler où une image fut prise le géopositionnement des images offre un système de classement, dans le meilleur des cas automatique, et une facilité de la recherche grâce aux sites ayant développé des interfaces cartographiques associées à un moteur de recherche. Depuis le début des années 2000 le nombre de personnes utilisant la géolocalisation comme système d'indexation, de classement et de marquage des images s'est considérablement accru. Comme le mentionne Carlo Torniai: « le résultat d'une recherche d'images ayant pour tag « geo:lat » sur Flickr retourne 16048 résultats entre octobre 2003 et octobre 2004, 89514 pour l'année suivante et 171574 pour la période allant de Octobre 2005 à Octobre 2006¹⁰⁶ ».

Nombreux sont aujourd'hui les sites offrant aux internautes la possibilité d'ajouter des métadonnées géographiques à leurs images utilisées pour le classement, qui est plus un positionnement, et la recherche qui devient alors navigation sur une interface cartographique. Les sites classiques de mise en ligne et en commun des collections d'images personnelles, par l'introduction d'un module cartographique de type Google le plus souvent deviennent ainsi des exemples courants de Mashups opérants.

Toutefois représenter des photographies sur une interface cartographique ne va pas sans poser quelques problèmes d'ergonomie, en particulier dans le cas d'une base de données amplement fournie. En effet comment faire figurer sur une carte même numérique, des centaines voire milliers d'images d'un seul et même lieu ? Cette question fut étudiée très tôt dans l'histoire du géotagging, au début des années 2000 des chercheurs de l'université de Berkeley en Californie réfléchissent à la création d'un site où les images seraient indexées, classées et présentées par leur lieu de prise de vue. Dans l'étude de Kentaro Toyama¹⁰⁷ apparaît la problématique de la représentation visuelle des images sur

¹⁰⁶Carlo Torniai et alt., *Sharing, Discovering and Browsing Geotagged Pictures on the Web*, 11 mars 2009, (en ligne : http://www.hpl.hp.com/personal/Steve_Cayzer/downloads/papers/geospatial_final.pdf).

¹⁰⁷ Kentaro Toyama, *op.cit.*

une même page cartographique garantissant à celle-ci une lisibilité optimale. Plusieurs solutions sont envisagées, de la miniature photographique au symbole de localisation comme le spot jusqu'à la grille de spots dont la taille reflète le nombre d'images localisées à cet emplacement. Aujourd'hui et comme nous allons le montrer dans la partie suivante, les divers sites dédiés au géotagging ont tous mis au point leur propre ergonomie et gestion de la densité des images en un point spécifique, allant de la sélection de certaines images à leur représentation par des symboles chiffrés en référence à leur nombre par localisation. Le problème de lisibilité étant la plupart du temps résolu avec un niveau de zoom élevé permettant d'atteindre un degré de détail nécessaire à une visualisation de la localisation exacte des images.

II. Les mises en application sur le Web

A. La géolocalisation dans les sites de partage

De nombreux sites de partage de photographies proposent aujourd'hui l'option de géotagging à la fois comme moyen de classement personnel des images et comme voie d'accès aux contenus du site.

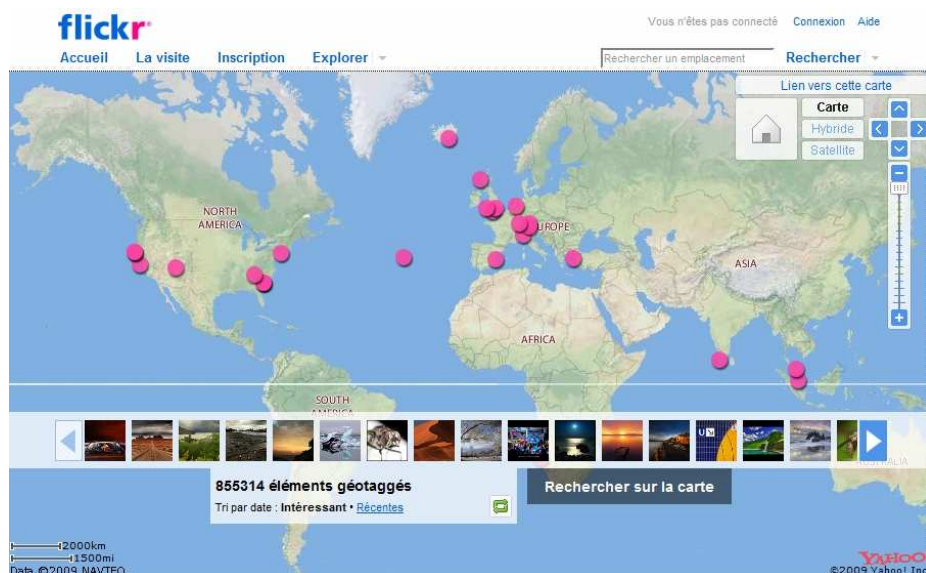
1. Flickr

Témoin de l'engouement soudain suscité par le Géotagging, Flickr¹⁰⁸ lance sa propre interface cartographique associée à un outil de géolocalisation durant l'été 2006. Le site comporte bien des portes d'accès vers ce contenu localisé, qu'il s'agisse de *la carte du monde*, de la page d'accueil du projet *Lieux* défini comme « un livre ouvert sur tous les lieux du monde » ou encore les différents champs assortis de la mention « Où avez-vous pris cette photo ? » qui invitent le visiteur à s'intéresser à ce principe et à renseigner ses images. Ce service qui vient compléter la palette déjà large de moyens mis en œuvre par le site pour explorer sa base d'images, permet aux utilisateurs de sélectionner sur une carte la région, le lieu où la photographie a été prise et d'y placer l'image, les données correspondant à la latitude et longitude de la position sont alors ajoutées aux informations conservées sur le site. Le géopositionnement de ses images intervient à la fin de la procédure classique d'upload, une indexation-positionnement automatique des images en fonction de leur localisation s'effectue en cochant l'option « Importation automatique des

¹⁰⁸ <http://www.flickr.com/>

informations GPS en tant que données géographiques ». Il est également possible d'attribuer manuellement des coordonnées géographiques à des images à partir de son gestionnaire personnel Organizr, par un simple glissé-déposé de la photographie au bon emplacement sur la carte.

La consultation des photographies stockées et géographiquement indexées sur Flickr se fait à partir de la page *Map*. Cette interface cartographique permet une requête par nom de lieu, mais aussi par simple navigation et zoom sur la carte, une barre d'images miniatures en lien avec la zone sélectionnée apparaît alors, ces images sont repérées sur la carte à l'aide de points de couleur. Il est alors possible de cliquer sur un point de la carte pour mettre en évidence l'image liée, ou au contraire de sélectionner une image dans la barre de navigation pour l'agrandir et mettre en évidence sa localisation à l'aide d'un symbole étoilé venant entourer le point.



III. 14 Capture d'écran de l'interface cartographique de Flickr

Pour toute image sélectionnée il est possible d'aller consulter le compte du photographe, ou de continuer la navigation en circulant d'images en images voisines. Les photographies géopositionnées sont regroupées en deux catégories, la fameuse sélection « Interestingness » de Flickr, regroupant les images les plus populaires du site, un onglet « recent » fait quant à lui apparaître les dernières images téléchargées en lien avec le lieu.

Toutefois géopositionner manuellement ses images sur Flickr est un exercice assez fastidieux, il en est de même de la navigation sur la carte Explore qui manque d'intuitivité. L'apparition des miniatures n'étant pas synchronisée avec le déplacement ou le niveau de zoom appliqué sur la carte mais nécessitant une réactualisation de la requête après tout

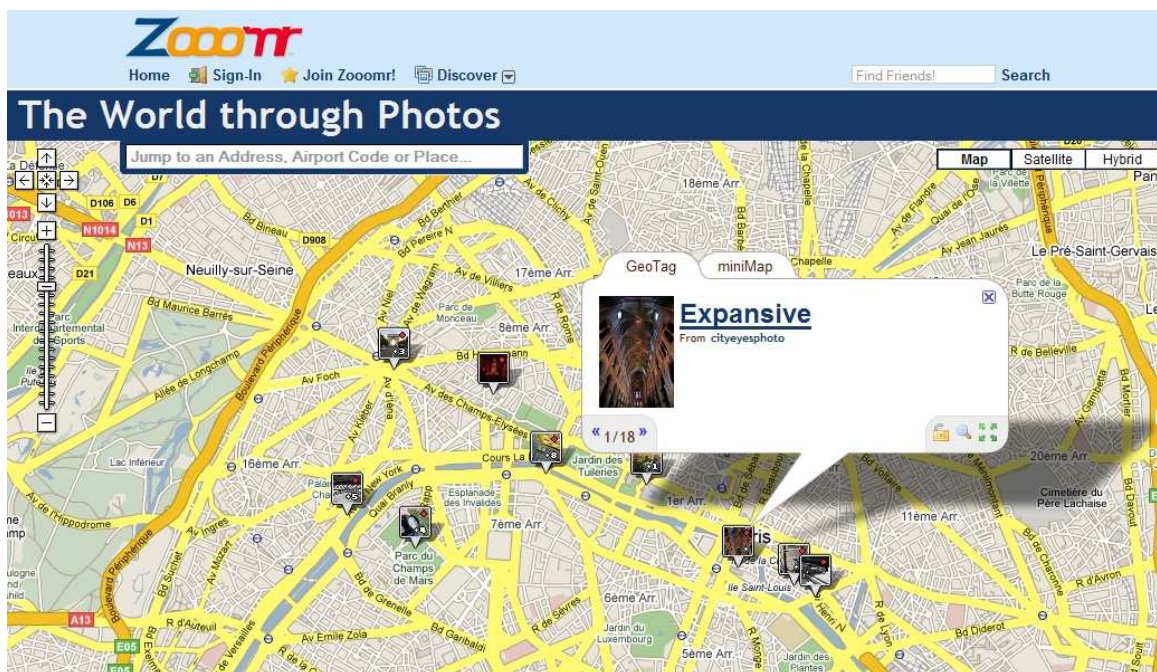
changement inhérent à la navigation cartographique. De plus le service utilise Yahoo Maps dont la précision et le degré de zoom est loin d'être suffisant pour positionner avec justesse l'image sur le lieu exact de la prise de vue. Il convient enfin de préciser que l'outil n'ajoute pas les informations, coordonnées géographiques de latitude et longitude aux données EXIF de l'image, mais les intègre seulement aux informations externalisées sur le site. Les images comportant déjà les métadonnées Exif, de type geo:long et geo:lat sont, quant à elles, automatiquement référencées sur la carte.

2. Zoomr

Zoomr¹⁰⁹ a été lancé en 2005 par Kristopher Tate. Le 19 juin 2006, Thomas Hawk, photographe et blogueur connu a été embauché par Zoomr en tant que "Chef évangéliste". En 2006 Zoomr sort "Zoomr: Mark II" introduisant les "SmartSets", l'amélioration du Geo-tagging et des améliorations dans la vitesse de chargement des pages. Si une image avec des métadonnées géographiques au format EXIF est téléchargée elle sera automatiquement placée sur la carte. Le procédé manuel est similaire à celui de Flickr à la seule différence qu'une carte Google est ici utilisée, permettant un degré de précision bien supérieur et de minimiser les erreurs ou approximations lors d'un positionnement manuel des images. Les recherches d'images s'effectuent soit par requête sémantique soit par navigation sur une carte, le système permet aussi la sélection d'une image et l'affichage des photographies prises aux alentours.

L'application de recherche par navigation sur la carte est accessible via l'onglet Discover, Zoomr n'offre pas de barre ou de mosaïque d'images en parallèle de la carte mais a mis en place un autre système de présentation des images : celles-ci apparaissent sous la seule forme de vignettes sur le fond de carte. Pour pallier le risque d'illisibilité qu'entraînerait la juxtaposition des miniatures, chaque vignette comporte un chiffre mentionnant le nombre d'images localisées au même point. Un simple clic de souris sur la vignette entraîne l'ouverture d'une fenêtre dans laquelle il est possible de faire circuler les photographies en lien avec le lieu précis. Si le principe offre une lisibilité à la carte c'est au détriment de celle des images qui ne sont regardables que par leur vignette de petite dimension, de plus une seule image est placée en illustration de la série rendant les autres invisibles à l'utilisateur.

¹⁰⁹ <http://www.zoomr.com/>



III. 15 Capture d'écran de la carte de Zoomr à l'échelle de Paris

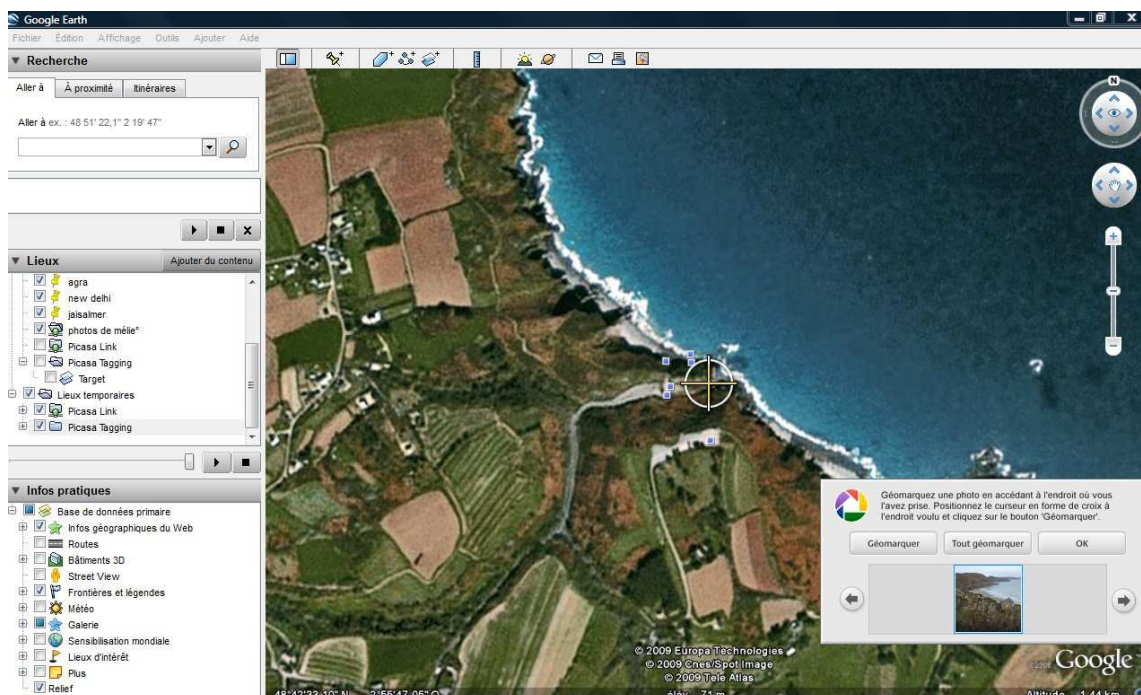
3. Picasa

Picasa n'est pas à proprement parler un site de partage d'images mais un logiciel proposant des applications de gestion, de retouche et de mise en ligne des photographies. Racheté en 2004 par Google, la société lance en 2006 *Picasa Web Album*, une nouvelle version de Picasa qui permet de publier très facilement sur le Web les photos des utilisateurs.

La fonction de géolocalisation sur Picasa¹¹⁰ se fait plus discrète que sur Flickr. Dans la barre d'icônes se trouvant sous chaque photo se trouve un symbole « Géomarquer », celui-ci commande l'ouverture d'une fenêtre GoogleEarth. Pour localiser les photos il convient alors soit de passer par le champ de recherche textuel, peut-être plus rapide mais moins intuitif que le glisser-déposer auquel Flickr fait appel, ou de déplacer manuellement le marqueur sur la carte de façon à bénéficier d'une plus grande précision. Une fois la localisation terminée, un marqueur en forme de cible se trouve associé à l'image pour signaler que l'opération est effective. Le même procédé peut être accompli à l'échelle d'un album, directement au moment de sa création, en renseignant le champ lieu de prise de vue. La fonction "Géomarquer avec Google Earth" permet d'enregistrer le lieu des prises de vue dans le dossier "Mes lieux préférés" de Google Earth. Il est alors possible

¹¹⁰ <http://picasa.google.fr/>

d'accéder aux lieux des prises de vue dans Google Earth, directement à partir de Picasa.



Ill. 16 Capture d'écran de l'outil de géotagging de Picasa sur l'interface Google Earth

Quand la photo est positionnée, les coordonnées GPS sont stockées dans les données EXIF de l'image, directement dans le fichier de la photo avec pour avantage de conserver ces informations lors par exemple de l'envoi de l'image à un tiers qui aura lui aussi accès au géo-positionnement de l'image sur Google Earth. Cette option est proposée dans le seul but que les utilisateurs profitent de leurs images dans un cadre différent, les photographies ne sont pas par la suite accessible sur GE par tout internaute.

B. Des exemples de sites dédiés

En parallèle des sites communautaires ou d'hébergement des images, des sites sont apparus ayant comme principale fonction d'offrir un système de géolocalisation efficace aux internautes. Si une dimension sociale cohabite le plus souvent avec la pratique du géotagging elle lui est le plus souvent totalement consacrée.

1. Panoramio

Créé en 2005 par deux informaticiens espagnols, se site fut le premier à être entièrement dédié à la géolocalisation des images. Dès le départ le positionnement de photographies sur une carte Google est le principal moyen de classer, organiser et chercher

ses images. Google, intéressé par les services proposés par le site, le rachète en 2007 et décide de faire de Panoramio¹¹¹ l'outil par défaut permettant d'interfacer les photos des internautes avec les cartes de Google Earth dans un premier temps puis de Google Maps. Le site se place dès sa création parmi les premiers en termes de photos géolocalisées avec 500000 en octobre 2006, mais l'accord fait bondir le nombre d'utilisateurs et de clichés mis en ligne puisqu'on passe à un million d'images en mars 2007 pour atteindre les deux millions en juin de la même année. Ce service permet d'héberger jusqu'à 2 Go d'images, de créer une carte du monde personnelle et d'afficher dessus des photos. L'interface est épurée et le procédé de géolocalisation d'une grande simplicité.

Le site reconnaît les métadonnées géographiques de type EXIF et les situe automatiquement sur la carte, la localisation manuelle des images est également possible par le lien « Placer cette photo ». Panoramio permet alors de positionner ses photographies sur la carte au moyen d'une recherche sur un champ texte « Lieu » ou par le biais d'un glisser déposer sur la carte après sélection de la zone géographique. Il suffit de Cliquer ensuite sur la miniature de l'image pour vérification : elle s'affiche à son juste emplacement dans une fenêtre d'aperçu Google Maps. Comme pour Zoomr l'interface cartographique du site est fournie par Google.



Ill. 17 capture d'écran de la carte de Panoramio

¹¹¹ <http://www.panoramio.com/>

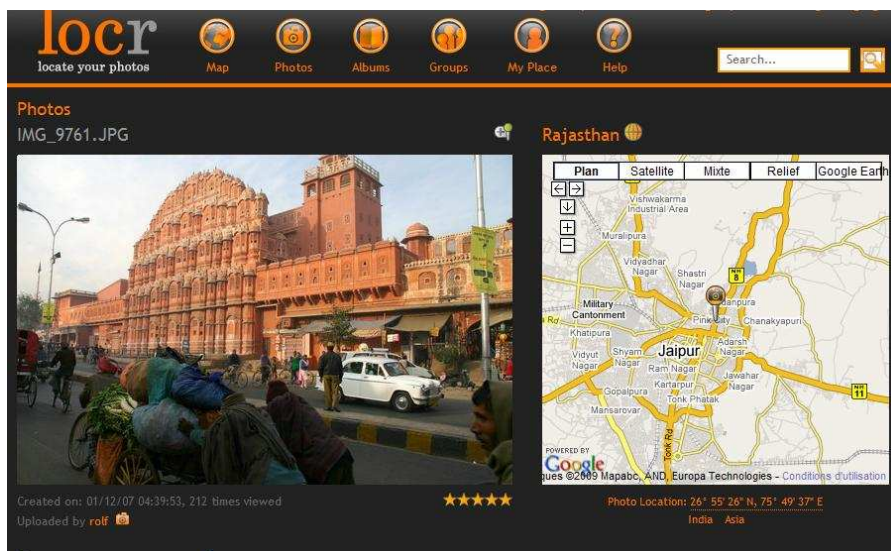
La recherche s'effectue soit par mots-clés de lieux soit en naviguant directement sur la carte par l'utilisation du zoom. La page est scindée en deux parties avec d'un côté la carte elle-même et de l'autre une mosaïque d'images en miniatures, à la différence de Flickr tout changement de direction ou de niveau de précision appliqué à sur l'espace cartographique induit une actualisation des images proposées. La sélection d'images peut donc s'effectuer aussi bien au niveau de la mosaïque que de la carte, l'utilisateur parvient alors à la page d'hébergement de la photographie en question. L'importance de la géolocalisation se révèle alors tout à fait par la présence sur la droite de l'image d'un nouvel encart cartographique situant doublement la photographie à la fois dans une fenêtre de localisation très détaillée et plus généralement à l'échelle du globe. Des images prises à proximité mais ne provenant pas forcément du même photographe sont également visibles et visualisables. Reprenant les propriétés plus classiques des sites communautaires les commentaires laissés par les autres utilisateurs apparaissent sous l'image. L'accent est bien mis ici sur le contexte et la proximité géographique des images, le nom du photographe apparaît en dessous de l'image et amène à sa page personnelle. Là encore une option de géolocalisation est proposée, voir toutes les images du photographe en question sur la carte du site où aller consulter directement celles retenues sur Google Earth.

2. Locr

Créé en 2006 par six informaticiens allemands, Locr¹¹² est un autre service d'hébergement et de partage de photos en ligne. Il se distingue notamment de Flickr par l'importance accordée à la localisation des clichés présentés et par l'intégration de contenus Wikipedia en lien avec la position géographique de l'image regardée.

Comme tous les sites les données de position géographique peuvent être entrées manuellement ou automatiquement à l'aide d'un GPS. Locr fournit à cet effet un logiciel sous Windows qui permet d'automatiser ce processus, mais à la différence de la plupart le site permet également de géotagger des photos prises par un Smartphone directement depuis ce dernier. Il n'est donc pas nécessaire de télécharger ces images sur son ordinateur mais le site est entièrement utilisable à partir de son téléphone portable. Cette option fait l'objet d'une importante promotion par les créateurs du site prônant qu'avec Locr les mobiles se font interactifs et deviennent de véritables tour opérateurs portatifs, « *Our daily lifes go mobile* » est le slogan de cette fonctionnalité.

¹¹² <http://www.locr.com/>



III. 18 Capture d'écran du site Loqr, recherche de photographies en lien avec la ville de Jaipur

Les collections de photographies sont consultables selon plusieurs modes, il est ainsi possible d'accéder aux groupes, albums personnels des utilisateurs, dernières images téléchargées mais l'option essentielle reste la carte. L'interface est très proche de celle de Panoramio – carte et mosaïque – de même lorsqu'une image est sélectionnée l'utilisateur est conduit sur la page de stockage de l'image. La photographie est là aussi représentée à côté d'un encart cartographique sous lequel sont disposées les images du site prises à proximité, si cette fonction est là aussi similaire à celle de Panoramio il est à noter la précision des informations données puisque les images ont chacune mentionné la distance et l'orientation les séparant de la photographie cible. Si les commentaires pour chaque image ne sont visibles que pour les membres du site, les informations contextuelles de Wikipedia sont par contre accessibles par tous offrant un contre point informationnel non négligeable et spécifique à Loqr.

3. Loc.alize.us

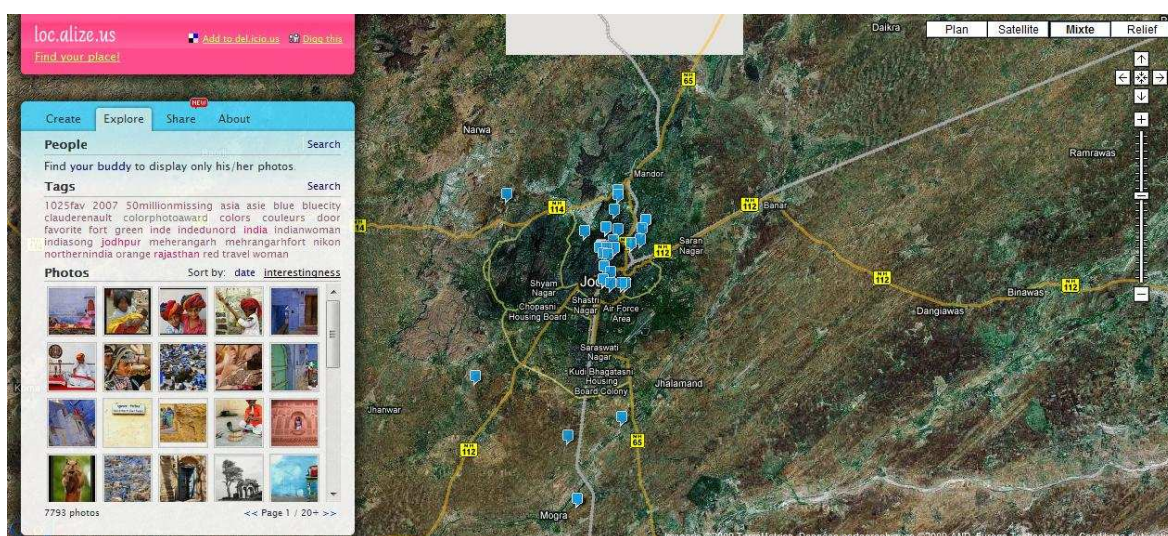
Loc.alize.us¹¹³ est un mashup utilisant à la fois l'API de Flickr et Google Maps et fut créé en 2006 par Martin Kleppe, jeune Web designer de l'agence Sumaato. Le site propose un bookmarklet, outil permettant le géotagging de ses images à partir de la lecture des fichiers EXIF. Il est également possible de télécharger directement ses photographies depuis le site Flickr et de les positionner sur l'interface cartographique de Google. Ce

¹¹³ <http://loc.alize.us/#/geo:0.000000,0.000000,2,k/>

mashup peut mettre en lien des photos et des tags de Flickr ainsi que les utilisateurs à partir de lieu géotagué.

L'interface se présente sous les mêmes traits que celles utilisant les cartes Google, avec une sélection d'une cinquantaine de photographies apparaissant à mesure que l'on zoom et se déplace sur la carte. Le positionnement sur la carte d'une image de la mosaïque est souligné par le passage de la souris, la photo sélectionnée apparaît dans une fenêtre de visionnage accompagnée de son titre, s'il en est, il est alors possible d'être dirigé vers la page Flickr de la photo correspondante.

Loc.alize.us se distingue aussi par son service de géocodage des images, permettant d'associer une image avec une adresse précise, la fonction de recherche d'adresses étant fournie par Google et GeoNames.org. Ce site relativement indépendant doit donc surtout être considéré comme une méthode efficace de visionner ses images Flickr sur une Google Maps, secondant l'option geotagging de Flickr loc.alize.us est une application plus ergonomique et efficace que le site référent.



III. 19 Capture d'écran du site de Loc.Alize.Us

III. Vers un déplacement de la sphère documentaire

A. Vers une recherche documentaire incarnée

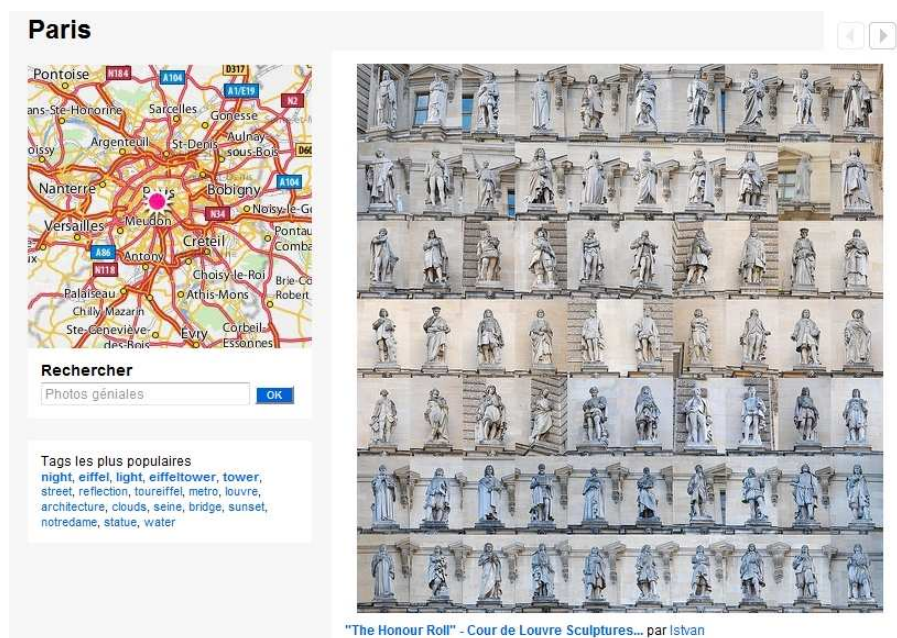
Au delà d'un moyen d'indexation d'une nouvelle forme, impliquant un nouveau rapport aux images parce que considérée dans leur contexte spatial, la pratique du géotagging implique également une approche renouvelée de la pratique même de partage d'images sur le Web en y insufflant une dimension documentaire, recherche et partage

s'agrément d'un versant informationnel, un échange d'expérience, où l'importance du vécu remplit une fonction jusque là inédite.

1. Les groupes Flickr

Conformément à la démarche qui caractérise Flickr en termes de possibilité d'accès aux collections, le géotagging en plus de la carte *Explore* aboutit à une autre application : les pages *Places*¹¹⁴. Cet outil, par rapport à la carte, se concentre sur un lieu, une ville le plus souvent, et détaille sur une même page les différents éléments et contenus en lien et visibles sur Flickr. L'interface, développée grâce au succès du géotagging sur le site, permet d'agréger des données multiples sur le même lieu, de réunir sur une même page les photographies d'un lieu. Une page Places contient un encart cartographique représentant la ville, une sélection d'images en lien jugées remarquables, des classeurs d'utilisateurs dédiés au lieu mais aussi depuis 2008 des vidéos et enfin les groupes dont le thème est centré sur le lieu. La démarche ici par rapport à *Explore* est plus documentaire et propose la découverte visuelle d'un lieu à partir des photographies géotagguées du site.

Les contenus visuels s'accompagnent souvent d'informations annexes telles que l'heure locale ou la météo.



Ill. 20 Capture d'écran de la page Flickr Places dédiée à Paris

¹¹⁴ <http://www.flickr.com/places/>

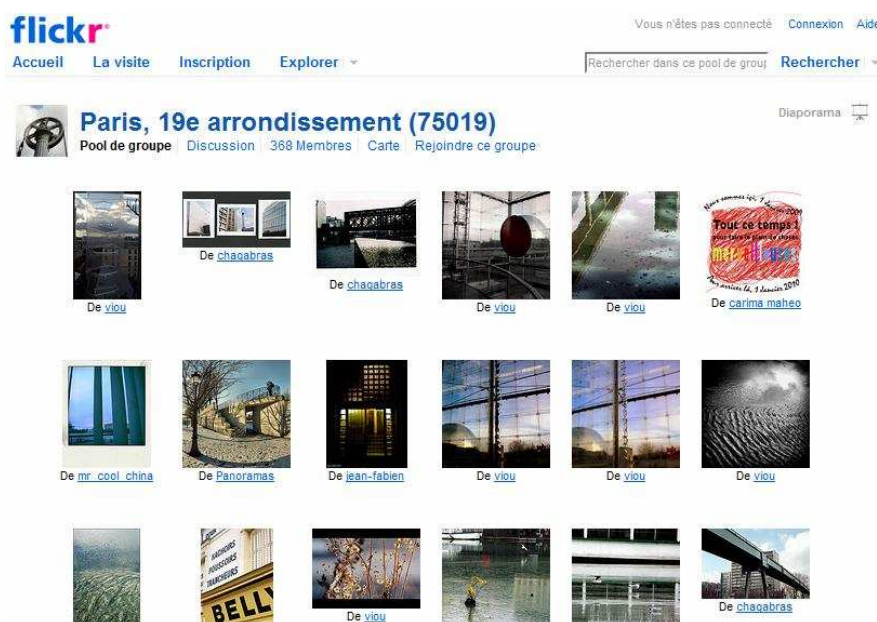
Site communautaire par excellence la pratique du géotagging sur Flickr se place, en plus de l'option *Lieux*, dans un horizon social. Si l'on effectue une recherche avec Geotagged comme intitulé au sein des Groupes Flickr, plus de 20000 résultats sont renvoyés nous présentant soit des groupes ayant directement le terme ou une acception proche pour intitulé (geotag, geotagging, etc.), soit comportant le mot dans leur texte de présentation, ou d'explication des règles. Plusieurs types de groupes sur le géotagging existent parmi lesquels les spécialistes ou férus de la technique rassemblés dans *GeoTagging Flickr*¹¹⁵. Les 4335¹¹⁶ membres recensés ont tous en commun la passion de ce système de classement et de localisation des images ; formant une communauté ils s'échangent les diverses techniques et noms de logiciel ou dernières trouvailles permettant un géotagging efficace et simplifié.

Plus nombreux sont les groupes dédiés au géotagging de photographies liées à une ville ou région spécifique comme *Geotagged Italy*, *Geotagged London*. Dans un esprit thématique seules les images prises dans l'aire géographique en question sont acceptées, l'importance étant dans le rassemblement d'images d'un même environnement et dans la possibilité de situer avec précision où la photo a été prise et donc où se situe tel monument ou paysage. Un esprit de photoclub d'un nouveau genre se retrouve dans les productions de ces collectifs, esprit qui se manifeste dans l'importance attribuée au territoire comme thème et à sa monstration, dans la volonté de donner une identité visuelle à une ville, région ou à un pays à travers la pratique photographique. Communauté de lieux, rassemblant les habitants aussi bien que les voyageurs. Il conviendrait certainement de s'intéresser au cas des groupes dédiés aux arrondissements de Paris, chacun comportant sa propre page et son réseau de membres particulier. Si la pratique du géotagging n'est pas ici centrale on retrouve l'importance du lieu et l'aspect communautaire forgé à partir de la notion de territoire. Ces groupes, qui associent tous une carte à leur pool, offrent pour la plupart une image renouvelée de la capitale, et permettent de mettre en avant une nouvelle vision des lieux. Ainsi le groupe du XIX^e arrondissement, propose une collection hétéroclite d'images rassemblées par des photographes habitant les environs ou par des touristes de passage ayant trouvé dans ce groupe une vitrine pour leur travaux, un lieu d'expression de leur vision du quartier. Ces photographies semblent pour la plupart accompagnées une expérience de marche dans la ville et dessinent en creux le portrait d'un photographe flâneur. À cette dimension de reprise en main du territoire s'ajoute pour certains groupes,

¹¹⁵ <http://www.flickr.com/groups/geotagging/>

¹¹⁶ Chiffre relevé le 23/03/09 ;

dont celui du XIXe arrondissement fait partie, une dimension communautaire locale qui apparaît au travers du forum du groupe par la présence d'informations concernant la vie sociale et les diverses manifestations culturelles se déroulant dans le quartier.



III. 21 Capture d'écran de la galerie du groupe Paris 19e arrondissement sur flickr

Les groupes dédiés au géotagging de villes, régions ou pays sont nombreux mais de petite taille, en moyenne une centaine de membres y sont inscrits pouvant néanmoins poster plusieurs milliers d'images sur le pool. Toutefois faire partie d'un groupe enclin au géotagging ne signifie pas forcément géotagger soi-même ses images. Si dans certains groupes en faveur du géotaggage des images de nombreuses photographies se retrouvent sur la carte du groupe, elles ne le sont jamais toutes, certains utilisateurs s'arrêtant à l'ajout des tags des noms géographiques. Le géotaggage véritable des images dans les groupes dédiés Flickr ne fait donc pas office de règle et son absence ne semble pas discriminante au niveau communautaire.

Si l'intérêt du géotagging se situe principalement dans la localisation de photos de lieux, de paysages et monuments, on constate que de nombreuses photographies sur Flickr sont géotagguées sans entretenir de charge iconique, de lien visuel direct avec la localisation référente, mais possédant une relation moins évidente et plus distendue au lieu. Ces pratiques se constatent par exemple dans le géotaggage de photos de nourriture, de repas.

Elles sont aussi typiques des communautés de goût qui abondent sur le site et usent du géotagging comme d'un outil de localisation et de regroupement idéal pour leur objet. Ainsi le groupe *Geotagged Street Art*¹¹⁷, qui rassemblent les amoureux des graffitis et autres œuvres graphiques urbaines demandent à ses membres de géotagger leurs images. C'est à l'émergence d'une nouvelle forme de carte que l'on assiste alors, où les objets situés se substituent peu à peu à la représentation géographique même pour devenir éléments d'une carte relationnelle et communautaire. De nombreuses pages sont ainsi dédiées à cette thématique et pour certains administrateurs de groupes le géotagging offre un nouveau moyen de connaissance des objets, une clé de lecture des images en conférant naturellement une localisation et donc un contexte spatial aux photos.

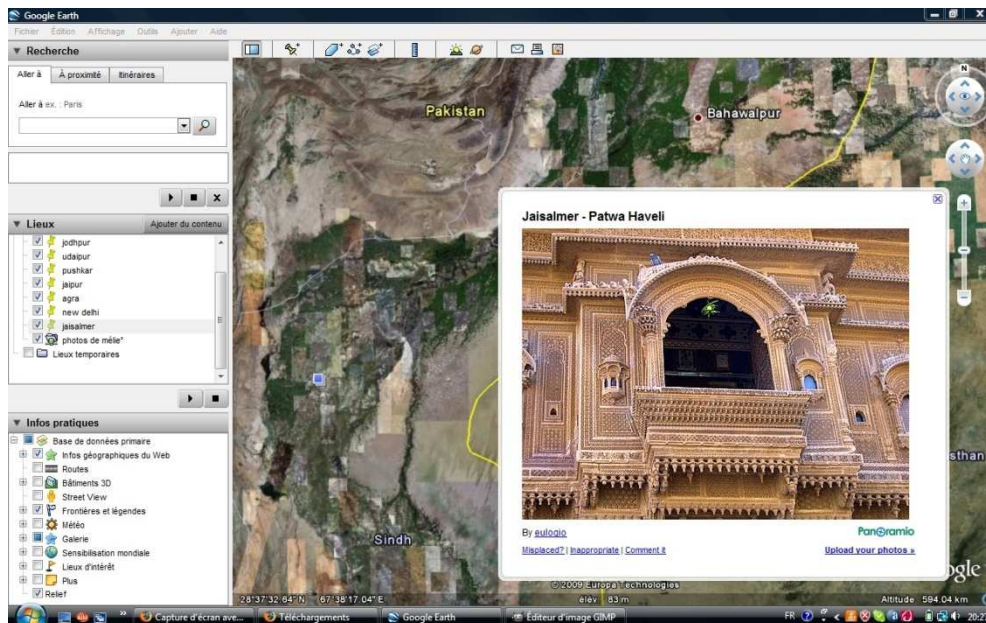
2. Panoramio ou le nouvel ordre documentaire

Panoramio est, comparativement à Flickr, moins axé sur le versant social de l'image, s'il est possible de commenter les photographies d'autrui, de se créer des contacts, aucune application n'existe pour rassembler les usagers autour de thématique commune, pour créer des groupes. Le site s'est développé comme une interface de partage spécifique, capable d'offrir une nouvelle approche du Monde, permettant de l'explorer de façon visuelle. Le géotagging est central dans l'utilisation de Panoramio, opération directement liée au téléchargement des images elle permet une déclinaison de services aussi bien pour le photographe possédant un compte, avec notamment la création de cartes personnelles avec itinéraire tracé, que pour l'internaute y naviguant. Le partage d'images se fait principalement, et de part l'ergonomie même du site et les applications proposées, autour de la notion de territoire, du contexte de prise de vue.

Les diverses discussions développées sur le forum français du site, l'abondance de post diffusant conseils et réclamations quant au refus de certaines images sur Google Earth, commentant l'acceptation de certaines images témoignent bien de l'engouement des usagers pour la liaison Panoramio-Google Earth. Sorte de coulisses, de backoffice de Google Earth, Panoramio engendre dès lors des pratiques non plus seulement sociale quant au partage d'images mais véritablement de documentation visuelle. Pour de nombreux utilisateurs du site poster ses images sur Panoramio consiste à faire œuvre documentaire. Comme l'explique un internaute sur le forum français du site : « *Ma passion de l'architecture m'a amené à proposer de nombreuses photos d'intérieur d'églises que j'ai*

¹¹⁷ <http://www.flickr.com/groups/geotagged-streetart/>

toujours choisies autant pour leur intérêt documentaire qu'esthétique »¹¹⁸, et le site d'apparaître comme un atlas photographique en devenir, aiguisant le désir de partager ses images non plus dans l'espace restreint d'un site mais à l'échelle de cette application grand public et ouverte à tous. Si toutes les images de Panoramio sont potentiellement susceptibles de figurer sur GE, de nombreuses sont refusées, décisions souvent difficile à comprendre pour les internautes dont le but était de faire figurer leurs images sur le site. Un exemple de ces déceptions courantes, « Bon ben décidément c'est très space la politique d'acceptation...J'ai fait mon deuil et accepté la triste vérité : seuls les plans larges de l'intérieur des églises sont acceptés... Sauf qu'un vitrail (d'une série de 3) vient d'être accepté et qu'à côté de ça les plans larges de la même cathédrale ont été refusés, de même qu'une photo en extérieur ...Je veux bien qu'on échafaude des théories mais j'ai tout de même l'impression que c'est du grand n'importe quoi. »¹¹⁹



III. 22 Capture d'écran d'une photographie de Panoramio sur Google Earth

Loin de ne constituer qu'un désir vain de visibilité cette volonté de figurer sur la carte mondiale de Google est souvent chargée d'un fort sentiment de légitimité, une considération positive de ses images comme parfaitement représentative des lieux. L'on constate à la lecture des diverses discussions que le projet même de GE est pris très au sérieux par les usagers de Panoramio, beaucoup regrettant le manque de qualité de certaines images acceptées, d'autres considérant l'importance de faire figurer des images

¹¹⁸ <http://www.panoramio.com/forum/viewtopic.php?p=89417&highlight=#89417>, en ligne le 23/03/09

¹¹⁹ <http://www.panoramio.com/forum/viewtopic.php?t=8527&postdays=0&postorder=asc&start=15>, en ligne le 23/03/09

illustrant les sites classés, monuments historiques et patrimoniaux, « Je n'ai pas posté ces photos en tant qu'œuvres religieuses mais en tant que photos du patrimoine national (Œuvres répertoriées à l'inventaire des monuments historiques). Je dois avouer que je ne suis pas très content quand je vois la quantité de photos médiocres tant par leur valeur intrinsèque que par la qualité de la prise de vue. »¹²⁰

Le but de l'association entre les deux sites est d'offrir aux internautes un contre point aux images satellitaires par l'ajout de ces vignettes photographiques proposant une vision spécifique de la zone. L'image sélectionnée dans Panoramio résonne dans l'usage de GE comme une incarnation du lieu. Résolument Web 2.0 puisque faisant appel aux internautes pour générer directement du contenu, la sélection des images pour GE est néanmoins rigoureuse et soumise à une charte¹²¹ d'appréciation des photographies selon laquelle seules les images prises en extérieur, celles de monuments, de paysages, de rues, bâtiments ou espaces verts sont acceptées. Il existe bien sûr de nombreuses exceptions comme pour la présence de personnes à l'image acceptée si les gens « sont une partie inévitable du lieu », ou pour les intérieurs de monuments qui doivent correspondre à « une large perspective », etc. Ces règles de publication des photographies sur Google Earth engendrent une modification automatique des images au sein même de Panoramio où les usagers désirent retrouver leurs images sur l'interface de Google vont privilégier le transfert d'images en adéquation avec la charte.

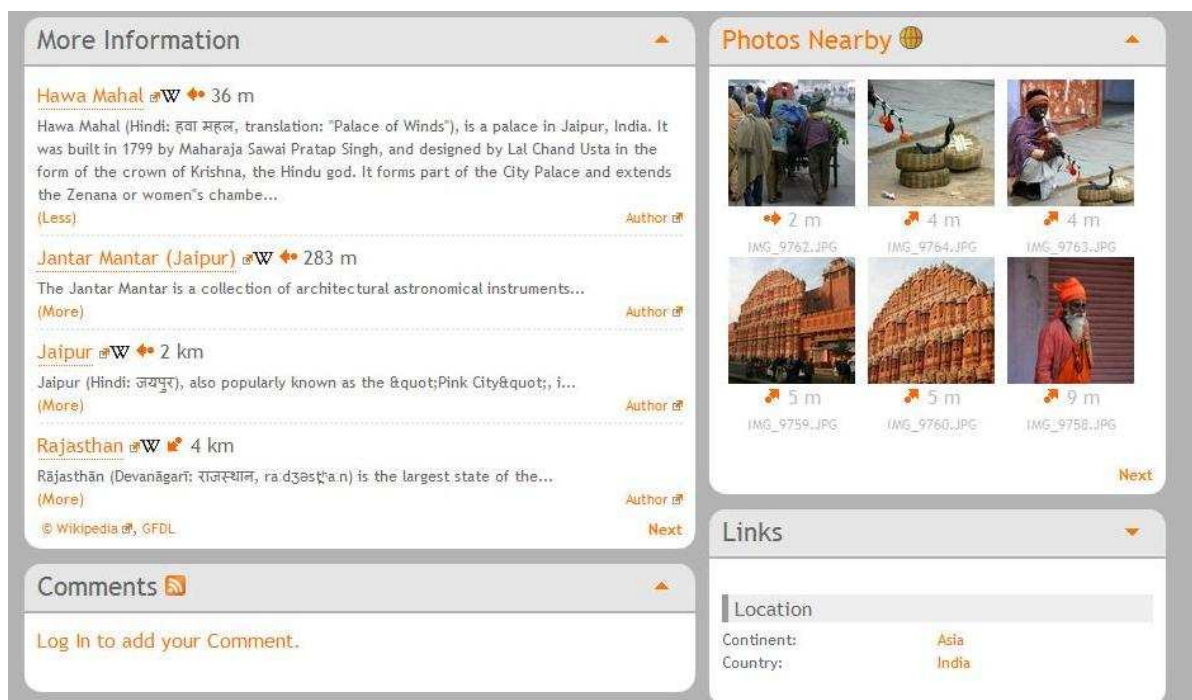
3. Echanger son expérience en plus de ses images

À côté du tout communautaire de Flickr et de l'importance primordiale du géotagging sur Panoramio d'autres sites ont su développer une approche spécifique de cette forme d'indexation, approche se basant sur la notion de partage d'expérience à partir de l'image toujours géolocalisée. Ces sites opèrent une prise en compte de la spécificité de l'image géotaguée, à savoir la prégnance des lieux, le fait que la majorité des images géopositionnées par les usagers sur une interface cartographique représentent un territoire, un monument et se trouvent naturellement en lien avec une expérience personnelle. Les photographies étant souvent prises dans un contexte de voyage, de vacances ou de contact quotidien avec l'endroit, ces sites considèrent l'importance d'offrir aux internautes à la fois de nouvelles formes d'albums de vacances et la création de sites de préparation au voyage inédits, le tout résolument Web 2.0 puisque s'appuyant sur les connaissances et

¹²⁰ <http://www.panoramio.com/forum/viewtopic.php?p=89417&highlight=#89417>, en ligne le 23/03/09

¹²¹ Cf annexes

productions de chacun. Le partage d'expérience sera ainsi au centre du lien communautaire et l'image apparaît comme un déclencheur de sociabilité, un support à la narration partagée du vécu. Le slogan de Loqr : « Partagez vos impressions et expériences avec des gens du monde entier. Montrez vos photographies préférées et utilisez Loqr comme une plateforme de communication internationale », témoigne bien de l'aspect communautaire et social du site engendré à partir du géotagging des images. Avec la possibilité de voir les images des autres internautes prises à proximité se multiplient les possibilités de rencontrer des personnes ayant expérimenté le même voyage, visiter les mêmes monuments, rendant plus aisée pour l'observateur la recherche d'informations sur un lieu. Loqr renforce sa dimension documentaire grâce aux champs Wikipedia intégrés à la page d'hébergement des images, ces informations supplémentaires qui entourent les images offrent en plus du contexte géographique des informations historiques, culturelles ou autre.



III. 23 Capture d'écran des informations Wikipedia sur Loqr

Autre site plus résolument tourné vers une nouvelle offre de tourisme, Zoom And Go¹²² a fait de la préparation au voyage et du témoignage d'expérience son champ d'opération :

¹²² <http://www.zoomandgo.com/>

“Zoom And Go a pour but de conseiller les voyageurs sur les lieux à visiter, les choses à voir, à faire et sur les lieux où se poser, à partir de l’expérience des autres routards ou des recommandations des locaux du monde entier”. Des informations et images sont ainsi fournies par les internautes concernant les villes et sites touristiques mais aussi sur les hôtels fréquentés et notés par les usagers. Ce partage d’expérience autour du lieu, se charge d’un versant lucratif certain et laisse présager les horizons économiques potentiels de ce genre de site et plus généralement du géotagging. On conçoit les possibilités pour les diverses entreprises intéressées de viser un public non plus seulement par une contextualisation sémantique des publicités à partir des recherches des usagers mais véritablement par une contextualisation spatiale, géographique se référant directement aux besoins localisés des internautes. Et l’on suit le raisonnement de Olivier Ertzscheid et Gabriel Gallezot quand ils affirment qu’ « Au delà de l’aspect pratique induit par ces techniques de géolocalisation, ces dernières font désormais de chacun d’entre nous des citoyens non plus simplement *connectés* mais *situés* »¹²³.

B. Vers de nouveaux modes de recherche d’images

Jusqu’à présent le géotagging a mis en évidence l’émergence d’un nombre conséquent de nouveaux usages, la plupart relevant du champ documentaire, et basé sur le rapport des contenus visuels, informationnels à leur représentation géographique. Cette pratique se démarque toutefois des autres systèmes d’indexation en créant une sphère iconique dédiée, favorisant son usage, que sont les photographies de lieu, de sites géographiques, mais aussi par sa capacité à renouveler les modes de visualisation des images et les principes de navigation au sein d’une collection.

1. Les applications composites : géolocalisation et locative media

Si pour l’instant le géotagging des contenus s’est plus ou moins cantonné à des pratiques proches des réseaux sociaux, de nombreuses perspectives se laissent apprécier et un horizon économique affleure pouvant laisser envisager un essor du système. La technologie du GPS, et les applications en découlant dont fait partie le géotagging, ont favorisé une interaction entre l’espace réel et mental, la création d’un lien entre une expérience physique, celle de la circulation en un lieu, de la connaissance physique d’un

¹²³ Olivier Ertzscheid et Gabriel Gallezot, « Etude exploratoire des pratiques d’indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », *Document numérique et société*, ADBS Éditions, 2006, (en ligne : http://affordance.typepad.com/mon_Weblog/files/sdndocsoc.pdf).

espace, et le monde cognitif de la représentation, venant ainsi modifier notre rapport au territoire. La géolocalisation semble pouvoir être mise à profit pour révéler le territoire, favoriser de nouvelles formes d'expression permettant de le "qualifier" ou de l'enrichir, jusqu'à créer ce que l'on nomme désormais un « espace augmenté ».

Pour Lev Manovitch, le déploiement des médias localisés, les technologies de géo-positionnement ont contribué à étendre et densifier le territoire, c'est en ce sens qu'il convient davantage, selon lui, de parler d'espace que de réalité augmentée. Nous naviguons donc aujourd'hui dans cet espace augmenté au sein duquel se distinguerait trois niveaux d'intervention : la localisation, la représentation de cette localisation par le biais d'un média et l'usage de cette localisation.

Il nous est désormais possible de déterminer notre position géographique ou notre itinéraire, et à partir de ces données, d'accéder à des informations de proximité ou à des services contextualisés, c'est-à-dire adaptés à notre position, situation spatiale. Appliquer les principes de localisation à des données, à des sites Web, à des systèmes d'information, permet ainsi aux informations numériques d'interagir avec le monde physique. Les systèmes de géo-positionnement et les technologies qui s'y rapportent nous propulsent dans l'expérience immatérielle de la navigation entre deux espaces.

Comme l'expliquent très bien les chercheurs Ertzscheid et Gallezot en prenant comme référence Google Earth : « la navigation sur ce support documentaire (...) est de nature schizophrénique. Les espaces de la carte et du territoire semblent se confondre au nom des deux prémisses suivantes : nous sommes sur la carte, et la carte est (tout au moins numériquement) « à l'échelle » du territoire. C'est cette inscription documentaire subjectivée (et qui peut être déclinée à l'infini selon la localisation géographique depuis laquelle sont envoyées les requêtes et le type de service qu'elles visent) que stigmatise et illustre l'essor des techniques de Geotagging.¹²⁴ ». Nous n'allons plus aux cartes pour trouver où nous sommes mais nous disons aux cartes où nous avons été, où nous nous situons. Ces géo-annotations personnelles fonctionnent donc sur le « j'étais là » ou « je suis ici », un échange perpétuel d'expériences passées ou présentes en lien avec un territoire, un nouveau flux informationnel motivé par une volonté de témoignage, ou par le désir de faire partager des informations pouvant être utiles à d'autres. On observe un mouvement de subjectivation de sa propre inscription au sein du Web et de la recherche documentaire devenant échange d'expérience.

¹²⁴ Olivier Ertzscheid et Gabriel Gallezot, *op.cit.*, p.8.

Une multiplication des offres mettant en jeu à la fois le contenu fourni directement par l'utilisateur et le géopositionnement de celui-ci émerge entraînant la création d'interfaces cartographiques, nouvelles pages lambda des sites Web. Google avait pressenti ce mouvement et propose des Google Maps personnalisables et surtout publiables sur le Web, autant d'outils démocratisés de géotagging et de cartographie nous reconnectent avec nos univers personnels, nous permettant d'inscrire nos expériences personnelles dans le flux informationnel de la toile.

Concernant les technologies de géotagging, de multiples perspectives de recherche apparaissent, avec pour toile de fond le déploiement à grande échelle de la téléphonie mobile connectée, ultime outil de connexion permanente, de géo-positionnement aisé, nous projetant dans un toujours connecté, un toujours situé. Il suffit pour s'en convaincre d'observer le nombre d'applications de sites de géotagging qui se développent sur mobiles, prônant un accès permanent aux fonctionnalités de recherche, aux cartes thématiques de Google et autres. Le développement de l'Internet mobile allié aux fonctionnalités déjà existantes des téléphones portables offre une instantanéité des recherches sur le net, une accessibilité à des informations ou à des services contextualisés, le tout en fonction de notre propre expérience et par la connectivité des usagers entre eux. C'est ainsi que « La frontière entre d'un côté, un espace documentaire factuel (adresses, heures d'ouverture des restaurants par exemple) et de l'autre un espace documentaire incarné et totalement subjectivé, tombe définitivement.¹²⁵ ». De ce nouvel état de fait apparaissent des requêtes documentaires d'un genre nouveau, des requêtes transactionnelles plaçant nécessairement la pratique du géotagging dans un contexte de ressource économique, sous couvert toujours d'un échange d'expérience, d'une utilité sociale de l'outil.

On parle de plus en plus de *Locative media* (que l'on pourrait traduire par média localisé) pour désigner des systèmes d'informations numériques dotés de coordonnées géographiques permettant aux mondes physique et numérique de s'interpénétrer à leurs points de convergence. Le terme de média localisé a été initialement proposé en 2003 par Ben Russel et Marc Tuters du Locative Media Lab, comme une tentative de catégorisation des nouvelles formes d'art visant à explorer l'intersection de l'espace virtuel et de l'espace physique. Les médias localisés ont attiré l'attention d'une nouvelle génération d'artistes et de techniciens, bientôt rejoints par des chercheurs et universitaires, qui ont cherché à

¹²⁵ Ibid, *loc.cit.*

explorer l'utilisation des technologies de géolocalisation dans des interventions artistiques, des cartographies personnalisées ou des définitions de réseaux sociaux.

Les enjeux des médias localisés touchent à la fois les réseaux relationnels, le marketing, l'univers du jeu, l'urbanisme... Dans leur diversité, les médias localisés forment un ensemble d'applications qui visent à se saisir d'un territoire physique, d'un "espace public" investi et enrichi par les outils numériques. Qu'on y navigue aujourd'hui par curiosité ou par intérêt ces premières applications visent à comprendre et à interroger la façon dont se tissent les réseaux humains, dont vivent les communautés, et tentent d'offrir de nouvelles voies à l'administration et au maintien de véritables "territoires", à la fois numériques et physiques.

2. L'image comme interface de navigation

Au-delà de l'aspect social, documentaire et potentiellement économique, le géotagging offre en plus d'un système d'indexation innovant une nouvelle voie dans la recherche d'images. Première interface disponible ou le recours au texte dans la phase de recherche n'est plus automatique et peut devenir secondaire, les systèmes de recherche par carte du monde apparaissent comme le premier moteur de recherche majoritairement visuel, une recherche d'images par l'image c'est-à-dire celle de la carte. Si les sites utilisant le géotagging conservent une possibilité de requête par mots clés, ceux-ci perdent de leur primauté à l'usage. Dans le cas d'une recherche d'images précise quant à une ville ou un site géographique, entrer directement le mot clé dans la barre de recherche paraît être une solution des plus rapides, si tant est que le terme choisi soit renseigné dans la base ; nous considérons toutefois cette pratique comme inaugurale dans le temps de la requête et vite remplacée par une navigation cartographique, une circulation dans l'espace géographique représenté. À la recherche directe par mot clé se substitue une recherche fluide, de proche en proche, de lieux en lieux favorisant une sérendipité des résultats, une flânerie au sein des collections qui fait écho à celles des nouveaux photographes parcourant leur ville et témoignant de leur quotidien. Et cette découverte des collections opérée par une circulation d'images en images à travers la carte fait du géotagging l'instigateur de la navigation par proximité visuelle. Parce que la carte, représentation visuelle, se substitue à l'index, au nuage de tags dans le cas de la folksonomie, cette pratique d'indexation nous permet également d'envisager les prémisses d'une recherche d'images par l'image.

Au-delà de fournir une navigation par l'image au sein de la base, cette pratique offre une contextualisation visuelle aux images, les remplaçant dans la mention de leur

environnement géographique. Les images peuvent sembler couper de leur contexte par leur inscription au sein d'une base de données, telle Flickr qui dispose de multiples biais d'exploration des collections, individualisant et isolant chaque image au sein de catégories. Si l'interface cartographique abstrait l'image de son corpus originel, elle lui confère néanmoins son contexte spatial.

À partir de cette notion de re-contextualisation géographique des images et de navigation par proximité visuelle, des chercheurs d'horizons divers ont poussé ces nouvelles caractéristiques dans une voie encore inédite. De nombreux projets voient peu à peu le jour avec pour objectif d'offrir une reconstitution du hors champs de chaque image à l'aide d'images géographiquement voisines. Une approche renouvelée du territoire apparaît s'accompagnant d'un nouveau mode de visionnage des images dans lequel le recours à la 3D joue un rôle essentiel. C'est à un véritable brouillage entre « virtuel » et physique, entre une visualisation plane et en perspective dont il s'agit alors.

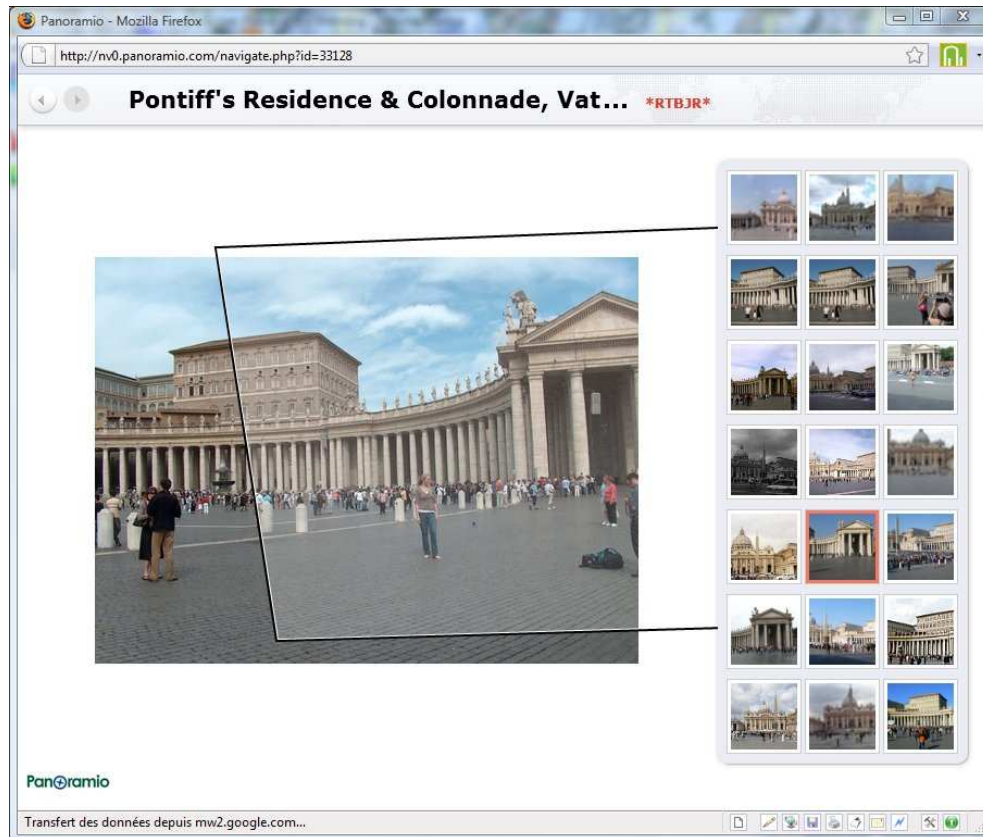
Parmi ces projets il convient de s'intéresser à l'option *Look Around*¹²⁶ de Panoramio. Apparue en 2008, cette application permet de visualiser un lieu, un monument selon plusieurs angles correspondants aux divers points de vue offerts par les photographies des membres du site. Les équipes de Panoramio ont ainsi indexé les photos qui montraient les mêmes lieux et les ont reliées entre elles en relevant les coordonnées qui leur étaient communes. Cette application pour l'instant seulement proposée pour les sites présentant une forte densité d'images comme la tour Eiffel, Notre dame, etc., permet de se déplacer de manière interactive d'une photo à l'autre en épousant, à chaque fois, le point de vue de la suivante. On passe ensuite d'une perspective à une autre par déplacement du curseur sur l'image, par sélection d'une des vues matérialisées par un cadre polygonale adoptant le point de vue de la photo correspondante. Cette application favorise la découverte dynamique d'un lieu et une nouvelle approche de la collection photographique.

PhotoTourism¹²⁷ est un autre projet qui apparaît dès 2005 dans les laboratoires de recherche de Microsoft. Soutenu par l'équipe de Noah Snavely, Photo Tourism vise à concevoir un système permettant de visualiser les sites géographiques, monuments, lieu urbain dans son ensemble en mettant bout à bout des images d'une même base de données et représentant un même objet, sujet. Il s'agit donc de jouer sur la similarité et la proximité géographique des images, de déterminer leur angle de prise de vue, la position du photographe en fonction de l'image pour permettre de relier les photographies entre elles.

¹²⁶ <http://www.panoramio.com/blog/hi-look-around/>

¹²⁷ <http://phototour.cs.washington.edu/>

Le logiciel prend appui sur les similarités présentes dans les différents clichés, les algorithmes étant chargés d'extraire une centaine de points distinctifs de chaque image. Dès lors qu'un même point distinctif est trouvé dans plusieurs images, le logiciel est en mesure de calculer sa position 3D. Ces différentes données permettraient ainsi la correspondance et la mise en lien des images entre elles.



Ill. 24 Capture d'écran de Look Around de Panoramio

Ce logiciel permettrait donc de naviguer à la fois virtuellement dans un paysage et de parcourir en même temps une collection d'images, le lieu observé n'étant accessible qu'à partir de ces images. Les usagers naviguent dans l'espace virtuel pour découvrir les images et leur représentation de l'espace en question. L'interface proposait un espace de reconstitution 3D, dans lequel les usagers peuvent se mouvoir, entouré d'une barre de visualisation des photographies composant le dit paysage et alors observables isolément. En sélectionnant une image il est alors possible d'obtenir des informations et d'accéder aux annotations d'origines de l'image.

Photo Tourism est demeuré un projet sans application complète sur Internet, ses concepteurs ont néanmoins appliqué leurs recherches à l'élaboration et à la mise en service

d'un autre logiciel de visualisation des images en 3D nommé Photosynth¹²⁸. Ouvert depuis peu au public, cette application de Microsoft fonctionne selon les mêmes principes que ceux mis en jeu par Photo Tourism, la reconstitution d'un lieu en 3D par l'agencement d'images photographiques contextualisées. Cet outil ne fonctionne toutefois et pour l'instant, qu'à partir de collections personnelles et thématiques, et ne prend pas encore en charge le croisement des corpus. Il propose en somme une façon originale de parcourir ses propres séries d'images, de mettre en relation ses photographies d'un même lieu de naviguer à travers elles par un biais uniquement visuel.



III. 25 Capture d'écran d'une reconstitution de la Tour de Pise en 3D sur Photosynth

Si le géotagging tend à se développer, ce type d'indexation des images reste encore relativement peu utilisé comparé au tagging par mots-clés notamment. L'explication du nombre restreint d'images géotagguées sur Flickr (par rapport aux milliards d'images que compte la base) et de la méconnaissance des sites dédiés par le grand public se trouve peut-être du côté du manque d'automatisation du système. Peu d'appareils numériques sont encore équipés de module GPS, obligeant les usagers du site à recourir au géopositionnement manuel, pratique considérée pour beaucoup comme fastidieuse. Elle s'avère, de plus, souvent propice aux erreurs de localisation, en particulier avec la carte Yahoo disponible sur l'interface de Flickr, trop peu précise même au niveau de détails le plus élevé.

¹²⁸ <http://photosynth.net/>

Les utilisateurs dans le contexte d'un géotaggage non automatique des images, se retrouvent souvent confrontés au problème de savoir quelle localisation ils sont tenus de tagguer, le lieu de la prise de vue et ou celui du sujet photographié. Une différence existe donc entre la pratique manuelle et celle automatisée, quand un utilisateur de plateforme de partage géopositionne manuellement ses images il semble avoir tendance à placer son image selon la situation non pas de la prise de vue mais de la localisation du monument, du lieu-dit ou remarquable. Au contraire les GPS eux enregistrent les coordonnées géographiques à l'endroit où ils se situent.

Le géotagging constitue pour l'instant une proposition de navigation par interface visuelle encore limitée malgré les possibilités lucratives qui s'esquissent. L'on peut être amené à penser que la généralisation des photophones avec GPS intégré va permettre de répandre et de développer cette pratique.

Chapitre 4

Indexation et recherche d'images par le contenu

Sommaire du chapitre

I. La recherche de l'image par l'image	110
A. La recherche par le contenu.....	110
1. Les principes	110
2. L'indexation et la requête	112
B. Les Caractéristiques symboliques	114
1. La couleur	114
2. La texture	116
3. Autres caractéristiques de l'image.....	116
II. Une observation des sites en lignes	118
A. Les recherche par l'exemple.....	118
1. La requête globale :.....	118
2. La requête partielle :	120
B. Les requêtes par esquisses	122
1. Retrievr	123
2. Gazopa	124
C. L'annotation automatique des images	125
1. Les principes	126
2. Les applications en ligne.....	127
D. Les systèmes de reconnaissance faciale	130
1. Les principes	130
2. Les applications grand public	132
3. Les applications du Web.....	135
III. Les enjeux de la recherche par le contenu	137
A. Le retour du texte ?.....	137
1. Les avantages du « bi-modal »	137
2. Conservation et partage des annotations.....	138
B. Des prospectives à visée commerciales.....	139
1. Pour quels usages dans la sphère de l'utilisateur non pro ?	140
2. La fin du piratage des images ?	142
3. L'enjeu des sites marchands	144

I. La recherche de l'image par l'image

Nous avons déjà mentionné l'intérêt actuel porté sur les bases d'images, les techniques d'indexation et celles de recherche d'images au sein de ces bases. Cependant, la problématique de la recherche d'images se résume soit en une problématique de recherche de mots en se basant majoritairement sur les attributs textuels des images tels que le nom du fichier ou des éléments de texte qu'on peut rattacher aux images, soit à l'emploi du géotagging et des interfaces cartographiques ne s'appliquant de façon pertinente qu'à une sphère iconographique spécifique. De nouvelles solutions tendent à être trouvées du côté de l'indexation et de la recherche par le contenu des images.

A. La recherche par le contenu

1. Les principes

L'archivage des images et des séquences vidéos aussi bien pour les chaînes de télévision, les journaux, les musées et mêmes pour les moteurs de recherche sur Internet ne se fait qu'au prix d'une étape d'annotation manuelle à l'aide de mots-clés. Cette indexation représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent aujourd'hui de plus en plus grandes. De plus, comme nous l'avons déjà mentionné, cette tâche est très subjective à la culture, à la connaissance et aux sentiments de chaque personne. Le besoin de méthodes d'indexation et de recherche directement basé sur le contenu de l'image n'est donc plus à démontrer. Le premier prototype de système fut proposé en 1970 et attira l'attention de la communauté des chercheurs. Les premiers systèmes d'indexation d'images par le contenu sont créés au milieu des années 90, non pas dans la perspective d'une application sur internet mais pour des bases de données fournies, spécialisées et pour la plupart fermées¹²⁹.

Les systèmes d'indexation par le contenu permettent de rechercher les images d'une base d'images en fonction de leurs caractéristiques visuelles. Ces caractéristiques, encore appelées caractéristiques de bas-niveau sont la couleur, la texture, la forme et tout autre caractéristique de l'image que l'on peut imaginer. Nous appellerons ces caractéristiques des caractéristiques symboliques. Les premiers systèmes créés pour la recherche d'images par le contenu sont du type *recherche par similarité*. La requête de

¹²⁹ Pour un aperçu des avancées et projets divers du domaine voir R.Datta, et al., *Image retrieval: Ideas, influences, and trends of the new age*, ACM Comput. Survey, Avril 2008, article 5, 60 pages.

l'utilisateur se fait sous la forme d'un exemple (croquis, image, etc.), le système réalise un calcul de similarité entre cet exemple et les images de la base puis sélectionne les meilleures d'entre elles. Ce calcul de similarité se fait à partir de *signatures* pré calculées sur l'image dont l'objectif est de résumer le contenu colorimétrique, textural ou autre de cette dernière. Dans une recherche d'images par similarité l'utilisateur fournit une image, dite image question ou requête, afin de trouver d'autres images qui lui sont similaires. La notion de similarité reste toutefois à définir mais, en général, c'est une similarité visuelle en couleur, texture et/ou forme qui est utilisée.

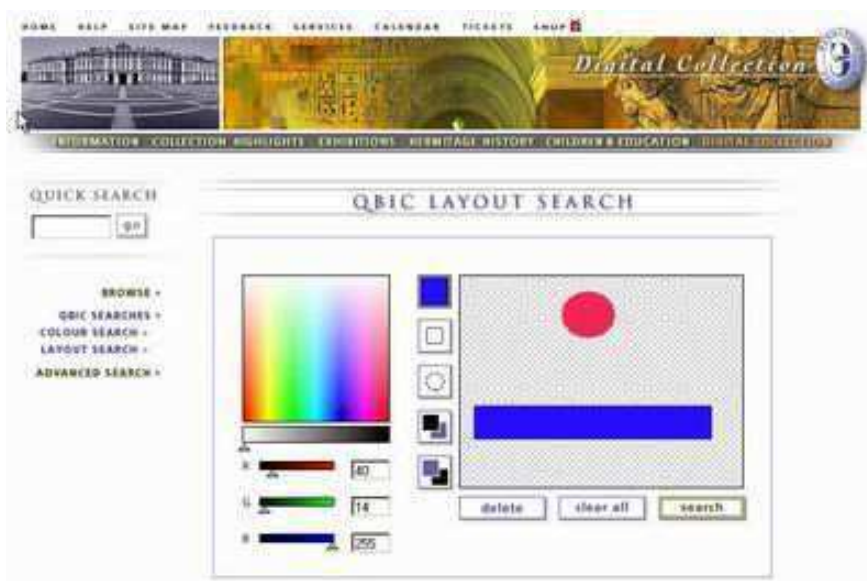
Le principal problème est alors de savoir comment choisir une (des) caractéristiques adéquates et comment les combiner. Dans cette section, nous présentons les principes des systèmes de recherche d'images par le contenu, la façon dont ceux-ci opèrent pour gérer la (les) requête(s), calculer la ressemblance et lancer les résultats. L'apparition de ce domaine de recherche, en ce qui concerne les images et les documents audiovisuels, date de la première moitié des années quatre-vingt-dix, et est donc encore récente. Comme l'indique Patrick Gros dans son article, l'émergence de cette discipline « a suscité un double mouvement, d'enthousiasme chez les chercheurs qui y ont vu un domaine nouveau d'investigation et qui, selon leur habitude, ont beaucoup promis afin d'attirer des financements pour mener leurs activités, et d'inquiétude chez certains professionnels de la documentation audiovisuelle qui y ont vu une remise en cause de leur métier, voire un danger de disparition de leur emploi. ¹³⁰».

Depuis le début des années 90, plusieurs systèmes d'indexation et de recherche d'image par le contenu (CBIR) ont été établis. Le système *QBIC* (Query by Image Content) a été développé par IBM en 1995, il permet notamment à l'utilisateur de faire la requête par le contenu visuel d'image, exemple par exemple : la couleur, la texture. Une de ses versions centrée sur l'appariement par couleur est aujourd'hui accessible sur le site du Musée de l'Hermitage¹³¹. *Virage*, qui est développé par la société Excalibur Technologies est un système de CBIR semblable à QBIC, il possède une caractéristique additionnelle permettant de combiner des requêtes en utilisant différents descripteurs : l'utilisateur peut ajuster l'importance liée aux caractéristiques visuelles selon leur signification. À la suite des travaux d'IBM sur le système QBIC, le domaine s'est développé sous le vocable

¹³⁰ Patrick Gros, « Description et indexation automatiques des documents multimédias : du fantasme à la réalité », *Interstices*, 29/05/07, en ligne : http://interstices.info/jcms/c_19256/description-et-indexation-automatiques-des-documents-multimedias-du-fantasme-a-la-realite

¹³¹ <http://www.hermitagemuseum.org/fcgi-bin/db2www/qbicColor.mac/qbic?selLang=English>

général d'indexation multimédia « par le contenu », pour le distinguer d'une indexation qualifiée de manuelle et de sémantique.



III. 26 Capture d'écran de l'interface de recherche par couleur du logiciel QBIC sur le site du musée de l'Hermitage

2. L'indexation et la requête

Ces systèmes s'exécutent en deux étapes : l'étape d'indexation et l'étape de recherche. Dans l'étape d'indexation, des caractéristiques sont automatiquement extraites à partir de l'image et stockées dans un vecteur numérique appelé descripteur visuel. Les techniques présentées ci-après, dites d'indexation, se proposent d'attacher à une image ou à une vidéo un ensemble de descripteurs de contenu, dans le but de mesurer la ressemblance avec les descripteurs correspondant à la requête. Dans l'étape de recherche, le système en fonction de la requête de l'utilisateur, lui renvoie les résultats correspondant la plupart sous forme d'une galerie d'images ordonnées en fonction de la similarité entre leur descripteur visuel et celui de l'image requête en utilisant une mesure de distance.

Il existe trois façons d'effectuer une requête dans un système d'indexation et de recherche d'images : la requête par mots clés, la requête par esquisse, la requête par exemple. Pour travailler avec des requêtes par mots clés, le système doit extraire des mots clés à partir d'une image. Le système se base alors sur l'annotation textuelle des images, réalisée manuellement, c'est aujourd'hui la méthode la plus employée, beaucoup de moteurs de recherche d'images tels que Google, Yahoo, l'utilisent. Elle est pourtant loin

d'être parfaite, quelques mots n'exprimant pas le sens d'une image, l'utilisateur désireux de trouver des images contenant un thème ou sujet précis devra filtrer le contenu des pages de résultats regroupant les photos pertinentes et celles hors de propos.

Puisque la requête par mots clés a des inconvénients, les systèmes de recherche d'images par le contenu utilisent souvent deux autres méthodes : la requête par l'exemple et celle par esquisse. Dans la première approche, soit le système choisit quelques images au hasard dans la base d'images et les montre à l'utilisateur, soit l'utilisateur parcourt la base d'images et choisit une requête ou encore donne sa propre image requête.

Dans la deuxième approche, la requête se fait par ébauche graphique (sketch), les systèmes permettent souvent de choisir entre une requête par ébauche de formes de contours ou une requête par ébauche de couleurs et textures de régions voire de combiner ces options. La forme est un élément également largement utilisé par les différents systèmes mis au point. QBIC propose ainsi une petite palette graphique : l'utilisateur trace une forme et le système retrouve les images possédant des formes semblables. Décrire une forme ne pose pas de problème majeur. En revanche la difficulté pour les systèmes informatiques est de retrouver la forme tracée dans l'image, de la délimiter. En effet, « dans une image normale, les objets se superposent : ils sont constitués de parties de couleur, de texture et de formes très différentes. La notion même d'objet est mal définie. Une voiture par exemple est un assemblage d'objets (roues, jantes, carrosserie, etc.). L'utilisateur humain pense la forme comme étant celle d'un objet (la voiture en l'occurrence), et lui associe par conséquent une valeur sémantique. Mais comment l'ordinateur peut-il savoir qu'on veut la voiture avec ses roues et ses enjoliveurs et pas seulement ces deux objets pris séparément ? Pourquoi la voiture serait-elle la forme à extraire plutôt que l'essuie-glace, qui a au moins une couleur uniforme ? Ce problème de délimitation des formes, QBIC l'a résolu en extrayant les formes de toutes les images manuellement. ¹³²». Il ressort de cet exemple qu'une étape manuelle de circonscription des formes est encore requise, s'il est envisageable d'opérer cette description sur une base de données close une telle entreprise est évidemment impossible à l'échelle d'Internet.

Des requêtes par régions sont également possibles, il s'agit alors de retrouver toutes les images de la base ayant certaines régions semblables à celles de l'image requête. Dans ce cas, soit le système segmente une image, soit on utilise une structure spatiale. Dans le premier cas, après avoir segmenté une image, l'utilisateur choisit les régions, formes

¹³² Patrick Gros, op.cit.

souhaitées à partir de l'image segmentée pour interroger la base de données. Dans le deuxième cas, la structure souvent utilisée est l'arbre quaternaire, celle-ci permettant de stocker les caractéristiques visuelles des différentes régions d'image et de filtrer les images en augmentant au fur et à mesure le niveau de détails. Cette technique est toutefois loin d'être idéale, « En ce qui concerne les régions, le problème est étudié de longue date. Les petites régions sont très instables. Celles correspondant à un objet rencontrent le problème de délimitation évoqué à propos des formes. Comment savoir quelles sont les régions les plus intéressantes dans une image ? Reste l'option de ne rechercher que quelques régions dans une image, moins d'une dizaine en général. Cela permet de séparer le centre de l'image de ce qui est soit au premier soit au dernier plan.¹³³»

Dans tous les cas, après avoir pris en compte les requêtes (images ou régions) le système va extraire un vecteur caractéristique de ces requêtes et va le comparer avec ceux des images de la base. Les requêtes par l'exemple sont les plus visuelles et également les plus souvent utilisées. Mais cette approche a un inconvénient, les systèmes supposent que plus il y a d'exemples de questions disponibles, meilleurs sont les résultats. Par conséquent, les utilisateurs pensent qu'ils peuvent avoir plus de résultats pertinents s'ils ajoutent de nouveaux exemples au résultat de la requête précédente. Cependant, cette supposition n'est pas toujours vraie, les exemples additionnels pouvant contenir des caractéristiques inappropriées et faire baisser la performance du système.

B. Les Caractéristiques symboliques

Une image contient plusieurs caractéristiques symboliques. Comme le précise Patrick Gros « En l'absence de toute indication sur l'image, on en est réduit à observer le signal bidimensionnel qui la constitue. Informatiquement, une image se présente comme une table rectangulaire d'éléments élémentaires appelés pixels, chacun codant sous la forme d'un ou de plusieurs nombres (généralement trois) l'information d'intensité lumineuse ou de couleur présente en un point. Ce sont donc ces nombres que l'on va utiliser pour décrire les images¹³⁴ ».

1. La couleur

Toujours engagé dans la course pour une indexation et une recherche des plus efficaces, Google a depuis peu mis au point un nouveau filtre de recherche concernant la

¹³³ Ibid, op.cit.

¹³⁴ Ibid. op.cit.

recherche d'images avec la possibilité de choisir la couleur dominante escomptée. Un filtre supplémentaire portant sur les couleurs est ainsi disponible depuis le mois d'avril 2009 pour la version mondiale du moteur, un déploiement rapide à l'ensemble des versions localisées de Google est annoncé. L'utilisateur peut décider d'effectuer une recherche d'images en spécifiant la couleur dominante qu'il désire. Un panel de seize couleurs est mis à disposition via un lien « All colors » disposé à droite des filtres usuels. Ce choix de la part de Google témoigne bien de l'attractivité du descripteur couleur. La couleur est en effet très souvent le premier descripteur qui est employé pour la recherche d'images, elle est le descripteur le plus accessible puisque directement codée au niveau de chaque pixel. Plusieurs travaux ont déjà prouvé que s'il s'agit d'un descripteur assez simple il demeure néanmoins efficace.

La technique classique de calcul des zones de couleur est celle de l'histogramme. Le principe est de déterminer la couleur présente dans l'image et la proportion de la surface de l'image qu'elle remplit. C'est ce qu'on appelle un histogramme de couleur, selon Wikipedia¹³⁵ « Les histogrammes de couleurs peuvent être construits dans plusieurs plages de couleurs, RVB (Rouge, Vert, Bleu) ou toute autre plage de couleurs de toute dimension. Un histogramme de couleurs est produit en découpant d'abord les couleurs de l'image dans un certain nombre de cases, puis en comptant le nombre de pixels dans chaque case ». Bien évidemment, un tel procédé de description est assez simpliste mais il est suffisant pour rechercher des images de coucher de soleil dans une collection d'images de villes ou maritimes... Cette technique est très utilisée, les histogrammes sont faciles et rapides à calculer, robustes à la manipulation des images comme la rotation et la translation.

Cependant l'utilisation d'histogrammes pour l'indexation et la recherche d'images pose quatre problèmes. Tout d'abord, ils sont de grandes tailles, donc par conséquent il est difficile de créer une indexation rapide et efficace en les utilisant tels quels. Deuxièmement, ils ne possèdent pas d'informations spatiales sur les positions des couleurs. De plus, ils sont sensibles à de petits changements de luminosité, ce qui est problématique pour comparer des images similaires, mais acquises dans des conditions différentes. Et enfin, ils sont inutilisables pour la comparaison partielle des images (objet particulier dans une image), puisque calculés globalement sur toute l'image. Plusieurs travaux récents ont été effectués pour améliorer ces problèmes. Il existe deux approches suivies par les

¹³⁵ http://fr.wikipedia.org/wiki/Histogramme_de_couleurs

chercheurs l'une des plus répandues consiste à ajouter des informations spatiales à l'histogramme en utilisant plus finement la distribution des couleurs. Patrick Gros remarque ainsi qu' « un pixel jaune entouré de bleu ne donnera pas du tout le même effet entouré de jaune. On peut ainsi pondérer l'importance de chaque pixel dans l'histogramme en fonction de son environnement immédiat, ce qui augmente grandement la puissance descriptive de l'histogramme. Par ailleurs, de nombreuses distances existent pour comparer des histogrammes et juger ainsi de la ressemblance entre les images. Le choix de cette distance influe bien entendu sur les résultats, et est un facteur déterminant de la rapidité du système. Malheureusement, les distances les plus intéressantes sont très onéreuses à calculer.¹³⁶ ».

2. La texture

L'idée est de décrire la texture de l'image, c'est-à-dire le fait qu'une image de nuage ne ressemble pas à une image d'herbe, de marbre ou de feuillage. Comme le précise Patrick Gros « Cette deuxième approche consiste donc à faire abstraction de la couleur et à s'intéresser uniquement aux distributions locales des intensités lumineuses. On va ainsi chercher à déterminer dans quelle mesure un pixel clair est plutôt entouré de pixels clairs ou foncés.¹³⁷ ». La texture est étudiée depuis une vingtaine d'années et plusieurs techniques ont été développées pour l'analyser. Certaines approches récentes semblent donner de bons résultats. Une méthode très connue pour analyser la texture est la matrice de cooccurrences de Haralick, qui calcule les niveaux de gris d'une image. Le problème est que la notion de texture n'a pas de définition formelle expliquant le nombre important de descripteurs proposés depuis les cooccurrences d'Haralick en 1973. Il est à noter toutefois que chacun d'eux a un domaine d'usage restreint, sans que ce domaine ne soit lui-même bien défini. Aussi trouve-t-on des descripteurs de texture dans de nombreux outils mais leur utilité n'est pas souvent prouvée.

3. Autres caractéristiques de l'image

Il existe plusieurs autres caractéristiques ou descripteurs de l'image qui peuvent être utilisées pour mesurer la similarité d'une image, telles que la forme, les points d'intérêt, les invariants, et plus encore. La forme est une caractéristique visuelle importante

¹³⁶ Patrick Gros, « Description et indexation automatiques des documents multimédias : du fantasme à la réalité », *Interstices*, 29/05/07, en ligne : http://interstices.info/jcms/c_19256/description-et-indexation-automatiques-des-documents-multimedias-du-fantasme-a-la-realite

¹³⁷ Ibid, Op.cit.

et elle est une des caractéristiques de base pour décrire le contenu d'une image. Cependant, la description des formes est une tâche difficile, surtout dans un contexte de « segmentation faible » et en particulier si l'on recherche la similitude et non l'exactitude. Ainsi dans les systèmes de recherche récents, peu emploient la forme comme critère de recherche. D'autres caractéristiques telles que des invariants, les points d'intérêts sont recherchés pour améliorer les performances des systèmes de recherche d'images. Dans leurs travaux Iqbal et Aggarwal¹³⁸ ont ajouté la structure de l'image à la couleur et la texture pour le système CIRES. D'autres chercheurs¹³⁹ ont présenté un système de recherche d'images basé sur la couleur, l'orientation d'objet et la position relative. Chacune de ces combinaisons a des avantages et des inconvénients et chacune peut être bien adaptée pour certaines bases d'images, mais ne fonctionnera pas correctement avec d'autres.

Les problèmes du bon choix de caractéristiques pour décrire l'image et de la pondération de ces caractéristiques ont été posés, mais ils ne sont pas encore résolus. Alors qu'un consensus minimal sur l'utilisation de la couleur et de la texture semble se dégager des travaux présentés ces dernières années, certains systèmes ajoutent d'autres caractéristiques, variables selon les systèmes et surtout selon les applications visées. Quant à la pondération des caractéristiques dans la fonction de mesure de la similarité entre les images, elle varie aussi fortement selon les systèmes, les applications et même selon les requêtes. Ces critères ne sont pas forcément d'ordre symbolique, mais plutôt d'ordre sémantique. Deux approches sont proposées pour résoudre ce problème : l'une demande à l'utilisateur de donner un ordre d'importance pour chaque caractéristique (par exemple 35% de couleur, 50% de texture...), l'autre décrète leur importance à partir des requêtes pour faire un retour de pertinence. La première approche est très difficile à effectuer car l'utilisateur n'est pas toujours capable de quantifier le poids accordé à chaque descripteur. La deuxième approche est utilisée de plus en plus dans les systèmes d'indexation et recherche des images à l'heure actuelle et donne des résultats très prometteurs.

¹³⁸ Qasim Iqbal, J.K. Aggarwal, "Combining structure, color and texture for image retrieval: a performance evaluation". *Proc. Of International Conference on Pattern Recognition (ICPR)*, 2002, Quebec (Canada).

¹³⁹ Sciascio, (1997), "Similarity Evaluation In Image Retrieval Using Simple Features". *Proc. of Storage and Retrieval for Image and Video Databases V*, 1997, San Jose (USA), vol. 3022.

II. Une observation des sites en lignes

De nombreux moteurs spécialisés dans la recherche d'images sont apparus depuis peu sur Internet. Ils ont tous développé des applications novatrices basées sur les principes de la recherche d'images par le contenu et la recherche d'images par similarité. La plupart de ces moteurs fonctionnent à partir de la base Flickr, ou de toutes autres bases de données déjà bien organisées, ce qui non seulement facilite le travail du moteur mais peut aussi faire du moteur une annexe du site source, une autre voie parmi les nombreuses déjà existantes dans le cas de Flickr d'explorer les collections.

A. Les recherche par l'exemple

1. La requête globale :

La requête globale consiste à soumettre une image exemple au moteur de recherche d'images qui devra alors prendre en compte tous les paramètres de la photographie pour fournir des images similaires.

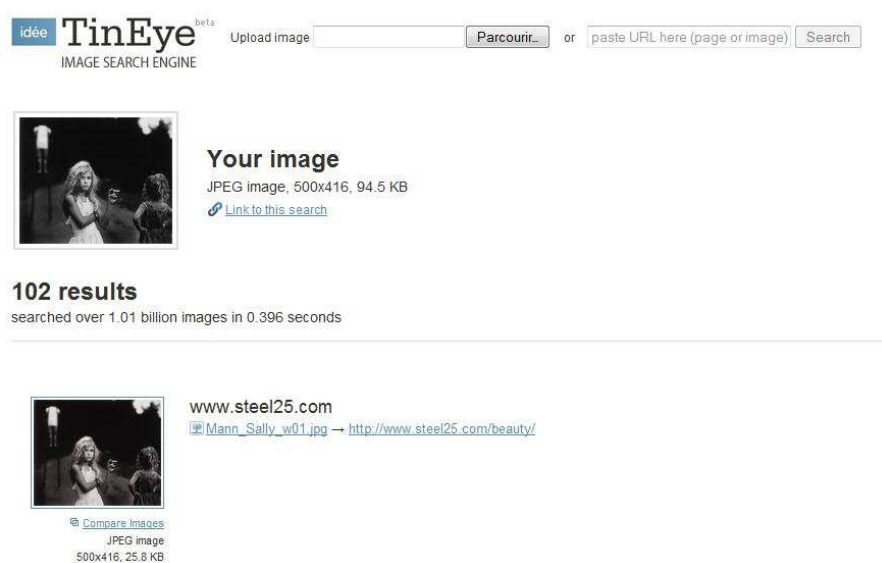
Apparu en version publique sur le Web depuis quelques mois seulement, Tineye¹⁴⁰ est un nouveau site fonctionnant sur ce mode de requête globale, la recherche se fait donc de manière purement visuelle en soumettant une image au logiciel. Cette approche renouvelle la recherche d'images sur le Web puisque le questionnement de la base de données ne se fait plus par saisie de mots clés mais par image-clé, image-exemple. Si l'on soumet une image au moteur du site, le logiciel ramènera toutes les occurrences de cette image qu'il a indexée, proposant ainsi toutes les adresses internet des pages sur lesquelles l'image est présente.

Les critères de recherche ne sont pas textuels, mais le logiciel de TinEye se base en réalité sur l'analyse des pixels d'une image pour identifier une image et ses caractéristiques, permettant au moteur de la retrouver même si elle à été fortement modifiée en taille ou altérée au niveau de couleurs. Il utilise pour cela un algorithme produisant une empreinte ou signature digitale de l'image recherchée qui par la suite sera comparée à l'index. Depuis son lancement, le moteur d'indexation du site n'a cessé d'analyser le Web pour agrandir sa base de données qui comporte aujourd'hui plusieurs centaines de millions d'items.

¹⁴⁰ <http://tineye.com/>

Dans son principe TinEye est donc plus un comparateur d'images qu'un moteur de recherche, à partir de l'image témoin la recherche s'effectue par similarité, les images trouvées seront identiques ou très proches de l'image fournie pour la recherche.

Bien que cette technique calcule effectivement la différence réelle entre deux images, elle ne prend pas du tout en compte le contenu sémantique de l'image. Elle est donc dépendante de l'angle de prise de vue de l'objet et non-robuste aux transformations classiques susceptibles d'être appliquées aux images comme les translations, symétries, rotations ou encore les changements d'échelle.



III. 27 Capture d'écran d'une recherche effectuée sur Tineye

Le système de calcul de Tineye se base donc sur l'utilisation de descripteurs locaux qui se sont révélés, dans la pratique, extrêmement performants notamment pour des tâches de détection de copie, permettant d'obtenir de faibles taux de fausse détection. Selon Patrick Gros « Cela est dû au fait qu'ils décrivent de manière très discriminante la texture locale de chaque point. Ils permettent donc de retrouver des objets précis. C'est pourquoi ils sont si adaptés pour la détection de copies ¹⁴¹ ». Il convient de noter dans le cas de Tineye que ce site ne repose pas encore sur une indexation par le contenu mais fonctionne par la création d'une signature digitale propre à chaque image et servant de référence. Si nous sommes bien dans la recherche d'images par l'image, cette technique nécessite de savoir précisément ce que l'on recherche pour être employée, de posséder une image source.

¹⁴¹ Patrick Gros, Opus cit., *Interstices*, 29/05/07, en ligne : http://interstices.info/jcms/c_19256/description-et-indexation-automatiques-des-documents-multimedias-du-fantasme-a-la-realite

2. La requête partielle :

Ce type d'interrogation porte sur une zone précise constitutive de l'image. Mais dans le cas des applications grand public cette terminologie consiste à choisir un ou plusieurs paramètres en particulier afin d'affiner la recherche sans pour autant avoir d'image au départ.

*Xcavator*¹⁴² est un moteur de recherche dédié aux images qui permet de rechercher des photographies au sein d'une base de plusieurs millions d'images comprenant également des illustrations flash, des dessins voire des vidéos. En théorie adressé aux professionnels de la recherche d'images, le site repose sur l'indexation de fonds de diverses agences d'images en ligne, il est toutefois possible d'effectuer une recherche ne prenant en compte que les images libres de droits (ce qui ne signifie pas gratuites), issues de Fotolia ou iStockphoto, et celles soumises à droits d'auteurs, la plupart des images proposées par *Xcavator* sont donc payantes.

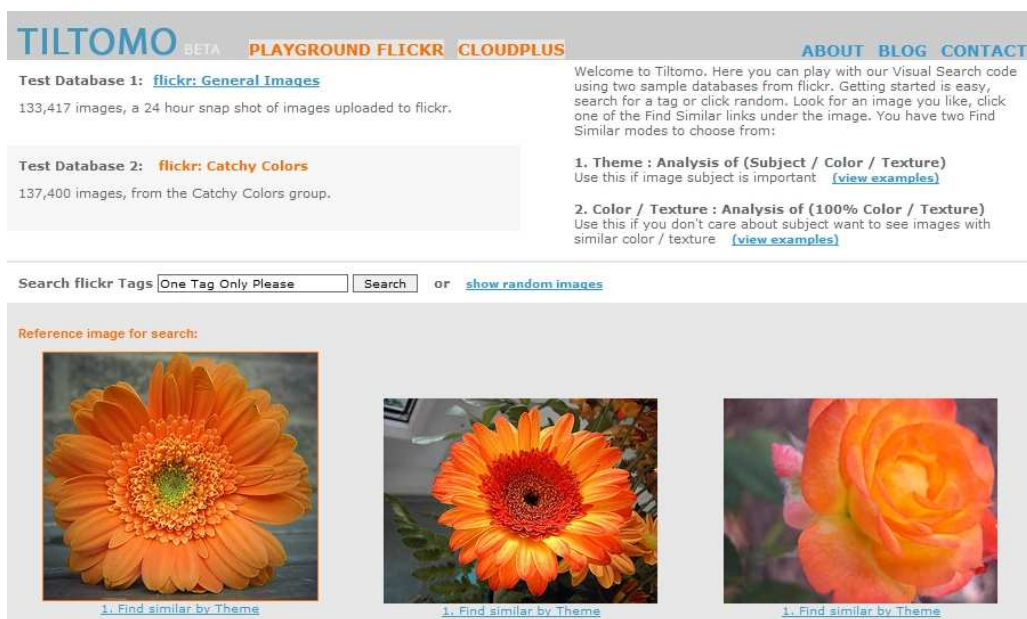


Ill. 28 Capture d'écran d'une recherche effectuée sur Xcavator

La requête débute par la sélection d'une catégorie ou l'inscription d'un mot clef dans la barre de recherche. Une galerie de dessins et de photographies est alors proposée. L'utilisateur a alors le choix d'effectuer une recherche basique par comparaison de la base avec une image sélectionnée comme exemple, ou une recherche experte dans laquelle lui est proposée une option de tri des images par couleur. Avec cette option si un utilisateur a besoin d'une photo dans des tons précis, il peut sélectionner la couleur souhaitée et *Xcavator* affiche une liste de fichiers dont les couleurs s'accrochent à la sélection.

¹⁴² <http://www.xcavator.net/>

Le site comprend également un outil assez sommaire d'esquisse de formes disponible dans les deux types de recherche, permettant à partir de l'image requête d'indiquer les formes ou lignes de force à privilégier lors de l'interrogation de la base. La petite taille de la vignette sur laquelle il nous est donné d'intervenir ne permet qu'une ébauche approximative nous faisant douter de l'utilité de cette option. Une fois qu'une image est définitivement sélectionnée elle s'affiche dans une barre sous la galerie, nommée « Lightbox » elle permet à l'utilisateur d'accéder à l'agence et/ou au site détenteur des droits de l'image.



III. 29 Capture d'écran d'une recherche effectuée sur Tiltomo à partir de la base Catchy Colors

*Tiltomo*¹⁴³ est un autre site proposant des recherches d'images par similarité et selon une sélection de descripteurs. Plusieurs caractéristiques peuvent être prises en compte, le thème, la couleur et la texture.

L'utilisateur doit d'abord interroger textuellement la base de *Flickr* pour obtenir l'image requête. Le site propose alors pour chaque image visible deux modes de requête :

- Par thème, le site propose alors une analyse du sujet des images. Les créateurs du moteur conseillent d'avoir recours à ce mode si le sujet de l'image exemple est ce qui prime dans la recherche d'images similaires. Il est également permis d'utiliser cette requête par thème couplé avec l'option "Show ONLY images from the reference photographer". En suivant ce lien seules les images du photographe de l'image requête apparaîtront dans la galerie.

¹⁴³ <http://www.tiltomo.com/>

- Par couleur et texture. Les résultats pour ce mode sont très hétéroclites, une recherche effectuée avec le tag « Tree » et la sélection d'une image-clé, aboutit à une sélection d'images de bébés, jeunes mariés, assiette repas. Sur cette page de résultats parmi une galerie de trente images seulement deux sont visuellement proches de l'image requête. *Tiltomo* est un moteur qui fonctionne seulement à partir de deux bases de données. Issues de Flickr elles demeurent indépendantes l'une de l'autre, les recherches s'effectuent donc soit :

- Sur la base générale, constituée d'une sélection de photographies issue de Flickr, et comptant environ 140 000 items. Cette base permet une grande variété dans les thèmes, sujets et les tags, et constitue donc la collection idéale pour tester la requête par thème.

- Sur la base issue du groupe « The Catchy Colors ». Elle comprend un grand nombre de photographies aux couleurs remarquables, et s'avère donc plus naturellement dévouée à la requête par couleur et texture.

Concernant ce site on regrettera la petite taille des deux bases de données ne permettant pas vraiment d'explorer en profondeur les fonctionnalités proposées.

Nous pouvons enfin mentionner un autre site permettant d'effectuer des recherches d'images par sélection de critères, il s'agit de *Picitup*¹⁴⁴. Si l'on accède en premier lieu aux images par le mode d'interrogation textuelle, par la suite une multiplicité de filtres visuels est alors proposée (visage, produit, couleur, orientation, taille...) ainsi qu'une option de recherche par similarité sous chaque image de la galerie de résultats, option qui pour l'instant ne s'avère que peu performante. *Picitup* explore les images de Yahoo!, Flickr et Picasa, et offre la possibilité de ne repérer que les images sous licence *Creative Commons* ou d'exclure le contenu commercial.

B. Les requêtes par esquisses

Ces outils permettent donc à l'utilisateur de dessiner des images requêtes. Il est bien sûr parfois difficile de faire l'esquisse d'un objet donné de façon précise et les logiciels ne sont pas toujours performants pour traiter les ressemblances entre le dessin et l'image. Les systèmes de recherche d'images par le contenu utilisent majoritairement le paradigme requête par image-clé ou combinent les deux approches ensembles.

¹⁴⁴ <http://www.picitup.com/picitup/index.jsp>

1. Retrievr

*Retrievr*¹⁴⁵ offre un moyen original d'envisager la recherche d'images sur internet. Créé par Christian Langreiter, à partir des travaux de Chuck Jacobs, Adam Finkelstein et David Salesin de l'université de Washington¹⁴⁶. Cette application effectue une recherche de similarité dans la base de données de Flickr essentiellement auprès des images « interestingness », en se basant sur les formes dessinées par l'utilisateur et les couleurs choisies. La reconnaissance est possible soit par « image », en téléchargeant une photo ou en indiquant son url, soit par « sketch » (*dessin*) en exécutant un croquis directement en ligne grâce à une application en Flash. Ce dernier mode est le plus surprenant. Il suffit de dessiner une forme et, automatiquement, s'affichent des planches de photos en rapport avec celle-ci. S'il est possible de représenter un dessin à l'aide du pinceau virtuel les esquisses peuvent également être agrémentées de couleurs.

Les résultats se révèlent la plupart du temps aléatoires et déroutants mais globalement, si l'utilisateur évite les détails inutiles et se concentre sur le choix des bonnes couleurs dans la palette, le logiciel offre quelques belles propositions d'images correspondantes. Une utilisation rapide du site permet d'observer que la recherche d'images s'effectue aussi bien en prenant en compte les couleurs que les formes dominantes, ou lignes de force. Prenons pour l'exemple un usager désirant rechercher des images de fleurs, si son premier tracé est exécuté en trait noir ne soulignant que la forme florale le logiciel ira puiser dans le stock d'images en noir et blanc de la base, et sortira les images comprenant des lignes proches de celles esquissées. Si au contraire il mentionne la couleur désirée lors de l'esquisse les images résultant de la recherche auront toutes pour teinte dominante celle utilisée.

De l'aveu même de son créateur *Retrievr* à ses limites. La première chose à avoir à l'esprit est que cette application ne fonctionne pas selon une reconnaissance exacte d'objets, de visages ou de texte, « so if you're drawing an outline sketch of a chair, it almost certainly won't get you one back (except your index only contains images of chairs)¹⁴⁷ ». Seules les formes les plus prononcées et les aplats de couleurs sont pris en compte par le logiciel. De même, s'il ne prend pas en compte les rapports de proportion

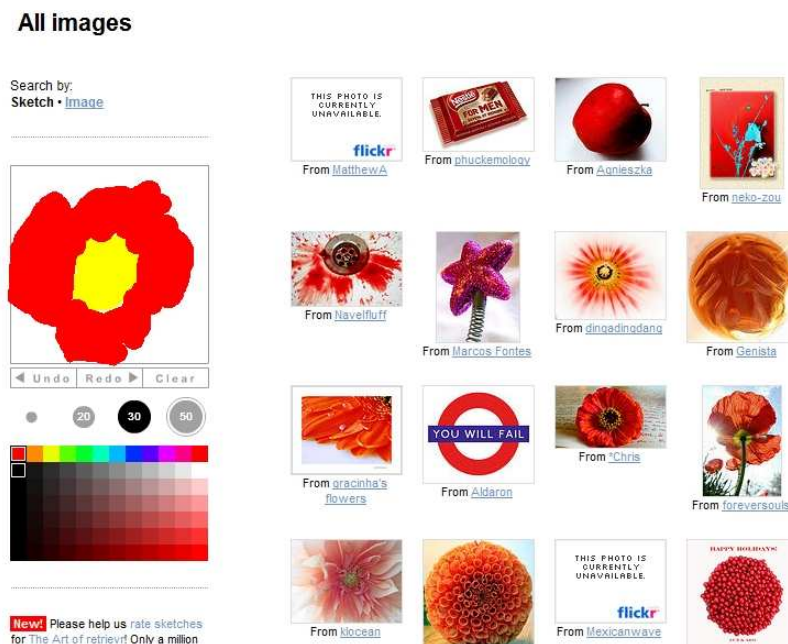
¹⁴⁵ <http://labs.systemone.at/retrievr/>

¹⁴⁶ Chuck Jacobs, Adam Finkelstein et David Salesin, *Fast Multiresolution Image Querying*, Proceedings of SIGGRAPH 95, Washington, 1995.

(en ligne <http://grail.cs.washington.edu/projects/query/mrquery.pdf>)

¹⁴⁷ Citation extraite des FAQ du site <http://labs.systemone.at/retrievr/about>

entre les éléments dessinés, il s'applique néanmoins à conserver l'agencement des formes esquissées au sein des images proposées.



III. 30 Capture d'écran d'une recherche effectuée sur Retrievr avec l'outil de requête par esquisse

2. Gazopa

Gazopa¹⁴⁸ est un autre moteur de recherche d'images similaires proposé par Hitachi et pour l'instant en version Beta privé. À côté de la requête par esquisse, le site propose de nombreuses autres fonctionnalités de recherche, l'utilisateur pouvant se servir de ses propres images uploadées sur le site, de l'adresse URL d'une image trouvée sur le Net et de mots-clés. Plusieurs types de descripteurs d'affinage des résultats sont également proposés aux usagers, tel que la couleur ou la forme, celle-ci pouvant être extraite de l'image source elle-même. Tout comme les sites vus jusqu'à présent et à l'exception de Tineye, Gazopa recherche les images par similarité, ressemblance, et non par exactitude. En ce qui concerne la recherche par esquisse, l'interface flash proposée par le site est plus développée que celle de Retrievr, malgré tout il n'est toujours pas facile de dessiner à main levée l'objet recherché, il vaut donc mieux souvent commencer par une recherche par mot-clé si l'on souhaite une requête efficace et pas forcément ludique.

L'une des particularités de Gazopa est, malgré sa version encore à l'essai, sa grande adaptabilité aux différents outils et appareils offrant accès à Internet. Il offre ainsi la

¹⁴⁸ http://www.gazopa.com/sign_in

possibilité de télécharger un plug-in permettant l'installation de l'application sur les iPhone. Grâce à l'écran tactile du mobile toutes les options de recherche sont conservées, il est donc possible d'effectuer une requête par esquisse directement sur le téléphone. Gazopa développe également une application sur Facebook, et un plug-in firefox qui permet de lancer une recherche depuis n'importe quelle image d'un site internet.



Ill. 31 Capture d'écran de la page d'accueil du moteur Gazopa

Si ce site peut être utilisé comme un moteur de recherche d'images lambda, il convient de prendre en compte sa dimension économique et ses applications commerciales, associé à différents sites de ventes en ligne, la recherche par similarité peut devenir recherche de produit et le procédé bien lucratif.

C. L'annotation automatique des images

Si la plupart des sites de recherche d'images observés jusqu'à présent disposent d'options de recherche par analyse de similarités, ils utilisent également les mots-clés comme mode d'interrogation des bases de données. On en revient ainsi au problème initial, à savoir comment permettre une indexation sémantique des images qui soient automatiques ou du moins le plus rapide possible quand les collections à traiter dépassent les millions de photos ? Comme nous l'avons déjà mentionné la seule indexation textuelle automatique des images est celle se basant sur le nom, la légende ou le texte entourant l'image, elle ne décrit donc pas le contenu de l'image et engendre un nombre important d'erreurs d'indexation car le texte pris en charge par le système ne correspond que rarement avec l'image. La plupart des sites cités plus haut ont résolu de façon partielle ce problème en

ayant recours aux collections de bases de données déjà sémantiquement indexées comme Flickr à partir desquelles fonctionne leur moteur.

De nouveaux projets se sont depuis peu développés tendant à utiliser les principes de la recherche par analyse du contenu visuel afin de générer du contenu sémantique. Il s'agit donc de décrire par le contenu des fichiers visuels pour aboutir à une automatisation de l'indexation sémantique, à des moteurs de recherche hybrides performants.

1. Les principes

L'annotation automatique d'images est un domaine du traitement d'images permettant d'associer automatiquement des mots-clés ou du texte à des images à partir de leur contenu afin de pouvoir ensuite rechercher des images par requête textuelle. L'image est d'abord analysée par l'extraction de descripteurs visuels, des mots exemples sont utilisés pour permettre au moteur d'être capable d'engendrer automatiquement des annotations à de nouvelles images selon des techniques d'apprentissage automatique. Les premières méthodes dans ce domaine se basaient sur un apprentissage des corrélations entre les descripteurs d'images et les annotations, puis des techniques utilisant des méthodes issues de la traduction automatique ont vu le jour, permettant de faire une traduction entre le vocabulaire textuel, et ce qu'on considérait comme un vocabulaire visuel, où chaque mot serait un groupe de régions similaires obtenues par regroupement (*clustering*).

Selon Christophe Millet, « l'annotation automatique d'images cherche à combler les lacunes des approches actuelles permettant la recherche d'images à partir de requête textuelle ¹⁴⁹ ». La première consiste à annoter manuellement les images, ce qui n'est plus envisageable avec le nombre croissant d'images numériques, et confronte l'indexation au problème de la subjectivité du choix du terme indexeur. La seconde approche, adoptée par les moteurs de recherche d'images sur le Web, est d'utiliser les mots de la page Web contenant l'image comme annotation de celle-ci, avec l'inconvénient de ne pas prendre du tout en compte le contenu de l'image. Si des systèmes d'annotation automatique d'images commencent à émerger, ils connaissent certaines limites, celui notamment de ne pouvoir prendre en charge qu'un nombre restreint d'images, « le nombre d'objets reconnus reste de l'ordre de 10 à 100, alors que les humains sont capables de reconnaître de l'ordre de 10

¹⁴⁹ Christophe Millet, *Annotation automatique d'images : annotation cohérente et création automatique d'une base d'apprentissage*, Thèse d'état, Telecom Paris, soutenue le 14 Janvier 2008, (en ligne <http://chr.millet.googlepages.com/these-christophe-millet.pdf>)

000 objets.¹⁵⁰ ». Un autre problème soulevé est celui de l'ambiguïté de l'image, de sa description et de son rattachement à une annotation, comment choisir parmi plusieurs hypothèses de reconnaissance d'une forme ou région donnée, celle qui permet une annotation globale de l'image sémantiquement juste. Deux sortes de désambiguïsation peuvent être développées comme le précise Christophe Millet, « la première utilise les relations spatiales, et s'assure que par exemple le ciel est toujours au-dessus de la mer. La deuxième tient compte du contexte de l'image, en utilisant la relation entre les objets et les milieux dans lesquels ils se trouvent : on a plus de chance de trouver un éléphant dans la savane, et un dauphin dans l'eau.¹⁵¹ »

2. Les applications en ligne

Si de nombreux travaux universitaires ont été menés sur l'annotation automatique des images, peu ont abouti à la création d'un moteur de recherche efficient. On retiendra ainsi le projet de chercheurs de l'université américaine de Pennsylvanie ayant développé un système de reconnaissance et de description automatisée d'images. Basé sur une méthode statistique, le projet s'intitule "Automatic Linguistic Indexing of Pictures in Real-Time" (ALIPR)¹⁵². Là encore, plutôt que de baser la recherche sur le texte environnant les photos, l'objectif est d'apprendre aux ordinateurs à reconnaître le contenu même des clichés, qu'il s'agisse d'immeubles, de paysages, de personnes etc. à partir d'un logiciel de reconnaissance des objets et des formes à l'aide de mots clés indexés manuellement.

La méthode utilisée par les chercheurs consiste dans un premier temps à référencer à la main une base de photos la plus large possible. Soixante mille photos ont déjà été taguées à l'aide de mots décrivant leur contenu. "Nous avons par exemple sélectionné une centaine de photos de parcs et les avons référencées avec des mots-clés tels qu'arbre, herbe, paysage etc." explique Jia Li, professeur à la Penn State. "Cela permet de construire un modèle statistique grâce auquel la machine apprend à reconnaître divers éléments tels que la couleur ou la texture de ces photos.¹⁵³ ».

¹⁵⁰ Ibid.


¹⁵¹ Christophe Millet, *op.cit.*, p.6.


¹⁵² <http://alipr.com/>

¹⁵³ Cf Les explications dans la section FAQ du site.

ALIPR™

Automatic Photo Tagging and Visual Image Search

Image Upload 

Or Image URL: 

http://..... - Try drag and drop from another website.

 + most-voted + emotions + pictures

or browse: animal people plant natural building city thing water flower macro art food sky female landscape ...
or random start: oak kittens point rivers vandalism queen cleaning walking hound self lebanon marble ...

Exploratory Features: **ACQUINE**- Rate Your Photos Aesthetically by Computer **NEW** <<<<<<
Story Picturing Engine - SPE | IMAGINATION captcha



australia bravo - similar >



landscape sky sunset dusk bravo - similar >



art paint gloss bottles - similar >



sport water racing dinghy - similar >

© alipr.com 2008-2009 Patent Pending. Patented. All rights reserved. Image copyright remains with the original owner. Parental discretion advised. Do NOT upload objectionable images.
About Us / In the News: Discovery, Tech Review, Sci. American, slashdot 1, 2, 3 / Research / Publication / Terms of Use

III. 32 Capture d'écran de la page d'accueil de ALIPR

Le système décrit les images à l'aide d'un vocabulaire de plus de 300 mots, pour l'instant en langue anglaise. Ces mots-clés permettent de référencer les images avec des tags spécifiques sans nécessiter l'intervention humaine. Le processus d'indexation de l'image se déroule en comparant l'image à plusieurs dizaines de milliers d'autres contenues dans une base de données. L'analyse suggère alors 15 mots-clés correspondant au mieux à l'image. James Wang, autre professeur en charge du projet déclare " avoir entraîné leur système de reconnaissance à reconnaître des concepts et des objets afin d'indexer automatiquement des images rencontrées pour la première fois ". Il ajoute que " la grande majorité des analyses donne 15 tags décrivant correctement l'image " permettant ainsi une grande précision dans la recherche. Grâce à cette technique, l'ordinateur doit par la suite être capable de reconnaître les mêmes éléments dans de nouveaux clichés. De sorte qu'une personne faisant une recherche pourra utiliser les mêmes mots-clés pour obtenir des images pertinentes.

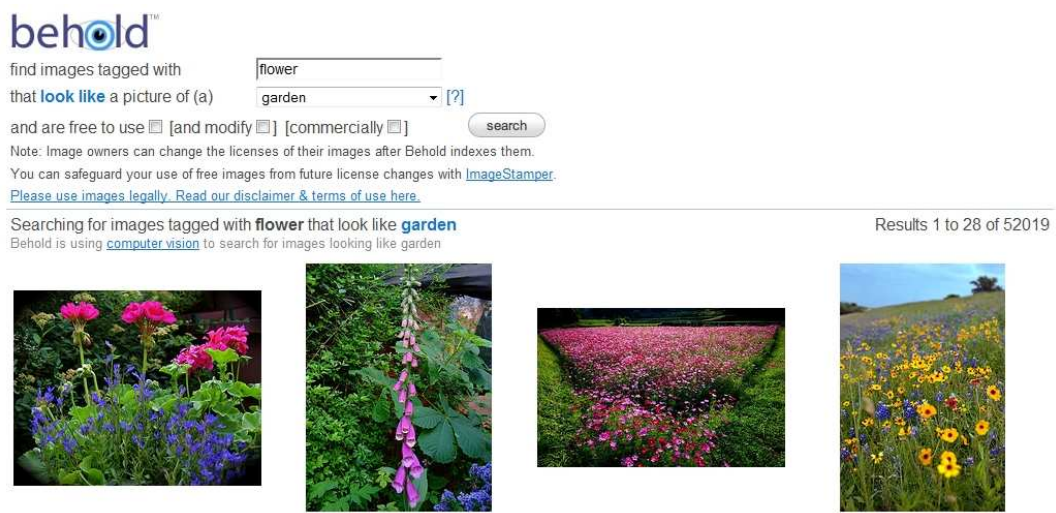
Depuis peu sur Internet ont émergé quelques sites de recherche d'images à vocation grand public et fonctionnant selon les principes et technologies d'attribution automatisée d'une indexation sémantique aux images. Nous retenons ici le site Behold¹⁵⁴, dont le principe de fonctionnement du moteur s'appuie sur un calcul des probabilités d'association

¹⁵⁴ <http://www.behold.cc/>

pertinente entre un concept et une image en analysant l'agencement des pixels de l'image en question. Chaque image analysée est ensuite rattachée à une catégorie.

Behold est l'un des sites ayant pour principe la liaison d'une indexation sémantique des images et une recherche par le contenu. La recherche d'images sur ce site se fait exclusivement à partir de requête par mots-clés, toutefois Behold ne se cantonne pas à une simple indexation textuelle des images mais est capable de reconnaître et d'associer un certain nombre de concepts à partir d'images. La recherche s'effectue à partir d'environ un million d'images extraites depuis les collections de Flickr.

L'utilisateur commence par taper librement dans la barre de recherche un mot-clé, une première galerie de résultats apparaît alors. Il est ensuite possible d'affiner la recherche en sélectionnant un des concepts proposé par le logiciel, celui-ci effectuant alors un tri des images en fonction de la demande. Par exemple si l'usager désire trouver des images des rues de Paris, il pourra affiner sa recherche en choisissant la catégorie « city », si au contraire il désire des vues plus larges de la ville la catégorie « skyline » sera plus adaptée.



III. 33 Capture d'écran d'une recherche effectuée sur Behold

Depuis peu, Behold suggère automatiquement à son utilisateur les filtres vecteurs de résultats par l'introduction d'une barre de miniatures au dessus de la galerie d'images. Les résultats semblent performants mais les recherches sont encore limitées, le moteur ne prenant en compte qu'une quinzaine de catégories sémantiques à partir desquelles l'image est analysée.

La plupart des moteurs mis au point selon cette technique d'annotation automatique requiert malgré tout une phase de renseignement manuel ne serait-ce que pour la

constitution de la base d'apprentissage à partir de laquelle le moteur effectue les indexations. Si cette opération est réalisable dans le cadre d'un projet universitaire, scientifique mené sur une collection d'items restreints, son application sur Internet est moins évidente en raison du nombre toujours croissant d'images circulant sur le Web et de leur grande hétérogénéité. Une solution développée par quelques sociétés spécialisées est d'avoir recours au « crowdsourcing », de même que pour l'indexation purement textuelle des images, le recours aux internautes pour effectuer une phase préliminaire d'indexation par mots-clés est envisageable et souvent efficace.

D. Les systèmes de reconnaissance faciale

Les technologies de reconnaissance faciale sont un champ de recherche largement mis en avant par les avancées de la biométrie. Elles ressortent plus précisément de la branche des analyses morphologiques, au même titre que la reconnaissance des empreintes digitales, de la forme de la main, des dessins du réseau veineux de l'œil, de l'iris, etc. autant d'éléments qui ont l'avantage d'être stables dans la vie d'un individu et ne subissent pas autant les effets du stress par exemple, que l'on retrouve dans l'identification comportementale. Les domaines d'applications de ces technologies sont plus largement dédiés à l'authentification d'identité, dans le cadre de sécurisation de donnée ou de lieux, à l'amélioration des systèmes de vidéo-surveillances, etc.

Depuis une dizaine d'années ce domaine se trouve au centre de projets en lien avec le traitement d'images, l'indexation et à la recherche des images par leur contenu. Dans ce champ des nouveaux médias, l'application aujourd'hui la plus répandue est la détection et de la reconnaissance des traits humains est celle de l'identification. Il s'agit de permettre un regroupement, un classement des images par similarité de traits et donc leur indexation sous un même tag.

1. Les principes

En raison de la grande diversité des apparences d'un visage humain, l'élaboration d'un logiciel dédié à sa détection est une entreprise des plus complexes qui suggère de surmonter la variabilité des visages en termes de taille, de forme, de couleur ou encore de texture. Les difficultés auxquelles font face les systèmes d'analyse de visages sont nombreux : forte variabilité des paramètres intrinsèques comme l'expression faciale et l'orientation du

visage ; mais aussi des paramètres extrinsèques comme les conditions de prises de vues de l'image (surexposition, ombres portées sur le visage ...) ou les occultations partielles.

Le but de la détection faciale dans le champ de la photographie numérique est donc d'identifier, à partir d'une image, les régions où apparaît un visage et ce quelle que soit sa position, son orientation, ou les conditions de luminosité. La détection des visages sur une image est la première étape indispensable à l'élaboration d'un système de reconnaissance facial. Le but est ainsi de localiser et d'extraire la région du visage de l'arrière plan de l'image.

Au début des années 1970, la reconnaissance par le visage se base principalement sur des attributs faciaux mesurables comme l'écartement des yeux, des sourcils, des lèvres, la position du menton, la forme, etc. Une grande majorité des systèmes se concentre également sur la détection des yeux comme caractéristique centrale. Les yeux semblent offrir un critère des plus aisés à détecter, l'apparence du nez changeant avec la pose du visage et celle de la bouche avec l'expression faciale¹⁵⁵. Depuis les années 1990, les différentes technologies utilisées exploitent toutes les découvertes effectuées dans le domaine du traitement d'image. Ainsi existent-il des techniques de reconnaissance par analyse du visage qui mettent en œuvre un ou plusieurs détecteurs « bas niveau » (contour, symétrie, couleur ...) et des méthodes basées sur l'apparence qui s'inspirent des algorithmes d'apprentissage statique ou dynamique à partir de modèles.

De nombreuses méthodes de détection et de localisation des caractéristiques faciales ont donc été développées ces dernières années prenant en compte des caractéristiques multiples. Pour que ces techniques soient des plus efficaces elles doivent se baser sur des éléments du visage qui sont le moins susceptibles aux changements. Ainsi les grands traits supérieurs des orbites, les secteurs entourant les pommettes, les côtés de la bouche et d'autres caractéristiques similaires sont les caractéristiques physiques les plus stables et permettent d'ignorer les changements qui interviennent avec le temps.

Trois méthodes sont principalement utilisées par les fabricants de systèmes biométriques : l'Eigenface¹⁵⁶, l'analyse de points particuliers et le traitement automatique de visage.

Le *traitement automatique de visage* est une technologie rudimentaire, elle caractérise les visages par des distances et des proportions entre des points particuliers comme les yeux, le nez les coins de la bouche. Aussi éprouvé que les autres technologies,

¹⁵⁵ Cf article sur l'analyse du visage en biométrie en ligne <http://www.biometrie-online.net/techno/visage/T-vis.php>

¹⁵⁶ Cf l'article Wikipedia pour plus d'informations <http://en.wikipedia.org/wiki/Eigenface>

le traitement automatique de visage est la plus efficace dans des situations de capture d'image avec peu d'éclairage.

La technique dite de *l'Eigenface* utilise quant à elle une représentation des éléments caractéristiques d'une image de visage à partir d'images modèles en niveau de gris. Des variantes de Eigenface sont fréquemment utilisées comme base pour d'autres méthodes de reconnaissance.

Enfin *l'analyse de points particuliers* est la technologie d'identification faciale la plus largement utilisée. Si cette technologie se rapproche de l'Eigenface, elle permet de mieux s'adapter à des changements d'aspect facial tels que le sourire, le froncement des sourcils, etc.

Le problème de la reconnaissance de visages dans le contexte d'une base de données d'images pourrait se poser ainsi : étant donnée une image de visage qu'elle est l'identité de la personne correspondante. Pour ce faire, il est nécessaire d'avoir des images de référence, sous la forme d'une base de données de visages de toutes les personnes connues par le système. A chaque visage est associé un vecteur de caractéristiques. Ces caractéristiques sont supposées être invariantes pour une même personne, et différentes d'une personne à l'autre. La reconnaissance consiste alors à comparer le vecteur de caractéristiques du visage à reconnaître avec celui de chacun des visages de la base. Une fois les opérations de détection et d'analyse de similarité des visages effectuées intervient alors l'étape d'indexation. Chaque ensemble de visages similaires doit être associé à un tag, en l'occurrence le nom de la personne, pour permettre non seulement le classement des photographies mais aussi leur recherche. Cette indexation résulte la plupart du temps d'une opération d'annotation automatique des images telle que nous en avons décrit le procédé auparavant, annotation qui permet d'associer à chaque catégorie de visages similaires créée l'identité nominative d'une personne.

2. Les applications grand public

Si les technologies de détection des visages sont largement investis par le domaine de la surveillance et de sécurité biométriques, depuis quelques années on assiste également à leur émergence au sein d'applications grand public. Depuis peu les constructeurs d'appareils photos numériques ont dotés les boîtiers reflex et bridge de fonctionnalités nouvelles de reconnaissance des visages. Par cette option de mise au point l'appareil est capable d'adapter ses paramètres en fonction des objets détectés dans le champ. En parallèle de la détection des visages comme outil de prise de vue, ces technologies sont

apparues depuis quelques années seulement au sein de logiciels de gestion personnelle des collections photographiques. L'enjeu pour les développeurs est de présenter cette technique comme une offre supplémentaire de classement des images, un système de tri automatique des images selon les personnes détectées mais facilitant également grandement la recherche d'une photo particulière.

En 2006 Google acquiert la start-up Neven Vision spécialiste des traitements de l'image et des logiciels de reconnaissance photographiques. Celle-ci a développé une technologie de reconnaissance faciale capable d'analyser des photos et documents vidéo à la recherche de personnes, de lieux, et d'objets. « Ce n'est pas toujours aisé de faire une recherche sur vos photographies personnelles, et cela est certainement beaucoup plus difficile que de chercher sur le Web. A moins que vous preniez le temps de les nommer et de les organiser (...) les chances de retrouver la photo cachée aux fins fonds de votre ordinateur sont minces » peut-on lire sur le blog de Google¹⁵⁷. Cette alliance doit donc permettre au moteur de recherche d'accroître sa gamme d'outils d'aide à la gestion des collections personnelles d'images et donc d'améliorer les potentialités de Picasa.



III. 34 Capture d'écran de l'application de reconnaissance faciale de Picasa

La fusion Google /Neven Vision a ainsi abouti dès 2008 à l'intégration d'une fonction de reconnaissance automatique des visages au sein de la dernière version de Picasa, un outil plus spécifiquement accessible avec la mise à jour de Picasa Web Albums. Avec sa

¹⁵⁷ Traduit par nos soins, En ligne <http://googleblog.blogspot.com/2006/08/better-way-to-organize-photos.html>

technologie, Google simplifie grandement le processus de reconnaissance des visages puisque le service est censé les détecter automatiquement. Après une analyse de chaque album, les photographies comportant un visage repérable sont sélectionnées, il suffit alors à l'utilisateur de renseigner le nom de la personne pour chaque visage détecté. Au fil des ajouts de photos dans les galeries, l'application est capable de suggérer automatiquement des *tags* en se basant sur les photographies précédentes. Mike Horowitz, chef de produit de Picasa, affirme dans une interview accordée à Stephen Shankland sur le site CNET¹⁵⁸, que leur technologie fonctionne bien « quand la personne fait face à l'appareil photo », avouant qu'il y a « un certain nombre de facteurs qui limitent les chances de réussite, en particulier les vues de profil et les conditions de lumières difficiles ».

Disponible également dans la nouvelle version de Picasa dédiées aux Macs, cette application vient en concurrence du logiciel de gestion et de traitement d'images spécifique à cette société, *iPhoto '09*, apparu à la même époque sur le marché. Apple dans la dernière version de son logiciel a ainsi inclus une option de reconnaissance faciale qui lui faisait jusqu'alors défaut. *iPhoto* passe en revue les images stockées dans l'ordinateur de l'utilisateur afin de détecter les visages. Lorsqu'un nom est associé par l'utilisateur à un visage, l'ensemble des clichés qui correspondent à une même personne sont alors localisés.



III. 35 Capture d'écran de la page *Visages* de *iPhoto 09*

¹⁵⁸ Stephen Shankland, « Revamped Google Picasa site identifies photo faces », *CNET.News*, 2 septembre 2008, en ligne http://news.cnet.com/8301-17939_109-10026577-2.html?part=rss&tag=feed&subj=Webware

Une sorte de trombinoscope baptisé *Visages* classe par nom les images de chaque individu identifié. De là, il est aussi possible par simple survol d'un cliché d'ajouter l'identifiant Facebook de la personne, un lien privilégié avec le réseau social étant proposé, au même titre qu'avec le site de partage de photos Flickr.

Il va sans dire que cette technologie nouvelle d'indexation des images en fonction des personnes représentées, appliquée au champ de la vidéo, offre une perspective des plus intéressantes en ce qui concerne la gestion des archives audiovisuelles. La reconnaissance et l'indexation automatique des personnes sur les séquences vidéos par la création d'un moteur de fouille dédié laisse présager un gain de temps de traitement documentaire important et une optimisation de l'usage des collections en terme de recherche.

3. Les applications du Web

Depuis peu sont également apparues des sites Web spécialisés dans la reconnaissance des visages et proposant des applications centrées sur cet outil.

*MyHeritage*¹⁵⁹ est un réseau social focalisé sur la famille et la généalogie. Il permet aux membres de créer leur propre site familial, de partager leur photos et vidéos, de créer leur arbre généalogique et de rechercher des ancêtres. MyHeritage permet de tagguer automatiquement ses photos grâce à un système de reconnaissance faciale. Le logiciel permet de détecter automatiquement les visages humains sans que l'utilisateur n'ait à tagguer ses photos, si la personne fait partie de l'arbre généalogique, alors elle sera automatiquement reconnue. De plus et dans un versant plus ludique MyHeritage offre une technologie de reconnaissance faciale qui permet de découvrir à quelle célébrité l'on ressemble le plus, ainsi qu'une application où les utilisateurs peuvent également comparer la ressemblance d'un enfant à ses parents.

En 2007 une jeune société suédoise a lancé Polar Rose¹⁶⁰, un logiciel en ligne de reconnaissance faciale qui fonctionne à partir des collections personnelles des usagers en ligne sur Flickr. Polar Rose permet aux usagers de classer, chercher et partager leurs images à partir de leur contenu.

Le site fonctionne ainsi comme un moteur de recherche à caractère privé, capable de reconnaître les visages de ses propres collections d'images grâce donc à une technologie de reconnaissance faciale associée à la création d'un modèle 3D à partir d'images 2D, modèle faisant abstraction de la luminosité et de tout autre facteur qui peuvent affecter le

¹⁵⁹ <http://www.myheritage.fr/>

¹⁶⁰ <http://www.polarrose.com/>

processus de reconnaissance. Plus concrètement le site fonctionne selon deux étapes, l'installation d'API sur des sites partenaires, avec pour le moment des partenariats établis entre le site, Flickr et Facebook, à partir desquels l'image en 3D est constituée et légendée. Le modèle est ensuite intégré à la base de données de Polar Rose. Un plug-in pour Firefox (une version pour Internet Explorer est à l'étude), permet de chercher des images dans cette base de données à partir d'un nom (et d'éventuellement corriger les mauvaises indexations).



III. 36 Capture d'écran de l'outil de reconnaissance faciale dans l'application Flickr

Comme tout système de reconnaissance faciale, si la technologie est capable de créer une "empreinte visuelle" unique à partir des photos, elle n'est pas capable de définir l'identité de la personne. S'il est basé sur des problématiques d'indexation et de recherche par le contenu, son fonctionnement s'avère encore largement tributaires de l'indexation textuelle, seule une étape de renseignement du nom de la personne est garante d'une utilisation probante du moteur. L'originalité du site est d'avoir fait appel aux internautes pour résoudre ce problème, à l'aide d'un plug-in mis à leur disposition, ils pourront indiquer le nom et les informations concernant la personne représentée. À partir d'une certaine quantité de données récoltées sur une même personne, le système pourra déterminer son identité sur une nouvelle photo, de façon automatique, et indépendamment de l'expression du visage ou de l'angle de prise de vue de l'image.

III. Les enjeux de la recherche par le contenu

A. Le retour du texte ?

1. Les avantages du « bi-modal »

Certains moteurs du Web privilégient la combinaison *texte-image* lors d'une session d'interrogation. On peut donc avec cette méthode avoir recours à une image requête et à la recherche par mot clé.

La recherche par image exemple circonscrit des champs d'usages très spécifiques et difficilement adaptables à un mode d'utilisation des moteurs grand public qui demeure bien souvent basé sur la requête par mot-clé. Le principe de la recherche de document sur un moteur est d'exprimer sa requête sous la forme d'une séquence de descripteurs empruntés au même vocabulaire que celui utilisé pour l'indexation, le logiciel compare ensuite la requête et l'index pour proposer les documents dont l'index coïncide tout ou en partie avec la requête. Dans le cas d'une indexation par des descripteurs dits de bas niveaux, ceux propres aux images que sont la couleur, la forme ou encore la texture, l'interrogation des bases d'images ne peut se faire qu'à partir de requêtes mettant en jeu un vocabulaire identique, on proposera donc des requêtes par l'image et non par des termes de haut niveau soit sémantiques.

On se trouve souvent face à un écart entre le besoin de l'utilisateur, qu'il désire exprimer en langue naturelle et qui est de très haut niveau sémantique, et les descriptions des images dont on dispose, qui sont, elles, de bas niveau sémantique. Entre l'homme capable d'appliquer une grille sémantique sur un contenu visuel et la machine capable de lire des pixels et de calculer des algorithmes complexes, le fossé sémantique n'a toujours pas été franchi, « On obtient de manière automatique des descripteurs qui reflètent le contenu physique des documents, et l'indexation manuelle permet d'obtenir des concepts interprétant le document dans son contexte. Il y a donc un gouffre sémantique entre les descripteurs physiques de l'analyse automatique et les concepts sémantiques de l'interprétation manuelle.»¹⁶¹ C'est ce qu'on nomme le *semantic gap* : comment traduire la

¹⁶¹ Bruno Bachimont : L'indexation automatique - enjeux, possibilités et limites [en ligne]. Disponible sur : <http://www.culture.gouv.fr/culture/dglf/rifal/indexation.htm>

recherche d'une image illustrant des thèmes ou concepts précis à partir d'une indexation par descripteurs de couleur ou de forme.

Tout le but des nouveaux projets sur la recherche d'images par le contenu est de générer du contenu sémantique sur des fichiers visuels, de dépasser le fossé sémantique. Il s'agit d'indexer par le contenu pour arriver à une automatisation de l'indexation sémantique, aboutissant à l'élaboration de moteurs de recherche hybrides, permettant à la fois une recherche en langage naturel et une sélection des images par similarité, proximité visuelle. Constatant les usages spécifiques qui découlent d'une requête entièrement visuelle, les recherches ont tendance depuis peu à s'axer de plus en plus vers la création d'outils bimodaux, privilégiant la combinaison *texte-image* lors d'une session d'interrogation. Cette méthode permet d'allier à la fois l'utilisation d'une image requête et d'y associer une annotation textuelle, ou de débiter par une recherche sémantique classique puis d'affiner les résultats par comparaisons d'images selon des descripteurs spécifiques.

Les avantages de l'annotation automatique d'images par rapport à la recherche d'image par le contenu (CBIR) sont que les requêtes peuvent être spécifiées plus naturellement par l'utilisateur, sous forme de requête textuelle. Dans les systèmes de CBIR, l'utilisateur doit rechercher les images par des concepts tels que la couleur ou la texture, ou alors doit donner une image exemple pour laquelle il souhaite trouver des images similaires. Les moteurs bimodaux confèrent une aisance d'utilisation aux usagers et sont également capables de gérer des collections d'images non seulement importantes mais aussi croissantes, puisque l'indexation textuelle est générée automatiquement.

2. Conservation et partage des annotations

L'un des problèmes dans la circulation des images sur le Web est la perte des données informationnelles entourant l'image. Si une image est mal renseignée, notamment au travers de ses métadonnées, l'on peut aboutir à l'obtention d'images esseulées, images se retrouvant dans le flux du Web sans informations contextuelles. Le système des moteurs de recherche de type Google Images, basé sur une indexation textuelle des images faite à partir des informations les entourant, s'avère d'une efficacité encore limitée, ces moteurs offrent toutefois la possibilité de replacer l'image dans le contexte de la page Web, de lui conférer parfois quelques éléments de contexte.

Le problème au sein des interfaces de recherche par l'image, de recherche par le contenu est l'absence de prise en compte des éléments contextuels de l'image. Là où le

géotagging replaçait tout du moins l'image dans un horizon géographique, les moteurs de recherche par le contenu semblent abstraire l'image de la sphère informationnelle pour ne garder que les caractères visuels alors seuls critères d'intelligibilité, de lecture de l'image. Cette lacune semble décuplée avec les interfaces fonctionnant en amont à partir des sites tel que Flickr, puisqu'on assiste à une sorte de double abstraction de l'image, celle-ci déjà dans le flux de son site source se trouve à nouveau extirpée par le moteur. Un nombre toujours croissant d'opérations est alors nécessaire pour retrouver sa source, son photographe et autres détails, s'il en est, de sa prise de vue.

En plus de son interface de visualisation 3D, le projet Photo Tourism¹⁶² avait également prévu de mettre au point un système permettant non seulement de conserver les annotations des images, mais aussi d'en générer pour celles s'en trouvant dépourvues. Dans cette application de mise en relation des images géographiquement proches jusqu'à la recreation d'un paysage en 3D, les auteurs avaient donc envisagé la possibilité de conférer aux images une strate informationnelle venant se surajouter à la dimension de contextualisation géographique. Il s'agissait là de prémisses à l'idée d'une annotation sémantique automatique des images, dans le projet toutes celles partageant les mêmes coordonnées ou certaines suffisamment proches se voient attribuer les mêmes informations. Un glissement d'annotations d'une image à l'autre était donc prévu, opération effectuée non en fonction d'une analyse visuelle des images mais selon leur localisation, critère alors conçu comme déterminant dans l'attribution ou non des informations aux images.

B. Des perspectives à visée commerciales

Si l'indexation par coordonnées géographiques et les interfaces cartographiques dédiées à la recherche d'images circonscrivent, comme nous l'avons vu, une sphère iconographique particulière, les moteurs de recherche d'images par le contenu tendent pour leur part à ne prendre en compte que certains types de recherche. Les différentes formes de requête permises par ces moteurs décrivent des désirs et pratiques documentaires spécifiques voire professionnelles, pas toujours adaptés à une utilisation grand public de ces techniques.

¹⁶² Voir Chapitre 3, p.

1. Pour quels usages dans la sphère de l'utilisateur non pro ?

Les besoins d'images, leurs utilisations et les façons d'interroger une base de données varient considérablement d'un usager à l'autre, selon qu'il soit journaliste, historien, artiste ou simple curieux, un internaute agit toujours selon un besoin particulier à l'œuvre derrière la question. Si certains internautes peuvent vouloir accéder aux images par l'intermédiaire de la description de leur contenu visuel, couleur, texture et forme, d'autres n'envisagent la recherche qu'à partir de mots-clés, du recours au sens sémantique.

Le point commun des moteurs de recherche d'images par le contenu est qu'ils fonctionnent selon le critère de similarité visuelle, les descripteurs de bas niveaux appliqués sur les images ayant pour fonction de permettre cette recherche de l'image ressemblante. Les systèmes de recherche d'images développés jusqu'à présent ont pour la plupart tenu compte des différents niveaux et types d'usages, ils ont ainsi multiplié les modes de requêtes dans le but de répondre aux divers besoins d'images manifestés par les utilisateurs. Tous proposent, et pour certains de façon exclusive, la requête par image exemple. Si ce mode d'interrogation de la base de données semble révolutionnaire et inscrit l'outil dans le pur champ de la recherche visuelle, permettant de se libérer du recours jusqu'alors systématique au mot, la question qui émerge à l'usage est alors celle de son intérêt. Car ce mode de recherche, aussi novateur qu'il puisse être, soulève par la même occasion quelques limites, privant de son emploi l'utilisateur indécis ou n'ayant pas d'idée, d'exemple à soumettre. On peut également se demander si l'utilisation d'une unique image exemple n'engendre pas une spécification limitée de la recherche, excluant de ce fait des images moins visuellement proches selon les critères des descripteurs mais tout aussi pertinentes dans le domaine du sens. Du point de vue des usages enfin, le public non expert peut trouver laborieux de manipuler et de juger de l'importance des paramètres de recherche détaillée que sont la couleur, la forme ou la texture qui permettent d'obtenir plus de précision mais se révèlent moins intuitifs que le texte.

On en revient ainsi au problème déjà évoqué du fossé sémantique, pouvant faire rencontrer à l'utilisateur des difficultés dans la formulation de ses besoins sur des interfaces où le recours au texte n'est le critère ni prédominant ni fonctionnel. L'utilisation des moteurs à image requête circonscrit donc un type de recherche particulière, précise et déterminée, loin des schémas de navigation hasardeuse, et semble donc s'adresser à un public averti. Nous pouvons pour l'exemple nous référer à TinEye, le moteur fonctionne à partir d'une image requête sur laquelle s'effectue une recherche globale. Le retour de résultats est basé non

sur la proximité visuelle mais sur l'exactitude iconique, seules les images identiques à l'exemple sont recherchées. Le recours à une image exemple est dans ce cas absolument nécessaire, sans elle aucune recherche n'est possible. À partir de ce principe une application phare du moteur, prônée par la société créatrice, est la détection des copies d'images illégales circulant sur le Web (cf. partie suivante). Dans l'optique d'une utilisation plus grand public, moins spécialiste, du moteur, ce dernier peut également permettre de rechercher l'origine des images du net. La circulation incessante des documents sur Internet engendre l'apparition de nombreuses images « orphelines », documents ayant au cours de leur emploi et réemploi fini par perdre une partie de leurs informations, de leur contexte, aboutissant souvent à une restriction de leur sens. L'utilisateur curieux peut envisager avec TinEye de remonter au site source d'une image, ou du moins de trouver des pages Web où le document possède encore certaines informations, où son histoire est relatée, et pourquoi pas de comparer, à travers les différents sites recensés, les divers modes de présentation de la dite image, l'usage qui en est fait.

D'autres moteurs, conscients du manque d'intuitivité de la requête exemple, permettent le recours au mot-clé pour débiter la recherche, et l'option de comparaison de similarité dans un second temps, encore faut-il pour que le système s'avère efficace que les images soient textuellement correctement indexées. On comprend dans ce contexte l'importance que revêtent les recherches sur l'annotation automatique des images, systèmes pouvant permettre comme nous l'avons vu à la fois une interrogation intuitive de la base parce que langagière et une comparaison des images selon leurs critères propres, visuels.

Si beaucoup de moteurs de recherche d'images peuvent par bien des côtés apparaître comme des solutions plus ludiques que véritablement performantes ou essentielles, un système se démarque toutefois : la reconnaissance faciale. Cette technique est aussi bien prisée des professionnels qui voient en elle la possibilité de repérer plus aisément les documents dans lesquels une personne apparaît, qu'elle se développe de plus en plus sous forme de logiciels ou de sites Internet à usage du grand public. Les utilisations possibles de la reconnaissance faciale sont nombreuses et facilement accessibles aux usagers les plus novices. Il devient ainsi possible de retrouver les photos d'une personne déterminée sur les sites communautaires ou les sites d'information, regrouper ses photos de famille dans le cas des logiciels ou sites de gestion des images, identifier des personnes sur des documents informatifs. Face à la profusion de clichés disponibles sur le net, ce type de solution est une nouvelle alternative aux méthodes de recherche et de classification

reposant essentiellement sur des informations textuelles. Le recours au mot-clé, devenu nom de la personne, est dans ce cas indispensable, et offre par conséquent une bonne adaptabilité du système aux besoins des utilisateurs même si ici aussi les usages sont spécifiques et se doublent d'une circonscription du champ iconique aux photographies, images de personne. Les sites en ligne qui développent ce système misent l'efficacité et l'exhaustivité de leur indexation sur la participation des usagers, il reste encore à déterminer si les utilisateurs joueront le jeu de cette étape d'identification, tout comme il reste à mesurer en quoi ce système offre une réelle efficacité de classement des images aux particuliers.

2. La fin du piratage des images ?

Pour de nombreux photographes et agences, la diffusion de photographies sur Internet est un moyen simple et rapide de faire connaître leur travail ou collection et donc d'accroître leur public. Si le nombre d'images disponibles sur le Web ne cesse d'augmenter, le nombre de copies, de réutilisation illégale des images augmente lui aussi. Publier une image sur Internet ne signifie pas l'abandon des droits d'auteurs, si des communautés de défenseurs de la culture libre ont développé ces dernières années des formes de licences permettant un libre réemploi des images par les internautes, comme le prône *Creative Commons*, la plupart des images circulant sur le Web sont toujours soumises à copyright et donc à une législation stricte quant à leur acquisition ou utilisation.

Il existe deux types de "vol" de photographies :

- La copie pure et simple, qui implique au voleur de prélever (par exemple par copier-coller, glisser-déplacer) et d'héberger à nouveau le fichier image sur son site, page personnelle de blog.

- Une autre technique est le *hotlinking* de l'image, de plus en plus répandu, et pourtant illégale. Le hotlink est l'affichage d'une image en la prenant directement à la source du site qui la propose. Au lieu d'enregistrer l'image sur son ordinateur et ensuite de l'insérer sur son blog ou site, on utilise directement l'url de cette image pour la publier. Cette technique, pour qui connaît les rouages des serveurs Web, est décelable et contournable relativement facilement. La copie d'image en revanche s'avère sensiblement plus difficile à contrer.

Des méthodes existent pour limiter le vol d'images sur Internet, en gérant la définition, la taille d'affichage des images, par la désactivation du clic-droit de la souris supprimant l'option de copiage de l'image. Il est également possible d'empêcher le glisser-déplacer

d'une image, ou l'ouverture de la barre de gestion des images. Toutes ces techniques demeurent toutefois limitées, fastidieuses et ne se présentent pas en outil dissuasif.

Jusqu'à présent la détection des images « volées », ne pouvait s'effectuer qu'à partir de Google Images. Le moteur s'avère dans ce cas de figure simple d'utilisation mais moyennement efficace. Google propose au travers de son outil dédié aux images de pouvoir rechercher selon divers critères, mots-clés tout ce qui est affichable sous forme de fichier bitmap : photographies, schémas, dessins, etc. Une recherche sur les mots-clés liés à vos images phare permet donc de déceler d'éventuels sites les utilisant sans autorisation.

L'une des facilités de *Google Images* est de permettre de rechercher toutes les images hébergées sur un domaine donné. Indexées dans sa base de données, cette option permet notamment de détecter les hotlinks. Les limites évidentes du recours à Google pour ce genre de détection d'images sont liées aux limites d'indexation connues par le moteur. Qu'il s'agisse des images du détenteur des droits ou de celles du « copieur », elles doivent être indexées sur le moteur, pour que la recherche soit opérante. Or, l'indexation des images sur GI est non seulement lente mais aussi fluctuante, certains blogueurs rapportant par exemple la désindexation puis réindexation régulière de leurs images, la prise en compte par le moteur d'une partie seulement de leurs images en ligne. De plus comme toute recherche effectuée sur Google la détection s'opère à partir des mots-clés de la requête ou du texte entourant l'image et non en fonction de l'image elle-même.

Dans ce contexte on mesure combien l'apparition du moteur de recherche TinEye est en passe de révolutionner la détection des usages illégaux d'images. Par sa capacité à détecter non pas les images ressemblantes mais identiques, sa robustesse au recadrage et variation de couleurs TinEye fait preuve d'une grande performance en la matière. L'application a déjà intéressé l'AFP et Getty Image clients qui ont passé des accords avec la maison mère de TinEye, Idée Inc. Le site offre ainsi la possibilité pour les auteurs, les agences ou tous les détenteurs de droits de vérifier les diffusions et usages illégaux des images sur internet, de retrouver les sites sur lesquels sont diffusés des photographies, logos, tout type d'images soumis à des droits.

Demain avec la recherche de vidéos, que TinEye annonce disponible pour la fin de l'année 2009, le moteur pourrait permettre à la fois de suivre la diffusion et la propagation d'une vidéo virale, tout en offrant une nouvelle solution pour surveiller et réaliser la protection de droits audiovisuels sur Internet. Dans un versant commercial toujours mais plus marketing, TinEye peut aussi permettre aux agences de communication, publicitaires de vérifier si une photographie, un logo est déjà utilisé ou non. Ce type de recherche

visuelle peut avoir un impact potentiel fort sur les ventes des images proposées par les microstocks (Fotolia, istockfoto,etc.) puisqu'il devient non seulement simple mais aussi rapide de vérifier si l'un des annonceurs concurrents utilise ou non le même visuel, et le cas échéant d'en choisir un différent. Des exemples de ces applications sont ainsi visibles sur la page *Cool Searches*¹⁶³ de l'interface. Bien sûr et au même titre que Google image, l'une des conditions de l'efficacité de la recherche est l'indexation des images par le moteur, à l'heure actuelle TinEye assure avoir indexé 1 043 884 366 images du Web et continue de le parcourir.

Cool Searches (click image to search)

This page, along with our [widgets](#) page, showcases some of the most interesting TinEye search examples we've come across. If you come across a TinEye search with great results and you want to share it, just click on 'Recommend as a Cool Search' next to your query image. We review every Cool Search submission we get.



With stores in 18 countries, it's no wonder that the 7-Eleven logo is everywhere online.



Little little energy star, how I wonder where you are?

III. 37 Capture d'écran du module Cool Searches de TinEye

3. L'enjeu des sites marchands

Si TinEye se place dans le crédo d'une utilisation commerciale, les autres sites de recherche par similarité cette fois ne dépareillent pas. Des applications telles que Gazopa ou Xcavator ne masquent pas leur visée lucrative, Xcavator par son affiliation avec des sites de microstocks ou des banques d'images en ligne se positionne, selon son propre slogan comme un moteur de recherche dévoué aux professionnels, permettant aux agences de communication de trouver plus facilement une image d'après leur souhait.

Gazopa quant à lui, se destine de façon plus affirmée au grand public, si comme nous l'avons mentionné, des options ludiques de recherche sont proposées, l'application phare du site est sa capacité à aller chercher des images similaires au sein des pages de site marchand. Ainsi un utilisateur effectuant une requête par dessin ou téléchargeant une image de botte, de voiture, se verra proposer des résultats issus de vendeurs en ligne mêlés

¹⁶³ http://tinEye.com/cool_searches

à des images à visée non commerciale. L'utilisation de Gazopa semble destinée à long terme à devenir bien plus un moteur de recherche de produit similaire que d'images à proprement parler. Si pour l'instant les sites et images indexés sont encore peu nombreux et essentiellement d'origines anglo-saxonnes, le développement de l'interface, encore en Beta privé, promet d'offrir une gamme variée d'utilisation aux usagers.

Là où Gazopa se démarque et semble envisager de façon pertinente les potentialités, à la fois pratiques et économiques, de la recherche par le contenu se retrouve du côté de son plugin pour iPhone. Par la combinaison du camphone, de l'écran tactile et de la 3G, il devient possible d'interroger la base de données de Gazopa de façon embarquée. La recherche sur mobile s'effectue en temps réel, instantanément, devant un objet devenant image, sujet de requête.

GazoPa
similar image search

GazoPa for iPhone

[GazoPa for iPhone in iTunes](#)



III. 38 Capture d'écran de la présentation de l'option pour I Phone de Gazopa

L'idée rejoint celle de l' " Internet des objets " rêvée pour les technologies sans fil à courte portée, une technologie créatrice de liens entre l'Internet et les objets physiques. Si ces nouveaux modes de recherche et d'interrogation des données du Web semblent favorables à l'utilisateur, le constituant en sujet toujours connecté, ces technologies sont également largement lucratives pour les entreprises pouvant leur assurer des retombées ludiques et publicitaires.

Selon un principe proche, le projet *Point & Find* de Nokia¹⁶⁴ est en développement depuis plus d'un an et demi et cherche à faire coïncider reconnaissance de forme et contenus mobiles. Une version Beta privée du logiciel et consacrée au cinéma est d'ores et déjà disponible au Royaume-Uni et aux États-Unis. Le principe est simple : un objet photographié avec l'appareil d'un téléphone portable et soumis à la base de données du projet permet de déclencher en retour certaines actions. Il est possible, d'après les vidéos de démonstration, de viser l'affiche d'un film, ce qui, après analyse, permet de recevoir en retour des informations sur celui-ci ou un lien vers sa bande-annonce.

Ce nouveau type d'application peut être également un excellent support publicitaire interactif et permet d'envisager de nouvelles pistes marketing, sous réserve de maîtriser les différents aspects : reconnaissance de forme, comparaison avec une banque d'images et identification de l'objet. La société invite d'ailleurs les professionnels intéressés par le procédé à se manifester pour imaginer les applications et campagnes publicitaires de demain. Nokia Point & Find et les autres applications du genre peuvent se constituer en maillon d'un environnement plus vaste prenant en compte notamment le positionnement GPS.

¹⁶⁴ Cf <http://pointandfind.nokia.com/?home>

Conclusion

Les systèmes d'indexation et de recherche par le contenu sont ils en passent de révolutionner le traitement des images sur le Web ? Apparues comme nous l'avons vu dès les années quatre vingt dix et largement développées depuis, ces techniques de description et de catégorisation visuelle des images peinent à donner des systèmes pertinents de recherche d'image en raison du manque de fiabilité de leur méthode d'indexation. À travers l'observation menée des divers moteurs de recherche par le contenu aujourd'hui en ligne nous avons ainsi pu constater les limites non seulement techniques mais aussi d'usages de ces derniers.

Au travers des divers exemples évoqués apparaît la difficulté de passer, pour l'utilisateur, d'automatismes de recherche basés sur le langage naturel à des requêtes exprimées par le sondage de l'importance des différents descripteurs, vecteurs visuels. Toutefois ces moteurs ne doivent pas être considérés comme des outils annexes ou anecdotiques, ils semblent de plus en plus à même d'offrir une aide efficace dans le cas d'un besoin spécifiques d'image. Que se soit par exemple pour la détection de copies circulant sans droit avec TinEye, pour une recherche avancée d'images professionnelles avec Xcavator, ou pour la reconnaissance faciale qui a fait son entrée de puis peu au sein des logiciels de gestion des bases d'images personnelles, les applications d'indexation et de recherche par le contenu s'envisagent dans des pratiques spécialisées. Une utilisation ludique de ces moteurs est bien sûr possible, la simplicité, l'ergonomie des interfaces y invitant souvent. Néanmoins, si l'on considère la fonctionnalité première de ces technologies, à savoir la recherche d'images sur le Web, leur usage appelle un public averti, ayant déjà défini ses besoins précis en termes de couleur, de forme, de rendu global de l'image.

Si le monde de la recherche scientifique s'est évertué dans l'intimité des laboratoires et depuis peu au grand jour à mettre au point des systèmes de traitement documentaire automatique des images, le Web n'a toutefois pas attendu après eux pour développer des méthodes d'indexation applicables aux sources iconiques.

Depuis ces dernières années de nouveaux moyens pour indexer, rechercher, naviguer au sein de bases iconographiques ont vu le jour. Du côté de la recherche par mots clés tout d'abord avec l'affinage des critères de requête sur Google ou plus spécifiquement avec

l'apparition de la folksonomie projetant l'indexation dans une sphère collaborative et sociale. Loin de ne figurer qu'un simple outil de gestion des images en ligne les tags apparaissent alors comme autant de voies de communication entre usagers, de lieux d'échanges et d'expressivité.

D'autres systèmes ont également émergé reposant pour leur part sur une prise en compte plus importante des données propres à chaque image et facilement exploitable que sont les métadonnées et en particulier les coordonnées géographiques. Le géotagging envahit peu à peu le Web et avec lui se dessinent une sphère iconique dédiée et de nouveaux usages documentaires centrés sur la notion de territoire. L'espace physique semble réinvesti, augmenté par nos pratiques visuelles. Mit en perspective avec ses diverses représentations cartographiques en ligne il s'enrichit d'autant d'éléments contextuels, historiques, sociaux, et dans un mouvement circulaire nos sources géotagguées profitent à leur tour de ce regain d'informations spatialisées.

La sphère visuelle du Web demeure constituée de systèmes d'indexation d'images plus ou moins bricolés, adaptés tant bien que mal du texte aux images. Si les méthodes de gestion documentaire des images semblent hétérogènes, il en est de même des offres, des possibilités de recherches demeurant plus que disparates mais pourtant fonctionnelles et adaptées semble-t-il à des usages spécifiques.

En l'absence d'un système d'indexation unifié des images tel qu'il en existe pour le texte, il convient alors peut-être de considérer chaque système existant comme un révélateur de pratiques visuelles particulières, qu'il a engendrées ou dont elles résultent. Cette étude nous fait d'une part repenser la nécessité d'un système uniformisé de traitement de l'image sur internet et nous laisse envisager d'autre part l'utilité d'une variété de techniques qui semblent aller de pair avec la diversité des usages et requêtes faites sur le web et de leurs perpétuelles évolutions. Les systèmes d'indexation existants sont donc compris comme des solutions certes limitées mais adaptées chacune à des pratiques particulières et à des corpus iconographiques spécifiques.

Cette étude a également permis de souligner l'importance du rôle du photographe amateur dans la constitution de corpus iconographiques, de collections d'images raisonnées, de base de données navigables, bref dans la circulation des images sur le Web. Une circulation de plus en plus documentée, vectrice d'échanges, de sociabilité et de

créativité. Selon Henry Jenkins¹⁶⁵, la possibilité offerte aux gens de s'appropriier et de diffuser du contenu (grâce aux outils d'archivage, d'annotation, et aux technologies d'échanges de données), aboutit à l'apparition de nouvelles formes de créativité, et à un réinvestissement actif du sujet dans la sphère culturelle. Ainsi se dessine très nettement le retour de l'amateur praticien, au moment même où un nouveau tournant social et économique surgit avec l'ancrage des sociétés modernes dans l'ère de l'hypermédia, du numérique. Nous assistons aujourd'hui à une démocratisation constante de l'acquisition des technologies informatiques et d'enregistrements. Les foyers s'équipant de plus en plus en appareils multimédia, cette démocratisation aboutit à l'émergence d'une figure renouvelée de l'amateur, d'un nouveau public-acteur dans le paysage artistique et culturel. Pour Jean-Louis Weissberg¹⁶⁶ les offres variées de logiciels de traitement d'images, d'échantillonnages musicaux, de boîtes à outils accessibles à tous, encourageant la création de documents au sein des familles. Pour l'auteur, ces « Home multimédia »¹⁶⁷ permettent de faire évoluer les pratiques de réception vers celles de production, remplaçant ainsi l'amateur dans une logique d'inventions, dans une pratique créative.

Cette implication croissante du public dans la production et les politiques d'échanges des informations sur le Web n'est pas sans avoir de conséquences sur la notion même de document. On assiste à l'ouverture du Web sur un champ non scientifique ou officiel, à son ouverture sur les données personnelles des gens devenant entités informationnelles, « Si le Web favorise une transgression de l'ordre documentaire ordinaire, inversement il «documentarise» des expressions qui relevaient autrefois de l'intime et de l'éphémère en les enregistrant, les «traçant», les indexant.¹⁶⁸ ». Le Web favorise conjointement deux mouvements jusque là opposés : le développement d'échanges spontanés, conversations verbales ou visuelles et leur fixation sur un support public, pérenne et documenté.

Enfin qu'en est-il des rêves originels d'une recherche de l'image par l'image, de l'affranchissement des images de la sphère du texte sensément aliénante et limitative qui avaient motivées les premiers projets d'une recherche d'images à partir, et à partir

¹⁶⁵ Henry Jenkins, *Convergence Culture. When old and new media collide*, New York University Press, New York, 2006.

¹⁶⁶ Jean-louis Weissberg, *Présence à distance, l'Harmattan, Paris, 1999, p 187-192.*

¹⁶⁷ Ibid, « L'amateur : émergence d'une figure politique en milieu numérique », in *Multitudes*, n°5, mai 2001.

¹⁶⁸ Jean Michel Salaün, Web, Texte, conversation et redocumentarisation, Presses Universitaires de Lyon, 2008, en ligne <http://jadt2008.ens-lsh.fr/IMG/pdf/JADT-2008-JMS.pdf>

seulement, de leur critère visuels ? Aux vues des implications du géotagging au sein des plateformes de partage d'images, du recours presque toujours nécessaire de tags lexicaux dans le cas du géotagging et en considérant enfin les projets les plus récents d'annotation textuelle automatique des images, il semble que les perspectives d'autonomisation des images ne soient plus aujourd'hui que de simples utopies. Les images ont besoin des mots pour être indexées, mais aussi recherchées, trouvées. Malgré les meilleures interfaces de navigation visuelles, nous en revenons toujours au champ lexical pour définir une requête. Les travaux de recherche dans le champ de l'indexation des images se poursuivent mais dans une optique de génération automatique d'annotation des contenus iconiques. Des recherches menées plus loin non seulement pour des raisons d'universalité des savoirs mais aussi et surtout parce que les mots sont toujours le seul moyen pour les régies publicitaires de contextualiser leurs annonces. En ce temps croissance fulgurante des données du Web, aussi bien textuelles que visuelles, où le Web devient une machine à documentariser, véritable continent documentaire sans précédent, les sociétés ont plus que jamais besoin d'outils leur permettant d'utiliser les informations disponibles comme support de leur visibilité.

Ainsi apparaît-il peu à peu au sein du Web une dichotomie entre pratiques visuelles vernaculaires et spécialisées, véritable partition entre une indexation formelle et technicienne d'une part et une indexation sociale de l'autre. La sphère officielle et professionnelle visible sur le Web révèle des besoins d'une mise au point de systèmes stables d'indexation visant un traitement des images à des fins bien souvent commerciales. A ses côtés émerge un Web dont les contenus ouverts et créés par des utilisateurs lambda engendrent un foisonnement de pratiques nouvelles. Notre travail a ainsi dévoilé comment le Web 2.0 tend à se concevoir comme un espace d'expérimentations touchant également les systèmes d'indexation et leur réappropriation par les usagers à travers le prisme des pratiques sociales.

Bibliographie

Histoire de la photographie et théorie de l'image

Barthes Roland, « Rhétorique de l'image », *Communication*, n°4, 1964.

Debray, Régis. *Vie et mort de l'image*, Paris, Gallimard, 1994. 526 p.

Gunning Tom, « La retouche numérique à l'index. Pour une phénoménologie de la photographie », *Études photographiques*, n° 19, décembre 2006, p. 96-119, en ligne <http://etudesphotographiques.revues.org/index1322.html>

Gunthert André, Poivert michel, *L'Art de la photographie, des origines à nos jours*, Paris, Citadelles & Mazenod, 2007, 620 pages

Jenkins Henry, *Convergence Culture. When old and new media collide*, New York University Press, New York, 2006

Jonas, Irene, « la photographie de famille au temps du numérique », *Enfances, familles, Générations*, n°7, automne 2007.

Joschke, Christian, « Aux origines des usages sociaux de la photographie », *Actes de la recherche en sciences sociales*, n°154, avril 2004, p 53-65.

Foucault Michel, *Les mots et les choses*. Gallimard, Paris, 1966. 405 p.

Krauss, Rosalind, « Notes sur l'index », *L'Originalité de l'avant-garde et autres mythes modernistes*, Paris, Macula, 1993, p. 69.

Rubinstein Daniel et Sluis Katrina, « À life more photographic », *Photographies*, Vol n°1, septembre 2008.

Weissberg, Jean-louis, « L'amateur : émergence d'une figure politique en milieu numérique », *Multitudes*, n°5, mai 2001.

Weissberg Jean-louis, *Présence à distance*, l'Harmattan, Paris, 1999, p 187-192.

Patrick Peccatte, « Notes sur la structure informationnelle de la photographie », *Du bruit au signal (et inversement)*, 5 aout 2008, en ligne

<http://blog.tuquoque.com/post/2008/08/05/Notes-sur-la-structure-informationnelle-de-la-photographie>

Gestion et traitement documentaire à l'ère du numérique

- **Généralités**

Association des archivistes français, « Abrégé d'archivistique : principes et pratiques du métier d'archiviste », Association des archivistes français, Paris, 2004, p. 259-263.

Roger T Pedauque, *Document : forme, signe et medium, les re-formulations du numérique*, Paris, ENSSIB, 2003, en ligne

http://archivesic.ccsd.cnrs.fr/docs/00/06/21/99/PDF/sic_00000511.pdf

Salaün, Jean-Michel, *Le défi du numérique : redonner sa place à la fonction documentaire*, ADBS, 2008, en ligne <https://papyrus.bib.umontreal.ca/jspui/bitstream/1866/2225/1/salaun-jm-defi-documentaliste-mktg.pdf>

Salaün Jean Michel, « La redocumentarisation, un défi pour les sciences de l'information », *Études de Communication* n° 30, Entre information et communication, Les nouveaux espaces du document, Université de Lille3, décembre 2007

Salaün Jean-Michel, *Web, texte, conversation et redocumentarisation*, Presses Universitaires de Lyon, 2008, en ligne <http://jadt2008.ens-lsh.fr/IMG/pdf/JADT-2008-JMS.pdf>

Sutter Eric, « Pour une écologie de l'information », *Documentaliste - Sciences de l'information*, no 2, vol. 35, 1998, p. 83-86

Zacklad Manuel, « Réseaux et communautés d'imaginaire documédiatisées », in Skare, R., Lund, W. L., Varheim, À., *À Document (Re)turn*, Frankfurt, Peter Lang, 2007, 279-297, en ligne :

http://hal.archivesouvertes.fr/docs/00/18/01/85/PDF/imaginaire_document_Skare_Lund_Varheim.pdf

- **Indexation des documents visuels**

Bachimont Bruno. « Bibliothèques numériques audiovisuelles : des enjeux scientifiques et techniques. », *Document numérique*, 1998, vol. 2, n° 3-4, p. 219-242

Bachimont, Bruno, « Image et audiovisuel : la documentation entre technique et interprétation. Critique et perspectives », *Documentaliste-Sciences de l'information*, 2005/6, Volume 42, p. 348-353

Bachimont Bruno, *L'indexation automatique - enjeux, possibilités et limites* [en ligne]. Disponible sur : <http://www.culture.gouv.fr/culture/dglf/rifal/indexation.htm>

Didier, Marie, « Indexation, structuration et encodage des fonds iconographiques : le fonds Léon Lefebvre de la Bibliothèque Municipale de Lille », mémoire d'étude, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques, Paris, 2005, (en ligne : <http://www.enssib.fr/bibliotheque-numerique/document-809>)

Kattnig, Cécile *En quoi les nouvelles technologies renouvellent-elles le traitement intellectuel de l'image ?*, Actes de colloques tenu les 23 et 24 octobre 2003 à Lyon dans le cadre des rencontres nationales cdidoc-fr, 2003, en ligne <http://www2.acrennes.fr/testscdi/rencontrelyon/kattnig/kattnig.pdf>

Manzarena Antoine, *l'indexation d'images*, [en ligne], disponible sur : <http://www.abhatoo.net.ma/index.php/fre/content/download/2563/28254/file/Indexation%20d'images.pdf>

- **Indexation du Web et métadonnées**

Boutin, Éric, Ertzscheid, Olivier, Gallezot, Gabriel *Perspectives documentaires sur les moteurs de recherche : entre sérendipité et logiques marchandes*, 2007, en ligne : <http://hal.archivesouvertes.fr/docs/00/17/21/69/PDF/ertzsgallbout.pdf>

Dublin Core Metadata Initiative, DC-Library Application Profile (DC-Lib), en ligne : <http://dublincore.org/documents/2004/09/10/library-application-profile/>

Ertzscheid Olivier et Gallezot, Gabriel « Etude exploratoire des pratiques d'indexation sociale comme une renégociation des espaces documentaires. Vers un nouveau big bang documentaire ? », *Document numérique et société*, ADBS Éditions, 2006, (en ligne : http://affordance.typepad.com/mon_weblog/files/sdndocsoc.pdf)

Kobayashi, M. and Takeda, K, *Information Retrieval on the Web*, ACM Computing Surveys, 2000, Vol.32, No.2, 144-173.

Peccatte Patrick. Métadonnées: *une initiation Dublin Core, IPTC, EXIF, RDF, XMP, etc.*, en ligne : <http://peccatte.karefil.com/Software/Metadata.htm>

Folksonomies et pratiques sociales de la photographie amateur

Aurnhammer M, Hanappe P, Steels L, *Integrating collaborative tagging and emergent semantics for image retrieval*, In Proc. of the Collaborative Web Tagging Workshop (WWW '06) (May 2006) <http://www.semanticmetadata.net/hosted/taggingws-www2006-files/17.pdf>

Boyd Danah, Marlow Cameron, Naaman Mor, Davis Marc, *Tagging Paper, Taxonomy, Flickr, Academic Article, To read*, Proceedings of Hypertext 2006, New York: ACM Press, <http://www.danah.org/papers/Hypertext2006.pdf>

Burgess Jean, *Vernacular Creativity, Cultural Participation and New Media Literacy: Photography and the Flickr Network*, AoIR 7.0: Internet Convergences, Brisbane, Australia, September 2006

Cardon Dominique et alt., “The Stength of Weak cooperation: À Case Study on Flickr”, 2008, [en ligne], disponible sur <http://hal.archives-ouvertes.fr/docs/00/25/66/49/PDF/weak-coop-flickr-v1.1.pdf>

Crepel, Maxime, « Les folksonomies comme support émergent de navigation sociale et de structuration de l’information sur le web », *Réseaux*, n° 152, juin 2008, p. 169-204, en ligne : <http://www.cairn.info/revue-reseaux-2008-6-p-169.htm>

Golder Scott À. & Huberman Bernardo, *The Structure of Collaborative Tagging Systems*, Information Dynamics Lab, HP Labs, 2005, En ligne : <http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf>

Le Deuff Olivier, « Folksonomies : Les usagers indexent le web », *BBF*, 2006, n° 4, p. 66-70, en ligne <http://bbf.enssib.fr>

Marieke Guy, Tonkin Emma, « Folk-sonomies. Tidying up Tags? », *Dlib Magazine* , N° 1, Volume 12, 2004, en ligne : <http://www.dlib.org/dlib/january06/guy/01guy.html>

Van House Nancy et alt, *From 'What?' to 'Why?': The Social Uses of Personal Photos*, School of Information Management and Systems, University of California, Berkeley, 2004, en ligne http://people.ischool.berkeley.edu/~vanhouse/van%20house_et_al_2004a.pdf

Géotagging

Kentaro Toyama et alt., *Geographic Location Tags on Digital Images*, in Proc. ACM Conference on Multimedia, October, 2003.

Seitz M. Steven, Snavely Noah, Szeliki Richard, *Photo Tourism: Exploring Photo Collections in 3D*, ACM Transactions on Graphics (SIGGRAPH Proceedings), 2006, 835-846, en ligne http://phototour.cs.washington.edu/Photo_Tourism.pdf

Torniai Carlo et alt., *Sharing, Discovering and Browsing Geotagged Pictures on the Web*, 11 mars 2009, en ligne

http://www.hpl.hp.com/personal/Steve_Cayzer/downloads/papers/geospatial_final.pdf

Naaman Mor., Paepcke Andreas et Garcia-Molina Hector, *Automatic organization for digital photographs with geographic coordinates*, In Proc. ACM/IEEE-CS Joint Conf. on Digital libraries,

Naaman Mor, Paepcke Andreas, et Garcia-Molina Hector, *From where to what: Metadata sharing for digital photographs with geographic coordinates*, 11th International Conference on Cooperative Information Systems, 2003, Catania, Sicily, Italy, en ligne :

<http://ilpubs.stanford.edu:8090/756/1/2003-62.pdf>

Naaman Mor, *Eyes on the World*, IEEE Computer Magazine, Vol. 39, No. 10, Oct 2006, en ligne : <http://infolab.stanford.edu/~mor/research/naamanComp06.pdf>

Indexation par le contenu

Aggarwal J.K., Qasim Iqbal “Combining structure, color and texture for image retrieval: a performance evaluation”, Proc of *International Conference on Pattern Recognition (ICPR)*, 2002, Quebec, Canada.

Celentano Augusto, Sciascio Eugenio Di, Di Bari Politecnico, “Similarity Evaluation In Image Retrieval Using Simple Features”, *Proc. of Storage and Retrieval for Image and Video Databases V*, 1997, San Jose (USA), vol. 3022.

Datta Ritendra, Joshi Dhiraj, Li Jia et Wang James Z., *Image retrieval: Ideas, influences, and trends of the new age*, ACM Computing. Surveys, Avril 2008, article 5, 60 pages, en ligne <http://infolab.stanford.edu/~wangz/project/imsearch/review/JOUR/datta.pdf>

Jacobs Chuck, Finkelstein Adam et Salesin David, *Fast Multiresolution Image Querying*, Proceedings of SIGGRAPH 95, Washington, 1995, en ligne <http://grail.cs.washington.edu/projects/query/mrquery.pdf>

Li Jia et Wang James Z., *Real-Time Computerized Annotation of Pictures*, MM'06, October 23–27, 2006, 10 p.

Millet Christophe, *Annotation automatique d'images : annotation cohérente et création automatique d'une base d'apprentissage*, Thèse d'Etat, Telecom Paris, soutenue le 14 Janvier 2008, en ligne <http://chr.millet.googlepages.com/these-christophe-millet.pdf>

Reconnaissance faciale

Zhao W., Chellappa R., Rosenfeld A. & Phillips P.J., *Face Recognition: A Literature Survey*, ACM Computing Surveys, 2003, pp. 399-458

Boon Kee Low, Hjelm^oas Erik, "Face Detection: A Survey", *Computer Vision and Image Understanding*, n°83, 2001, p 236–274, en ligne http://www.sandrosufba.hpg.com.br/inteligencia_artificial/artigos/Hjelmas2001.pdf

Webographie

- **Sites et Blogs scientifiques**

Actualité de la Recherche en Histoire Visuelle : <http://www.arhv.lhivic.org/index.php/>

Affordance.info : http://affordance.typepad.com/mon_weblog/

Du bruit au signal : <http://blog.tuquoque.com/>

Internet Actu : <http://www.internetactu.net/>

- **Sites et Plateformes de partage d'images**

Flickr: <http://www.flickr.com/>

Loc.alize.us: <http://loc.alize.us/>

Zoomr: <http://www.zoomr.com/>

- **Plateformes d'images géotagguées**

Locr : <http://www.locr.com/>

Panoramio : <http://www.panoramio.com/>

- **Moteurs de recherche d'images**

Behold: <http://www.behold.cc/>

Gazopa : http://www.gazopa.com/sign_in

Google Image: <http://images.google.fr/>

Tiltomo: <http://www.tiltomo.com/>

Tineye: <http://tineye.com/>

Xcavator : <http://www.xcavator.net/>

- **Articles en ligne**

Allard Laurence, "L'amateur : une figure de la modernité esthétique", *Lowiki*, 1995
<http://www.culturesexpressives.fr/doku.php>

Chatonsky Gregory, « La révolution a lieu à New-york », *Incident.net*, 03/07/2003, en ligne : <http://incident.net/users/gregory/wordpress/03-la-revolution-a-eu-lieu-a-new-york/>

Douine Rémi, « Extension du domaine de la recherche vidéo », *ARHV*, le 8 décembre 2008, en ligne <http://www.arhv.lhivic.org/index.php/2008/12/08/885-extension-du-domaine-de-la-recherche-video>

Ertzscheid Olivier, « La bibliothèque du congrès et FlickrR : des accès desaxés », *Affordance.info*, janvier 2008, en ligne : http://affordance.typepad.com/mon_weblog/2008/01/la-bibliothque.html

Fievet Cyril, « de recherche vidéo, une nouvelle étape », *InternetActu*, 16 mars 2007, en ligne : <http://www.internetactu.net/2005/03/16/moteurs-de-recherche-vido-une-nouvelle-tape/>

Gros Patrick, « Description et indexation automatiques des documents multimédias : du fantasme à la réalité », *Interstices*, 29/05/07, en ligne : http://interstices.info/jcms/c_19256/description-et-indexation-automatiques-des-documents-multimedias-du-fantasme-a-la-realite

Guillaud Hubert, « Quand YouTube remplacera Google », *InternetActu*, le 11 décembre 2008, en ligne <http://www.internetactu.net/2008/12/11/quand-youtube-remplacera-google/>

Gunthert André, « L'empreinte digitale. Théorie et pratique de la photographie à l'ère numérique », *ARHV*, 3 octobre 2007, en ligne : <http://www.arhv.lhivic.org/index.php/2007/10/03/506-l-empreinte-digitale>

Gunthert André, « Flickr une des choses les plus importantes qui soit arrivée à la photographie », *Internet actu*, en ligne <http://www.internetactu.net/>

Nobis Pierre, Chercher des images par l'image, *Documentation Rouen*, 15 novembre 2008, en ligne : <http://documentaliste.ac-rouen.fr/spip/spip.php?article205>

O Reilly Tim, « What is Web 2.0 ? », *O'Reilly*, 09 septembre 2005, (en ligne <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

O'Reilly Tim, "Why Dell.com (was) more enterprise 2.0 than Dell Ideastorm", *O'Reilly Radar*, 29/09/2008, en ligne <http://radar.oreilly.com/2008/09/why-dell-dot-com-is-more-enterprise.html>

Sinha Rashmi, « À cognitive analysis of tagging. », *Rashmi's blog*, 27 Septembre 2005, en ligne <http://rashmisinha.com/2005/09/27/a-cognitive-analysis-of-tagging/>

Table des Illustrations

Ill. 1 Capture d'écran d'une recherche effectuée sur Google Image à partir du mot-clé LHIVIC.....	28
Ill. 2 Capture d'écran des options de recherche avancée sur Google Image.....	29
Ill. 3 Capture d'écran d'une session de Google Image Labeler.....	31
Ill. 4 Capture d'écran de l'outil Google Audio Indexing	32
Ill. 5 Capture d'écran du nuage de tags de Flickr.....	49
Ill. 6 Capture d'écran du Pool du groupe Flickr BlackandWhite	52
Ill. 7 Capture d'écran de la liste de tags de la photo « Fofie aux fourneaux »	54
Ill. 8 Capture d'écran d'une photographie ayant été tagguée par le groupe « Viremoi ».....	56
Ill. 9 Capture d'écran d'une partie du Mini Game Fastr	59
Ill. 10 Capture d'écran de la galerie du groupe One Letter	60
Ill. 11 Capture d'écran de la galerie de photographies de la Bibliothèque du Congrès sur Flickr.....	64
Ill. 12 Capture d'écran de la galerie d'exemples de photographies fournies par les Institutions participant à The Commons sur Flickr.....	65
Ill. 13 Capture d'écran de la page Creative Commons de Flickr.....	67
Ill. 14 Capture d'écran de l'interface cartographique de Flickr	85
Ill. 15 Capture d'écran de la carte de Zoomr à l'échelle de Paris.....	87
Ill. 16 Capture d'écran de l'outil de géotagging de Picasa sur l'interface Google Earth... ..	88
Ill. 17 capture d'écran de la carte de Panoramio	89
Ill. 18 Capture d'écran du site Locr, recherche de photographies en lien avec la ville de Jaïpur	91
Ill. 19 Capture d'écran du site de Loc.Alize.Us	92
Ill. 20 Capture d'écran de la page Flickr Places dédiée à Paris.....	93
Ill. 21 Capture d'écran de la galerie du groupe Paris 19e arrondissement sur flickr.....	95
Ill. 22 Capture d'écran d'une photographie de Panoramio sur Google Earth	97
Ill. 23 Capture d'écran des informations Wikipedia sur Locr	99
Ill. 24 Capture d'écran de Look Around de Panoramio	105
Ill. 25 Capture d'écran d'une reconstitution de la Tour de Pise en 3D sur Photosynth ...	106
Ill. 26 Capture d'écran de l'interface de recherche par couleur du logiciel QBIC sur le site du musée de l'Hermitage.....	112

III. 27 Capture d'écran d'une recherche effectuée sur Tineye.....	119
III. 28 Capture d'écran d'une recherche effectuée sur Xcavator	120
III. 29 Capture d'écran d'une recherche effectuée sur Tiltomo à partir de la base Catchy Colors	121
III. 30 Capture d'écran d'une recherche effectuée sur Retrievr.....	124
III. 31 Capture d'écran de la page d'accueil du moteur Gazopa.....	125
III. 32 Capture d'écran de la page d'accueil de ALIPR.....	128
III. 33 Capture d'écran d'une recherche effectuée sur Behold	129
III. 34 Capture d'écran de l'application de reconnaissance faciale de Picasa	133
III. 35 Capture d'écran de la page <i>Visages</i> de iPhoto 09	134
III. 37 Capture d'écran de l'outil de reconnaissance faciale dans l'application Flickr	136
III. 38 Capture d'écran du module Cool Searches de Tineye	144
III. 39 Capture d'écran de la présentation de l'option pour I Phone de Gazopa.....	145

Table des Matières

Introduction.....	3
Chapitre 1	
I. Internet : un nouvel espace documentaire.....	11
A. Moteur de recherche et évolution des contenus du Web	11
1. Du Web aux moteurs de recherche	11
2. Extension des espaces documentaires indexables	14
B. Indexation des contenus et publicités contextualisées.....	16
1. Double articulation et indexation textuelle	16
2. Une indexation à visée économique	18
II. L’indexation des documents visuels par les moteurs de recherche	21
A. Accroissement des contenus visuels par le Web 2.0	21
1. Nouvelle approche du document	21
2. L’explosion visuelle du Web	23
B. Les systèmes d’indexation de la firme Google.....	26
1. Chercher le texte avant l’image	26
2. Google Image et Google Image Labeler.....	28
3. Les solutions pour la vidéo	31
Chapitre 2	
I. L’apparition d’une indexation sociale	36
A. L’indexation des collections iconographiques	36
1. Une indexation sémantique surajoutée	36
2. Indexation, classification et thésaurus	38
3. Indexation et métadonnées.....	40
B. La folksonomie	43
1. L’Apparition de la folksonomie sur le Web	43
2. Les principes de la folksonomie	46

3. Une étude de cas : Flickr.....	48
II. Vers les nouveaux usages de la folksonomie	50
A. Les pratiques sociales et communautaires.....	51
1. Des tags aux groupes	51
2. Commenter par les tags.....	53
3. Tags et jugement de goût.....	55
B. Les détournements créatifs de l'indexation.....	57
1. Définition d'un mashup	57
2. Mashups et jeux de mots.....	58
3. Mashups artistiques.....	60
C. Flickr <i>Commons</i> et les nouvelles pratique documentaires	61
1. The <i>Commons</i> sur Flickr.....	62
2. Flickr comme moteur de recherche	66
III. Les limites de la folksonomie	68
A. Polysémie, synonymie et subjectivité.....	68
1. Problèmes de l'indexation sémantique des images.....	68
2. Polysémies des tags	69
B. Problème des multilinguismes et Restriction du champ sémantique	70
1. Régulation par les foules.....	70
2. Un vocabulaire limité, Folksonomie et multilinguisme.....	72
 Chapitre 3	
I. Les présupposés d'une indexation par coordonnées géographiques	76
A. L'indexation par les métadonnées EXIF	76
1. La remise en question de la photographie comme index.....	76
2. La structure informationnelle des images.....	78
3. Les métadonnées comme outil de classement	79
B. La géolocalisation.....	81
1. Les principes	81

2.	Les photographies géolocalisées.....	82
II.	Les mises en application sur le Web.....	84
A.	La géolocalisation dans les sites de partage	84
1.	Flickr.....	84
2.	Zoomr.....	86
3.	Picasa	87
B.	Des exemples de sites dédiés.....	88
1.	Panoramio	88
2.	Locr.....	90
3.	Loc.alize.us	91
III.	Vers un déplacement de la sphère documentaire.....	92
A.	Vers une recherche documentaire incarnée	92
1.	Les groupes Flickr	93
2.	Panoramio ou le nouvel ordre documentaire	96
3.	Echanger son expérience en plus de ses images.....	98
B.	Vers de nouveaux modes de recherche d'images.....	100
1.	Les applications composites : géolocalisation et locative media.....	100
2.	L'image comme interface de navigation.....	103
 Chapitre 4		
I.	La recherche de l'image par l'image	110
A.	La recherche par le contenu.....	110
1.	Les principes	110
2.	L'indexation et la requête	112
B.	Les Caractéristiques symboliques	114
1.	La couleur	114
2.	La texture	116
3.	Autres caractéristiques de l'image.....	116

II. Une observation des sites en lignes	118
A. Les recherche par l'exemple.....	118
1. La requête globale :.....	118
2. La requête partielle :	120
B. Les requêtes par esquisses	122
1. Retrievr	123
2. Gazopa	124
C. L'annotation automatique des images	125
1. Les principes	126
2. Les applications en ligne.....	127
D. Les systèmes de reconnaissance faciale	130
1. Les principes	130
2. Les applications grand public	132
3. Les applications du Web.....	135
III. Les enjeux de la recherche par le contenu	137
A. Le retour du texte ?.....	137
1. Les avantages du « bi-modal »	137
2. Conservation et partage des annotations.....	138
B. Des perspectives à visée commerciales.....	139
1. Pour quels usages dans la sphère de l'utilisateur non pro ?	140
2. La fin du piratage des images ?	142
3. L'enjeu des sites marchands	144
Conclusion.....	148
Bibliographie.....	152
Table des illustrations.....	161