

Diplôme national de master

Domaine - sciences humaines et sociales

Mention - sciences de l'information et des bibliothèques

Spécialités - archives numériques

## **L'indexation des documents iconographiques par les métadonnées internes**

**Fabien Van Den Berghe**

Sous la direction de **Éric Guichard**  
Maître de conférences - ENSSIB



## ***Remerciements***

*Mes remerciements vont en priorité à mon directeur de mémoire, M. Éric Guichard, qui m'a aiguillé vers un domaine d'étude très intéressant et que j'étais bien loin de connaître, et qui a suivi mon travail avec bienveillance et enthousiasme.*

*Je tiens également à remercier M. Patrick Peccatte, qui m'a accordé un entretien de plus d'une heure en janvier 2013, et m'a communiqué à cette occasion de nombreuses informations.*

*Le Capitaine Christophe Cazorla (ECPAD), M. Jeffrey Friedl (créateur de l'outil « Jeffrey's Exif Viewer »), M. Phil Harvey, Mme Anne Meyer (Bibliothèque municipale de Lyon) et M. David Riecks doivent également être remerciés pour les informations qu'ils m'ont envoyées en réponse à mes mails de sollicitation.*

*Enfin, je voudrais remercier M. Philippe Martin (Aproged) et M. Stéphane Pouyllau (CNRS) de m'avoir orienté respectivement vers l'ECPAD et l'archive ouverte MédiHAL, ainsi que M. Michel Jacobson (SIAF) et M. Olivier Rouchon (CINES) pour avoir répondu à mes interrogations.*

### **Résumé :**

*Un fichier image numérique peut contenir des informations sur ce fichier ou l'image qu'il transporte, qui sont appelées métadonnées internes. Le but de ce mémoire est de présenter les principaux standards de métadonnées internes des images fixes, de développer certains exemples d'utilisation, et d'exposer des pistes d'évolution de ces métadonnées, en réfléchissant à leurs champs d'application potentiels. Notre développement portera surtout sur la question de l'indexation des images numériques à partir de leurs métadonnées internes.*

### **Descripteurs :**

Indexation - Image fixe - Image numérique - Document iconographique – Photographie numérique - Métadonnées internes - Métadonnées embarquées – IPTC – IIM (Information Interchange Model) – Exif (Exchangeable Image File) – XMP (Extensible Metadata Platform) – Adobe – CIPA (Camera & Imaging Products Association) – International Press Telecommunications Council – JEITA (Japan Electronics and Information Technology industries Association) – Géolocalisation - Géotagging

### **Abstract :**

*Embedded data in digital pictures files can give some information about a file or about the image it shows. This report aims to present the main formats for embedded metadata in digital pictures, to give some examples of use, to define some of their potential applications, and to present some possible evolutions. Picture indexing is the principal subject of this document.*

### **Keywords :**

Indexing – Digital image – Digital picture – Digital photograph – Embedded metadata – IPTC – IIM (Information Interchange Model) – Exif (Exchangeable Image File) – XMP (Extensible Metadata Platform) – Adobe – CIPA (Camera & Imaging Products Association) – International Press Telecommunications Council – JEITA (Japan Electronics and Information Technology industries Association) – Geolocation - Geolocalization

## *Droits d'auteurs*



Cette création est mise à disposition selon le Contrat :

**Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France**

disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/2.0/fr/> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.



# Sommaire

<b>SIGLES ET ABRÉVIATIONS.....</b>	<b>9</b>
<b>INTRODUCTION.....</b>	<b>11</b>
<b>PREMIÈRE PARTIE : LES MÉTADONNÉES INTERNES AUX IMAGES FIXES ET LEUR CONTEXTE TECHNIQUE .....</b>	<b>19</b>
<b>I. Présentation des métadonnées internes aux images fixes.....</b>	<b>19</b>
1.1. <i>Présentation des standards IPTC/IIM, Exif, XMP, IPTC Core, IPTC Extension et Exif 2.3 metadata for XMP.....</i>	<i>21</i>
<b>II. Comparaisons, champs d'application et conjonction des métadonnées IPTC, Exif et XMP.....</b>	<b>29</b>
2.1. <i>Exif et IPTC.....</i>	<i>29</i>
2.2. <i>IPTC/IIM et XMP.....</i>	<i>30</i>
<b>III. Exploitation des métadonnées .....</b>	<b>31</b>
3.1. <i>Les outils de gestion des métadonnées internes aux images.....</i>	<i>31</i>
3.2. <i>Archivage des métadonnées et indexation.....</i>	<i>33</i>
<b>DEUXIÈME PARTIE : EXEMPLES D'UTILISATION DES MÉTADONNÉES INTERNES AUX IMAGES.....</b>	<b>35</b>
<b>I. Réseaux sociaux et services de partage de photographies en ligne.....</b>	<b>35</b>
1.1. <i>Les services de partage de photographies en ligne.....</i>	<i>37</i>
1.2. <i>Les réseaux sociaux : Facebook et Twitter.....</i>	<i>43</i>
<b>II. Des acteurs de la recherche, de la culture et du patrimoine.....</b>	<b>44</b>
2.1. <i>L'archive ouverte MédiHAL.....</i>	<i>45</i>
2.3. <i>L'ECPAD (Établissement de Communication et de Production Audiovisuelle de la Défense).....</i>	<i>48</i>
2.4. <i>La Bibliothèque municipale de Lyon.....</i>	<i>52</i>
<b>III. La presse, l'édition et les photographes professionnels.....</b>	<b>54</b>
<b>TROISIÈME PARTIE : ÉLÉMENTS SUR LES CHAMPS D'APPLICATION ET SUR LES POSSIBILITÉS D'ÉVOLUTIONS DES MÉTADONNÉES INTERNES AUX IMAGES.....</b>	<b>57</b>
<b>I. Les champs d'application possibles des métadonnées internes aux images .....</b>	<b>57</b>
1.1. <i>L'association étroite entre l'image et ses métadonnées : transport et préservation des informations.....</i>	<i>57</i>
1.2. <i>Indexation et renseignement des métadonnées.....</i>	<i>61</i>
<b>II. Quelques évolutions possibles dans le domaine des métadonnées internes aux images.....</b>	<b>64</b>
2.1. <i>La place des métadonnées internes dans les pratiques relatives aux images : entre désintérêt et pédagogie.....</i>	<i>64</i>
2.2. <i>Regards sur les possibilités d'évolution technique des métadonnées internes aux images .....</i>	<i>67</i>
<b>CONCLUSION.....</b>	<b>71</b>
<b>SOURCES.....</b>	<b>73</b>
<b>BIBLIOGRAPHIE ET WEBOGRAPHIE.....</b>	<b>79</b>
<b>GLOSSAIRE.....</b>	<b>83</b>
<b>TABLE DES ILLUSTRATIONS.....</b>	<b>85</b>

**TABLE DES MATIÈRES.....87**

## *Sigles et abréviations*

**CIPA** Camera & Imaging Products Association

**DICoD** Délégation à l'information et à la communication de la défense

**ECPAD** Établissement de Communication et de Production Audiovisuelle de la  
Défense

**EHESS** École des hautes études en sciences sociales

**Exif** Exchangeable Image File

**ICC** International Color Consortium

**IIM** Information Interchange Model

**IPTC** International Press Telecommunications Council

**JEITA** Japan Electronics and Information Technology industries Association

**XMP** Extensible Metadata Platform



# INTRODUCTION

---

« Une image sans description est une image perdue »<sup>1</sup>. En écrivant ces mots sur son carnet de recherche en ligne, en février 2012, Patrick Peccatte, chercheur associé au Lhivic (Laboratoire d'histoire visuelle contemporaine) de l'EHESS, et spécialiste des métadonnées internes des images, résume la nécessité de documenter un objet iconographique.

L'ensemble des documents qui peuvent aujourd'hui être consultés – sur des supports et pour des usages variés – forme un volume considérable. Dans cet ensemble, une image doit être associée à des informations qui permettent de savoir ce qu'elle est, et de la retrouver : autrement dit, selon les besoins documentaires d'une personne ou d'une organisation donnée, il faut pouvoir connaître ce que montre l'image, d'où elle vient, qui l'a créée, comment, etc., et il faut pouvoir y accéder quand il est à peu près certain qu'elle répond aux besoins du chercheur<sup>2</sup>. Pour cela, et notamment pour la fonction d'accès que nous venons d'évoquer, cette image doit être indexée.

Tout d'abord, il faut s'entendre sur ce que désigne le terme « indexation ».

Selon Amélie Segonds, il s'agit d'une pratique qui consiste à associer à un document ou à un objet (texte, image, vidéo, etc.) des descripteurs relatifs à son contenu<sup>3</sup>. Ce document – ou cet objet<sup>4</sup> – est considéré comme un support d'informations, et c'est à ces informations que se rapportent les descripteurs<sup>5</sup>. Toujours selon Amélie Segonds, l'indexation s'appuie sur des langages documentaires, ce qui signifie que les descripteurs doivent être normalisés, qu'ils

---

<sup>1</sup> PECCATTE, Patrick, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118>] (consulté le 15 novembre 2012 à 22h43)].

<sup>2</sup> Ce terme est employé ici au sens large de chercheur d'information, de chercheur de documents.

<sup>3</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], pp. 3-4. Selon elle, l'indexation doit « 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 » (p. 4).

<sup>4</sup> À cet endroit de notre introduction, nous proposons, par commodité, d'inclure dans la notion de « document » un objet porteur d'information et/ou auquel il soit possible d'associer des informations. Ce choix ne reflète aucunement une prise de position dans le débat autour des limites de la signification du terme « document » en sciences de l'information.

<sup>5</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 4. L'indexation se différencie du catalogage, qui 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.). »

doivent appartenir à un groupe de termes prédéfinis, et être liés entre eux par une certaine syntaxe<sup>6</sup>.

Cette définition correspond à ce que nous avons lu dans d'autres sources<sup>7</sup>. Cependant, il est possible d'envisager le concept d'indexation d'une manière un peu plus large.

Ainsi (d'après l'édition de 2012 de l'*Abrégé d'archivistique*), le *Dictionnaire des archives*, édité par l'École nationale des chartes et l'AFNOR<sup>8</sup>, définit l'indexation comme « l'opération destinée à représenter par les éléments d'un langage documentaire ou naturel des données résultant de l'analyse du contenu d'un document ou du document lui-même » : le seul champ du contenu du document est donc dépassé.

Ghislaine Chartron, dans sa communication du 16 novembre 2004 sur l'indexation des ressources pédagogiques numériques<sup>9</sup>, a présenté trois types d'indexation, qui sont :

- l'indexation qui rend compte du sens du contenu du document ;
- une indexation qui décrit le contexte technique du document – plate-forme, liens avec d'autres objets, formats –, afin de permettre sa réutilisation (cette indexation appartenant au monde de l'informatique<sup>10</sup>) ;
- une indexation centrée sur la « gestion des documents » : cycle de vie, droits relatifs aux documents, différentes versions et exemplaires, etc.

Ainsi, pour Ghislaine Chartron, le terme « indexation » peut recouvrir d'autres pratiques que l'expression du contenu informationnel d'un document selon la définition donnée, par exemple, par Amélie Segonds. Bien que cette typologie à trois facettes soit tirée d'une communication ayant pour sujet une certaine catégorie de documents (les documents pédagogiques numériques), nous pensons que cette vision de l'indexation peut être appliquée à d'autres domaines documentaires. Nous

---

<sup>6</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 38.

<sup>7</sup> Voir par exemple LÉVESQUE, Michel. « L'indexation : luxe ou nécessité ? ». *Archives*, volume 33, n°1, 2001-2002, pp. 17-45 [En ligne : [http://www.archivistes.qc.ca/revuearchives/vol33\\_1/33-1-levesque.pdf](http://www.archivistes.qc.ca/revuearchives/vol33_1/33-1-levesque.pdf) ], notamment pp. 33-34.

<sup>8</sup> *Dictionnaire des archives. De l'archivage aux systèmes d'information*. Paris : École nationale des chartes / AFNOR, 1991 ; cité dans : Association des archivistes français. *Abrégé d'archivistique. Principes et pratiques du métier d'archiviste*. Paris : Association des archivistes français, 2012. 346 p., p. 180.

<sup>9</sup> *L'indexation des ressources pédagogiques numériques : questions transversales*. Compte-rendu rédigé par l'ENSSIB à partir d'une transcription de la communication orale de Ghislaine CHARTRON du 16 novembre 2004 [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-1241>], pp. 4-5. Il s'agit de la seule ressource que nous ayons trouvée qui propose une vision de l'indexation à la fois précise et plus large que la seule restitution du contenu d'un document.

<sup>10</sup> *Ibid.*, p. 4 : « l'indexation d'un contenu pour sa réutilisation sur différentes plates-formes appartient au monde de l'informatique ».

avons donc choisi, dans notre développement, d'adopter une définition très large, en considérant l'indexation comme la manière de restituer un ensemble d'aspects relatifs à un document/objet – son contenu documentaire, son histoire, ses caractéristiques techniques, etc. – pour donner accès à cette ressource<sup>11</sup>.

L'indexation contribue à alimenter un ensemble plus large de renseignements relatifs à un document et à son contenu informationnel : les descripteurs qu'elle produit font en effet partie des « métadonnées » d'un document, que nous allons maintenant tenter de définir.

D'une manière très simple, on peut définir une métadonnée comme une donnée qui fournit une information sur une ou plusieurs autre(s) donnée(s) – après avoir été comprise par une machine ou un humain.

Pour expliquer ce qu'est une métadonnée, Sylvie Dalbin cite notamment la norme expérimentale XP X50-435 *Management des systèmes – Gestion documentaire – Concepts généraux* (septembre 1995) en donnant la définition suivante : « représentation réinterprétable d'une information, sous forme conventionnelle convenant à la communication, à l'interprétation ou au traitement », sachant que le terme « réinterprétable » signifie ici que cette représentation « n'est, en principe, pas utilisable en l'état »<sup>12</sup>. La métadonnée est codifiée, comme la donnée qu'elle documente, et anticipe l'utilisation de cette donnée<sup>13</sup>. Patrick Peccatte définit plus simplement la métadonnée comme un « ensemble structuré d'informations décrivant une ressource quelconque »<sup>14</sup>.

<sup>11</sup> Nous employons le mot « ressource », sauf mention contraire, comme synonyme de « document », avec la signification large expliquée en note 4, page 11.

<sup>12</sup> DALBIN, Sylvie. « Représentation et accès à l'information : transformation à l'œuvre », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p. [En ligne : [http://www.atd-doc.com/xmedia/publications/INRIA-2008\\_DALBIN\\_Chap01.pdf](http://www.atd-doc.com/xmedia/publications/INRIA-2008_DALBIN_Chap01.pdf)], p. 11. La référence citée pour cette définition est la norme expérimentale XP X50-435 *Management des systèmes – Gestion documentaire – Concepts généraux* (septembre 1995).

<sup>13</sup> DALBIN, Sylvie. « Représentation et accès à l'information : transformation à l'œuvre », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p. [En ligne : [http://www.atd-doc.com/xmedia/publications/INRIA-2008\\_DALBIN\\_Chap01.pdf](http://www.atd-doc.com/xmedia/publications/INRIA-2008_DALBIN_Chap01.pdf)], p. 12.

<sup>14</sup> PECCATTE, Patrick. « Métadonnées : une initiation. Dublin Core, IPTC, Exif, RDF, XMP, etc. ». *Soft Experience* [site web]. Août 2002, mis à jour le 13 décembre 2007 [En ligne : <http://peccatte.karefil.com/software/metadata.htm> (consulté le 10 juin 2013 à 10h40)].

Pour le terme « ressource », Patrick Peccatte renvoie à la page de *Wikipédia* [site web] « Ressource du World Wide Web », où il est expliqué que le concept de ressource a évolué, durant l'histoire du Web, « vers une définition plus générale et abstraite [que la notion primitive], recouvrant maintenant toute chose ou entité susceptible d'être identifiée, nommée, manipulée à travers ses représentations, par quelque moyen que ce soit, sur le Web en général ou dans n'importe quel système d'information utilisant les technologies du Web » [En ligne : [http://fr.wikipedia.org/wiki/Ressource\\_%28Web%29](http://fr.wikipedia.org/wiki/Ressource_%28Web%29) (consultée le 10 juin 2013 à 10h47)].

Une métadonnée peut aussi bien se rapporter à une ressource numérique qu'à un document papier<sup>15</sup>.

Pour expliquer davantage le rôle des métadonnées, on peut reprendre l'exemple des étiquettes sur des canettes dans un supermarché<sup>16</sup>: si ces canettes n'étaient pas étiquetées, il serait nécessaire de les ouvrir pour voir ce qu'elles contiennent ; alors qu'avec des étiquettes, le potentiel consommateur sait quel est le produit, avec quoi il est fabriqué, d'où il vient, éventuellement comment l'utiliser, etc.

Les métadonnées peuvent être associées plus ou moins étroitement à la ressource à laquelle elles se rapportent : elles peuvent se trouver dans un fichier autonome, dans une base de données, ou être intégrées totalement à la ressource<sup>17</sup>. Dans ce dernier cas de figure, pour les documents numériques, les métadonnées sont dites « internes » ou « embarquées » (*embedded*, en anglais) : cela signifie qu'elles se trouvent à l'intérieur même du fichier sur lequel elles fournissent des informations.

Le sujet de ce mémoire étant l'indexation des images fixes (sous forme numérique, obligatoirement) par leurs métadonnées internes, il convient de s'arrêter sur ce qui fait la spécificité et la complexité des images du point de vue de leurs métadonnées et de leur indexation.

Les images présentent, au-delà même de la sphère numérique, des caractéristiques qui rendent difficile leur description – par comparaison, notamment, avec un texte. (Dans cette introduction, quand nous parlerons du contenu informationnel de l'image, nous nous concentrerons uniquement sur l'image en tant que telle, en faisant abstraction de son support).

Le recours au texte semble la plupart du temps indispensable pour documenter une image et lui donner accès. Ainsi, Michel Melot a expliqué que « l'identification des images, même si ces dernières gagnent de plus en plus d'autonomie par rapport au texte, ne peut se passer de légende dès qu'on cherche

---

<sup>15</sup> FOULONNEAU, Muriel, RILEY, Jenn. *Metadata for digital resources : implementation, systems design and interoperability*. Oxford : Chandos Publishing, 2008. 203 p., p. 4.

<sup>16</sup> Exemple cité dans FOULONNEAU, Muriel, RILEY, Jenn. *Metadata for digital resources : implementation, systems design and interoperability*. Oxford : Chandos Publishing, 2008. 203 p., p. 4. La référence donnée pour cet exemple est : BOURDA, Y. « Définitions et concepts ; un exemple du processus de normalisation : les métadonnées pour les ressources pédagogiques ». Exposé au séminaire « Normes et standards pour les activités numériques dans l'enseignement ». Lyon : 9 octobre 2003.

<sup>17</sup> DALBIN, Sylvie. « Représentation et accès à l'information : transformation à l'œuvre », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p., p. 25.

la moindre précision »<sup>18</sup>. Nous sommes d'accord avec cette affirmation. Premièrement, il est difficile de rendre compte d'aspects techniques ou de gestion avec d'autres outils que du texte. De plus, en ce qui concerne l'indexation du contenu informationnel de l'image, il est difficile d'être précis, ou suffisamment abstrait pour certains concepts. Une normalisation d'éléments iconographiques, destinée à permettre une recherche d'images par des images uniquement, nous semble possible dans certains cas, mais est probablement extrêmement délicate la plupart du temps : par exemple, comment exprimer autrement que par du texte le concept « monument » – détaché de tout exemple précis de monument – ? Comment rechercher la photographie d'une personne particulière, sans utiliser son nom sous forme textuelle<sup>19</sup> ?...

Il ne semble donc pas possible de se passer de texte pour trouver des documents iconographiques correspondant à certains critères, ou pour obtenir certains types de renseignements sur des images. Cependant, pour l'indexation du contenu informationnel de ces documents, il est absolument impossible de procéder à une extraction de mots-clés : il faut donc opérer une transcription de ce contenu sous forme de texte. Cela nous amène à évoquer d'autres difficultés : comment restituer par le texte le contenu d'une image ?

L'indexation du contenu informationnel d'une image pose, semble-t-il, plus de problèmes que l'indexation d'un texte, notamment car :

- elle doit restituer plusieurs niveaux de lecture (description de ce qui se trouve sur l'image, description de ce qu'évoque l'image, ...) <sup>20</sup>
- il est difficile (voire impossible) d'épuiser le contenu informationnel d'une image<sup>21</sup>
- une partie de la description de ce contenu (celle qui concerne ce qu'évoque l'image), peut, *a priori*, fortement varier d'un individu à l'autre suivant son ressenti, son interprétation de l'image.

Nous n'irons pas plus loin sur ces problèmes d'indexation dans notre introduction : notre but était de rappeler, si cela était nécessaire, que l'indexation des images peut constituer un sujet d'étude à part entière.

<sup>18</sup> COLLARD, Claude (dir.), MELOT, Michel (dir.). *Images et bibliothèques*. Paris : Éditions du Cercle de la librairie, 2011. 240 p., p. 11.

<sup>19</sup> Exemple tiré de la référence citée dans la note précédente, pp. 11-12.

<sup>20</sup> PECCATTE, Patrick. « Description des images et rhétorique documentaire ». *Déjà vu. Carnet de recherche visuel* [site web]. 30 juin 2010 [En ligne : <http://culturevisuelle.org/dejavu/255> (consulté le 17 novembre à 16h45)].

CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 33.

<sup>21</sup> COLLARD, Claude (dir.), MELOT, Michel (dir.). *Images et bibliothèques*. Paris : Éditions du Cercle de la librairie, 2011. 240 p., pp. 11-12.

Les images numériques se distinguent des autres documents numériques par les formats de fichiers qui leur sont dédiés. Les métadonnées internes de ces fichiers sont enregistrées conformément à des modèles spécifiques aux fichiers images. Patrick Peccatte, récemment<sup>22</sup>, considérait qu'il existait essentiellement trois catégories de métadonnées internes pour les images fixes (JPEG ou TIFF) :

- les informations IPTC/IIM
- les informations Exif
- les informations XMP.

Il s'agit de standards, c'est-à-dire (au sens informatique) de « spécification[s] documentée[s] permettant de développer des produits interopérables »<sup>23</sup>.

Ce mémoire aborde l'utilisation des métadonnées internes pour l'indexation des images numériques, en s'intéressant à la fois aux usages actuels de ces métadonnées et à leurs utilisations potentielles – en prenant en compte leurs possibilités d'évolution. Notre propos se concentrera surtout sur les informations IPTC (IIM, Core et Extension), Exif et XMP, et s'articulera en plusieurs segments. Dans un premier temps, nous présenterons les métadonnées internes des images, et notamment les standards que nous venons de citer, en développant de manière succincte leur histoire, leurs particularités et leur contexte technique. Nous exposerons ensuite quelques exemples d'utilisations de ces métadonnées, en particulier pour l'indexation des photographies. Cela nous permettra de terminer avec une partie consacrée aux champs d'applications possibles de ces métadonnées et à des possibilités d'évolution des standards mis en jeu.

Nous n'avons pas cherché à approfondir le fonctionnement technique des métadonnées qui nous intéressent, ce qui nous aurait sans doute demandé beaucoup plus de temps que nous n'en avons pour finir ce mémoire ; aussi les spécifications des formats Exif, XMP et IPTC (IIM, Core et Extension) n'ont-elles été que peu (ou pas du tout) consultées et exploitées. Le lecteur ne trouvera pas non plus dans ce document de présentation de tous les types d'utilisation qui sont faits ou qui peuvent être faits de ces métadonnées : seuls quelques exemples ont été cités, avec pour objectif de faire intervenir des acteurs les plus variés possibles, sans que nous

<sup>22</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 1<sup>er</sup> décembre 2012 à 19h15)].

<sup>23</sup> PECCATTE, Patrick, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

puissions parvenir à l'exhaustivité. Enfin, bien que les métadonnées citées puissent parfois s'appliquer à d'autres documents qu'à des images, nous nous sommes presque exclusivement concentré sur les aspects relatifs aux documents iconographiques.

Nous serons souvent amené à citer Patrick Peccatte, que nous avons présenté au début de cette introduction : cette personne est en effet l'un des spécialistes de notre sujet les plus visibles et les plus cités, en tout cas sur le Web francophone et plus généralement dans les sources rédigées en français. De plus, il nous a accordé un entretien en janvier 2013, au cours duquel il nous a communiqué de nombreuses informations que nous avons reprises, pour la plupart, dans ce mémoire.

Pour la partie consacrée aux exemples d'utilisation effective des métadonnées, nous avons réalisé quelques tests sur des services présents sur le Web, à l'aide de l'outil « Jeffrey's Exif Viewer », accessible en ligne à l'adresse : <http://regex.info/exif.cgi><sup>24</sup>. Cet instrument de travail, créé par Jeffrey Friedl<sup>25</sup>, permet d'afficher les métadonnées contenues dans des fichiers images, et est notamment utilisé par David Riecks, qui a participé à l'élaboration des formats IPTC Core et IPTC Extension<sup>26</sup>.

Nous n'avons malheureusement pas toujours pu, au cours de nos recherches, identifier précisément le standard IPTC (IIM, Core ou Extension) mentionné dans nos sources. Nous nous sommes cependant efforcé, dans notre développement, d'être le plus précis possible à ce sujet.

Comme certaines de nos sources datent d'il y a un an ou plus, nous avons fait le choix d'indiquer, lorsque nous n'avons pas pu vérifier une information, et faute de pouvoir être plus précis, que notre propos devait être exact au moins à la date associée au document dans lequel nous avons trouvé cette information (ce qui, donc, peut ne pas être totalement vrai, puisqu'il peut s'agir (notamment) d'une date de publication).

---

<sup>24</sup> (Lien vérifié le 10 juin 2013 à 14h24).

<sup>25</sup> Voir la présentation de cette personne sur son blog : « About me ». *Jeffrey Friedl's Blog* [site web] [En ligne : <http://regex.info/blog/jfriedl-links/about> (lien vérifié le 10 juin 2013 à 14h30)].

<sup>26</sup> « about Controlled Vocabulary : ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/aboutus.html> (consulté le 10 juin 2013 à 14h33)]. David Riecks est le fondateur du site *Controlled Vocabulary*.



# PREMIÈRE PARTIE : LES MÉTADONNÉES INTERNES AUX IMAGES FIXES ET LEUR CONTEXTE TECHNIQUE

---

## I. PRÉSENTATION DES MÉTADONNÉES INTERNES AUX IMAGES FIXES

Les métadonnées internes d'un document iconographique numérique sont contenues dans le fichier qui sert de support à cette image : cela veut dire que, à l'intérieur de ce fichier, on trouve des informations textuelles, susceptibles d'être lues par un être humain.

Prenons un exemple, en utilisant un outil disponible sur le Web, « Jeffrey's Exif Viewer »<sup>27</sup>, qui permet d'afficher les métadonnées internes d'une image :

- Voici une photographie, prise avec un appareil photo numérique:



*Illustration 1: Photographie testée  
avec l'outil « Jeffrey's Exif Viewer »  
(photographie de l'auteur)*

---

<sup>27</sup> En ligne : <http://regex.info/exif.cgi> (lien vérifié le 10 juin 2013 à 14h24).

- Voici maintenant une capture d'écran montrant une partie des métadonnées internes de cette photographie, révélées grâce à l'outil utilisé<sup>28</sup> (en l'occurrence, il s'agit d'informations Exif contenues dans le fichier image):

**EXIF** — this group of metadata is encoded in 11,820 bytes (11.5k)

Exif Image Size	2,816 × 2,112
Make	Panasonic
Camera Model Name	DMC-LS65
Orientation	Horizontal (normal)
Software	Ver.1.0
Modify Date	<b>2008:10:12 16:32:18</b> 4 years, 7 months, 1 day, 20 hours, 46 minutes, 43 seconds ago
Y Cb Cr Positioning	Co-sited
Exposure Time	1/160
F Number	4.30
Exposure Program	Program AE
ISO	100
Exif Version	0221
Date/Time Original	<b>2008:10:12 16:32:18</b> 4 years, 7 months, 1 day, 20 hours, 46 minutes, 43 seconds ago
Components Configuration	Y, Cb, Cr, -
Compressed Bits Per Pixel	4
Exposure Compensation	0
Max Aperture Value	2.8
Metering Mode	Multi-segment
Light Source	Unknown
Flash	Auto, Did not fire

*Illustration 2: Métadonnées de la photographie présentée en Illustration 1 (page précédente), affichées par « Jeffrey's Exif Viewer ». Capture d'écran.*

Les informations internes aux images peuvent porter sur des caractéristiques techniques propres à l'image qui les contient (c'est le cas sur notre capture d'écran), mais aussi, par exemple, sur le titre ou l'auteur du document, voire sur des descripteurs et des indications de propriété intellectuelle. Cette liste n'est pas exhaustive : d'autres types d'informations internes seront signalés plus loin.

Dans notre développement, nous nous intéressons surtout (pour les raisons données en introduction) aux standards de métadonnées internes que sont l'IPTC/IIM, l'IPTC Core, l'IPTC Extension, le standard Exif et le standard XMP.

L'inclusion de métadonnées internes dans une image est une pratique apparue essentiellement avec le standard IPTC/IIM, qui doit beaucoup son succès aux agences photographiques, pour lesquelles il était fondamental d'avoir la description

<sup>28</sup> Essai réalisé le 14 mai 2013.

de l'image et l'image elle-même en même temps chez les clients – en l'occurrence, la presse<sup>29</sup>.

## **1.1. Présentation des standards IPTC/IIM, Exif, XMP, IPTC Core, IPTC Extension et Exif 2.3 metadata for XMP**

### ***A. Les métadonnées IPTC/IIM***

L'International Press Telecommunications Council (IPTC) est un organisme créé en 1965 par plusieurs organisations dans le domaine de la presse. L'IPTC s'occupe prioritairement de développer, publier et maintenir des standards pour l'échange de données de presse<sup>30</sup>. Actuellement, il regroupe plus d'une soixantaine d'entreprises, organisations et associations du secteur de la presse, dont l'Agence France-Presse (AFP), Associated Press (AP), Reuters... et d'autres membres de toutes les parties du monde, excepté l'Océanie et l'Amérique du Sud (on ne trouve que deux pays africains représentés, l'Égypte et l'Angola)<sup>31</sup>.

Le plus ancien standard de l'IPTC pour les métadonnées est l'Information Interchange Model<sup>32</sup>, connu également sous le nom d'IPTC/IIM, développé avec la NAA (The Newspaper Association of America), association des éditeurs de presse américains<sup>33</sup>. On parle parfois pour désigner l'IPTC/IIM (ou IIM) de « standard IPTC », mais il s'agit d'un abus de langage : en effet, il existe d'autres standards publiés par l'IPTC. (Nous n'avons malheureusement pas toujours pu, au cours de nos recherches, identifier précisément le standard IPTC mentionné dans nos sources. Nous nous sommes cependant efforcé, dans notre développement, d'être le plus précis possible à ce sujet).

---

<sup>29</sup> Ces informations nous ont été communiquées par Patrick PECCATTE lors de l'entretien qu'il nous a accordé le 23 janvier 2013.

<sup>30</sup> « The IPTC... as a membership organisation ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/site/Home/About/> (consulté le 11 mai 2013 à 14h39)].

CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., pp. 35-36.

<sup>31</sup> « The IPTC... as a membership organisation ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/site/Home/About/> (consulté le 11 mai 2013 à 14h39)].

« IPTC Members ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/site/Home/Members/> (consulté le 11 mai 2013 à 15h)].

<sup>32</sup> PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu . Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre 2012 à 14h58)].

<sup>33</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., pp. 35-36.

Lancé en 1991, l'IIM a connu plusieurs versions, dont la dernière en date est la spécification 4.1 de 1999<sup>34</sup>. Au départ, ce format était destiné à s'appliquer à tous les types de données utilisables dans l'univers des médias et de la presse – texte, image ou multimédia<sup>35</sup> – ; mais son utilisation s'est élargie, pour l'image, au-delà du domaine de la presse<sup>36</sup>.

L'IPTC/IIM a été intégré dans de nombreux logiciels de gestion d'images, le premier d'entre eux à l'avoir adopté étant Photoshop<sup>37</sup>, qui a inclus, en 1994, un certain nombre de champs de l'IIM. Ce sous-ensemble a été pris en charge dans les formats JPEG, TIFF et PSD<sup>38</sup>. Selon Patrick Peccatte<sup>39</sup>, c'est l'intégration des informations IPTC/IIM dans Photoshop qui a contribué au succès de ce standard, lui permettant d'acquérir une grande importance dans l'industrie des agences de presse. L'utilisation de ce format de métadonnées et son impact vont cependant au-delà de l'univers où il est né. Pour Patrick Peccatte, le moment où Adobe a adopté un sous-ensemble de l'IIM a permis d'enclencher le développement de l'utilisation des métadonnées internes aux images (d'autres initiatives, autres que celles de l'IPTC, n'ont pas abouti)<sup>40</sup>.

D'autres formats de métadonnées IPTC ont été créés pour adopter le formalisme XMP développé par Adobe : nous en parlons plus loin dans cette partie. Néanmoins, les champs IPTC/IIM restent toujours très utilisés par les professionnels, y compris par Adobe<sup>41</sup>.

---

<sup>34</sup> « IIM Information Interchange Model ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0134> (consulté le 2 décembre 2012 à 15h28)].

<sup>35</sup> « IIM Information Interchange Model. IIM for photos : a good solution for more than a decade ». *International Press Telecommunications Council* [site web] [En ligne : [http://www.iptc.org/site/Photo\\_Metadata/IIM/](http://www.iptc.org/site/Photo_Metadata/IIM/) (consulté le 10 juin 2013 à 22h12)].

CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., pp. 35-36.

<sup>36</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., pp. 35-36.

<sup>37</sup> *Ibid.*, pp. 35-36.

<sup>38</sup> PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre 2012 à 14h58)].

<sup>39</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>40</sup> *Ibid.*

<sup>41</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 13h58)].

## B. Les métadonnées Exif

Le format Exif (Exchangeable Image File) a été développé en 1995 par le JEIDA (Japan Electronic Industry Development Association<sup>42</sup>), qui est devenu ensuite le JEITA (Japan Electronics and Information Technology industries Association)<sup>43</sup>. Le JEITA s'est donné pour objectif de contribuer au développement des industries de l'électronique et des technologies de l'information (« IT ») du Japon, en encourageant notamment le commerce international et la consommation de produits et de composants électroniques<sup>44</sup>.

Selon Patrick Peccatte, le format Exif est un « standard de fait », car tous les fabricants d'appareils numériques l'ont adopté – avec, cependant, autour d'un noyau commun assez important, des variantes propriétaires suivant les constructeurs<sup>45</sup>. La spécification la plus récente de ce format est apparemment la version « Exif Unified 2.3 » d'avril 2010, élaborée conjointement par le JEITA et le CIPA (Camera & Imaging Products Association)<sup>46</sup>, organisme regroupant la plupart des fabricants japonais d'appareils photographiques numériques<sup>47</sup>.

Les métadonnées Exif sont désignées dans certaines sources comme des informations d'ordre « technique »<sup>48</sup>. Presque toutes les métadonnées Exif sont renseignées automatiquement à la prise de vue, par l'appareil photo<sup>49</sup> – par exemple, les coordonnées GPS correspondant à l'endroit où la photographie a été prise, si l'appareil est équipé d'un GPS.

---

<sup>42</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 44.

<sup>43</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 13h58)].

<sup>44</sup> « JEITA's Activities and Organization. What is JEITA ? » *JEITA* [site web] [En ligne : <http://www.jeita.or.jp/english/about/what/index.htm> (consulté le 10 juin 2013 à 22h45)].

<sup>45</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013, et PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 13h58)].

<sup>46</sup> JEITA. *Exchangeable image file format for digital still cameras : Exif Unified Version 2.3*. Préparé par : AV&IT Standardization Committee. Japan Electronics and Information Technology Industries Association. Révisé en avril 2010 [En ligne : [http://www.jeita.or.jp/japanese/standard/book/CP-3451B\\_E/#page=1](http://www.jeita.or.jp/japanese/standard/book/CP-3451B_E/#page=1) (lien vérifié le 11 juin 2013 à 9h50)]. Les informations ont notamment été collectées à la deuxième page et aux pages 3-4, où est présenté un historique des révisions du format.

Cette spécification est consultable sur le site du JEITA à partir de la page « JEITA Standards. AV&IT Technology Standardization. Digital Cameras » [En ligne : [http://www.jeita.or.jp/cgi-bin/standard\\_e/list.cgi?cateid=1&subcateid=4](http://www.jeita.or.jp/cgi-bin/standard_e/list.cgi?cateid=1&subcateid=4) (consultée le 11 juin 2013 à 0h07)].

<sup>47</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 15h10)].

<sup>48</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 13h58)].

CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 44.

<sup>49</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 13h58)].

Il est vrai que de nombreux champs stockés en Exif peuvent concerner des éléments techniques relatifs à la prise de vue (temps d'exposition, informations sur l'utilisation du flash, appareil utilisé, ...), au fichier généré (par exemple, à sa compression) ou aux traitements opérés sur la photographie après la prise de vue (logiciel utilisé, date des modifications...). Cependant, il est possible de renseigner des champs Exif qui concernent le titre de l'image, l'auteur et le titulaire des droits d'auteur, ainsi que des informations de description de l'image. Certains champs Exif sont d'ailleurs équivalents à des champs IPTC (IIM et Core) : ainsi, par exemple, le champ Exif « Artist » correspond aux champs IPTC/IIM « By-line » et IPTC Core « Creator », et le champ Exif « Copyright » au champ IPTC (Core et IIM) « Copyright Notice »<sup>50</sup>.

### C. Les métadonnées XMP

C'est en septembre 2001 qu'Adobe a présenté le format XMP (eXtensible Metadata Platform). Ce type de métadonnées utilise le langage à balises XML ainsi que le schéma de métadonnées du Dublin Core comme fondation<sup>51</sup>, et repose sur une version simplifiée de RDF. Il est devenu une norme ISO (ISO 16684-1:2012) – à ce titre, il ne doit pas être considéré comme un format propre à Adobe<sup>52</sup>.

Avant d'aller plus loin, peut-être faut-il rappeler brièvement ce que sont XML, le Dublin Core et RDF :

- XML<sup>53</sup> (eXtensible Markup Language) est un langage à balises, recommandé par le W3C. Il permet de structurer des informations textuelles grâce à des balises. Cette structuration, qui obéit à des principes simples, permet un transport et une conservation de l'information.

---

<sup>50</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 44.

RIECKS, David. *"IPTC Core" Schema for XMP. Version 1.0. Supplemental documentation. Custom Panels User Guide. Document Revision 13*. 2005 [En ligne : [http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore\\_1.0-doc-CpanelsUserGuide\\_13.pdf](http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore_1.0-doc-CpanelsUserGuide_13.pdf)], p. 20.

<sup>51</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 47.

<sup>52</sup> PECCATTE, Patrick, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

Le site de l'IPTC indique qu'Adobe possède un droit de propriété intellectuelle sur le format XMP (« IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consulté le 3 décembre 2012 à 18h51)].

<sup>53</sup> « Introduction to XML ». *W3schools.com* [site web] [En ligne : [http://www.w3schools.com/xml/xml\\_what.asp](http://www.w3schools.com/xml/xml_what.asp) (consulté le 1<sup>er</sup> décembre 2012 à 16h18)].

– Le Dublin Core<sup>54</sup> est un jeu de métadonnées défini en 1995, et qui est devenu une norme ISO en 2003 (la norme ISO 15836:2003). Ce format avait notamment pour objectifs d'être simple à gérer, d'avoir une sémantique largement comprise, d'être utilisable avec HTML et XML, d'être applicable au nombre maximum de formats, et de prendre une envergure internationale. Il devait être interprétable à la fois par des machines et par des humains.

Le Dublin Core est composé de quinze éléments de base, optionnels et répétables, qui concernent le contenu (« Title », « Description », « Subject », « Source », « Coverage », « Type » (catégorie de ressources), « Relation »), la propriété intellectuelle (« Creator », « Contributor », « Publisher », « Rights ») et la version du document (« Date », « Format », « Identifier », « Language »). Il est possible de procéder à des ajouts à cet ensemble d'informations.

– RDF<sup>55</sup> (Resource Description Framework) est un standard publié dans sa première version en 1999 et finalisé par le W3C en 2004. Son but est de décrire des ressources, en proposant un modèle basé sur des triplets.

Ces triplets sont constitués comme ceci :

« sujet, prédicat, objet »

où le « sujet » correspond à la ressource à décrire, le « prédicat » à une propriété applicable à ce sujet, et l'« objet » à une valeur de cette propriété.

RDF utilise des URI pour identifier le sujet et le prédicat d'un triplet, l'objet (valeur de la propriété exprimée par le prédicat) pouvant être une ressource ou une donnée. Ce standard est conçu pour permettre la réutilisation de vocabulaires définis indépendamment les uns des autres, grâce aux URI.

RDF n'est pas lié à une syntaxe en particulier. Cependant, la plupart des vocabulaires standards tels que RDFS, OWL, SKOS, FOAF, etc., sont publiés en RDF/XML (c'est-à-dire RDF exprimé avec une syntaxe XML).

---

<sup>54</sup> DALBIN, Sylvie. « Métadonnées et normalisation », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p., pp. 127-128.

SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 41.

<sup>55</sup> VATANT, Bernard. « Des métadonnées à la description des ressources : Les langages du web sémantique », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p., pp. 164, 167 et 172.

Pour résumer, XMP est basé sur RDF, pour le formalisme, et sur le Dublin Core pour la signification des champs.<sup>56</sup>

Avec XMP, toutes les métadonnées telles que les informations IPTC ou les données Exif des appareils photographiques, ainsi que les métadonnées personnalisées, peuvent être enregistrées<sup>57</sup>. XMP est d'ailleurs un format extensible, ce qui signifie que l'utilisateur peut définir ses propres schémas et métadonnées (en respectant le formalisme RDF)<sup>58</sup>. Des schémas peuvent être utilisés pour étendre le Dublin Core en XMP : la plupart sont fournis par Adobe, mais ils peuvent également provenir d'autres organismes, fabricants, industries, etc.<sup>59</sup>

Les métadonnées enregistrées en XMP sont stockées soit à l'intérieur du fichier (JPEG, TIFF, PSD, DNG ainsi que d'autres types de fichiers, y compris les PDF<sup>60</sup>), soit dans des fichiers XMP extérieurs à l'image – ce qui peut être le cas, par exemple, pour les fichiers Raw.

XMP a été introduit dans tous les logiciels d'Adobe<sup>61</sup>.

#### ***D. Les métadonnées IPTC Core et IPTC Extension***

L'IPTC Core Schema a été développé grâce à une collaboration entre Adobe et l'IPTC, et approuvé par l'IPTC en 2004<sup>62</sup>. Le schéma IPTC Core a été conçu pour transférer les valeurs des champs IPTC basés sur l'IIM dans le formalisme XMP. Il inclut une spécification qui explique comment synchroniser les valeurs des champs IPTC/IIM avec l'IPTC Core, et vice-versa<sup>63</sup>. Certains champs de l'IIM ont

---

<sup>56</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>57</sup> PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre 2012 à 14h58)].

<sup>58</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013, et CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 47.

<sup>59</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>60</sup> PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre 2012 à 14h58)].

<sup>61</sup> *Ibid.* (consulté le 17 novembre 2012 à 14h58).

<sup>62</sup> « IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consulté le 3 décembre 2012 à 18h51)].

<sup>63</sup> *Ibid.* (consulté le 3 décembre 2012 à 18h51).

tout de même été supprimés, et d'autres ont été ajoutés<sup>64</sup> pour améliorer la pertinence du standard pour la gestion de l'image au-delà du monde de la presse<sup>65</sup>.

L'IPTC Core devait prendre la suite de l'IIM, désormais considéré comme obsolète<sup>66</sup> – le développement des métadonnées IIM a d'ailleurs été stoppé définitivement par l'IPTC<sup>67</sup>. La version 1.0 de l'IPTC Core a été annoncée officiellement en mars 2005<sup>68</sup>. La version 1.1 est la plus récente, et a été approuvée en juillet 2008<sup>69</sup>.

Selon Patrick Peccatte<sup>70</sup>, la technologie XMP présente plusieurs avantages par rapport à l'IPTC/IIM :

- la taille des champs est illimitée
- les problèmes d'accents sont résolus
- l'intégration de légendes multilingues devient possible
- les métadonnées sont extensibles et personnalisables.

L'IPTC Extension est un schéma qui vient compléter l'IPTC Core avec des champs qui concernent notamment :

- le contenu de l'image (ce qui est montré sur l'image : lieu, événement, etc.)
- un objet montré sur la photographie (titre, créateur, date de création, etc.)
- la propriété intellectuelle, et la licence d'utilisation correspondant à l'image.

L'IPTC Extension 1.1. a été approuvé en juin 2009<sup>71</sup>.

---

<sup>64</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 48.

RIECKS, David. *"IPTC Core" Schema for XMP. Version 1.0. Supplemental documentation. Custom Panels User Guide. Document Revision 13*. 2005 [En ligne : [http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore\\_1.0-doc-CpanelsUserGuide\\_13.pdf](http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore_1.0-doc-CpanelsUserGuide_13.pdf)], p. 20.

<sup>65</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 48.

<sup>66</sup> PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre 2012 à 14h58)].

<sup>67</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 48.

<sup>68</sup> *Ibid.*, p. 48.

<sup>69</sup> « IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consulté le 11 juin 2013 à 14h39)].

<sup>70</sup> PECCATTE, Patrick. « Métadonnées : une initiation. Dublin Core, IPTC, Exif, RDF, XMP, etc. ». *Soft Experience* [site web]. Août 2002, mis à jour le 13 décembre 2007 [En ligne : <http://peccatte.karefil.com/software/metadata.htm> (consulté le 11 juin 2013 à 14h27)].

<sup>71</sup> « IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consulté le 8 juin 2013 à 16h18)].

L'IPTC Core et l'IPTC Extension utilisent la technologie XMP d'Adobe<sup>72</sup>, et sont publiés dans la spécification *IPTC Standard Photo Metadata*, dont la dernière version date de juillet 2010<sup>73</sup>.

### ***E. Les métadonnées Exif 2.3 metadata for XMP***

Le 30 janvier 2012, un nouveau standard, l'Exif 2.3 metadata for XMP, a été publié par le CIPA (Camera & Imaging Products Association)<sup>74</sup>. Ce standard s'appuie sur un principe déjà mis en œuvre dans les versions récentes de Photoshop, à savoir la mise en correspondance des données Exif et de leur expression en XMP (ou, pour le dire autrement, l'enregistrement des informations stockées en Exif dans le format XMP<sup>75</sup>).

Selon Patrick Peccatte<sup>76</sup>, la mise en application de ce standard amènerait, pour la première fois, à une expression en RDF à la fois de métadonnées générées automatiquement (Exif) et de métadonnées rédigées par des humains (IPTC/IIM).

---

<sup>72</sup> « IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consulté le 11 juin 2013 à 14h39)].

<sup>73</sup> IPTC. *IPTC Standard Photo Metadata. Document Revision 1.* (Juillet 2010) International Press Telecommunications Council, 2010 [En ligne : [http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007\\_1.pdf](http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007_1.pdf) (lien vérifié le 12 juin 2013 à 12h46)].

<sup>74</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 15h10)].

<sup>75</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>76</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 15h10)].

## II. COMPARAISONS, CHAMPS D'APPLICATION ET CONJONCTION DES MÉTADONNÉES IPTC, EXIF ET XMP

L'existence de plusieurs standards de métadonnées oblige à s'intéresser aux différentes sortes d'informations qui peuvent être enregistrées dans chaque format, et aux utilisations possibles ou recommandables de ces formats.

Il est possible de trouver à la fois des métadonnées IPTC, Exif et XMP dans une même image : par exemple, une image composée à partir d'une autre photographie peut contenir un numéro IPTC identifiant cette photographie initiale, des données Exif mentionnant la date de prise de vue, et des données XMP sur les dates et le nombre de modifications<sup>77</sup>. Ces différents types de données sont alors stockés dans des endroits différents du fichier image.

### 2.1. Exif et IPTC

Certaines sources opposent les métadonnées IPTC et les métadonnées Exif : Emmanuel Georjon explique que les premières sont renseignées manuellement et les secondes, globalement, de manière automatique<sup>78</sup>, et Barbara Contassot affirme que les métadonnées IPTC donnent des informations sur la sémantique de l'image, alors que les métadonnées Exif concernent plutôt les caractéristiques techniques des images<sup>79</sup>. Pour Patrick Peccatte<sup>80</sup>, il est possible de distinguer métadonnées « automatiques » et métadonnées « non automatiques », c'est-à-dire de faire la différence entre les informations données par l'appareil qui fabrique le fichier image, et les métadonnées qui, au contraire, sont fournies par un opérateur (en général, en post-production, après la prise de vue ou la numérisation de l'image). Pour lui, les informations contenues dans les photographies peuvent, de toute façon, toujours être qualifiées de « sémantiques », puisqu'il s'agit de descriptions.

---

<sup>77</sup> Exemple tiré de : PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 1<sup>er</sup> décembre 2012 à 19h15)].

<sup>78</sup> GEORJON, Emmanuel. « Photos numériques : présentation des métadonnées ». *Emmanuel Georjon. Photo numérique et technologies Web* [site web]. 10 octobre 2008 [En ligne : <http://www.emmanuelgeorjon.com/photos-numeriques-presentation-des-metadonnees-270/> (consulté le 7 décembre 2012 à 18h04)]. L'auteur ajoute dans son billet que certaines données techniques des métadonnées Exif peuvent être modifiées après la prise de vue, comme l'orientation de l'image si une rotation des clichés est opérée.

<sup>79</sup> CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 44.

<sup>80</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

Comme nous l'avons vu précédemment<sup>81</sup>, il existe des recoupements entre les standards IPTC (Core et IIM) et le standard Exif. De plus, certains champs Exif (ceux qui concernent l'auteur et le titre de l'image, par exemple) ne sont pas remplis automatiquement, et certains champs IPTC sont faciles à renseigner de manière automatique (tels les champs de dates, par exemple la date de prise de vue)<sup>82</sup>.

Alors, quelle attitude adopter face aux recouvrements entre Exif et IPTC ? Selon Jean-Paul Figer (dans un article plutôt destiné au grand public)<sup>83</sup>, une bonne pratique consiste à remplir manuellement les champs IPTC (/IIM) qui concernent « le contenu de l'image (créateur, description, copyright, ...) », et à laisser telles quelles les informations données automatiquement en Exif lors de la prise de vue. Pour Emmanuel Georjon, « l'IPTC [Core] offre des fonctionnalités bien supérieures à EXIF pour la description du contenu »<sup>84</sup>.

Actuellement, de plus en plus de logiciels – par exemple, les versions récentes de Photoshop – renseignent de manière automatique les plages de recouvrement entre IPTC et Exif, ce qui amène à retrouver des informations en plusieurs exemplaires dans les fichiers<sup>85</sup>.

## **2.2. IPTC/IIM et XMP**

Une comparaison entre l'IPTC/IIM et XMP permet de distinguer deux notions : « extensible » et « personnalisable »<sup>86</sup>.

Patrick Peccatte<sup>87</sup> explique que « personnalisable » signifie que l'utilisateur est libre de s'affranchir des règles dictées par le standard : par exemple, les champs définis en IPTC/IIM sont numérotés, mais il est possible de créer d'autres numéros

---

<sup>81</sup> Voir **B. Les métadonnées Exif**, dans la première partie de ce chapitre, plus précisément p. 24.

<sup>82</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>83</sup> FIGER, Jean-Paul. « Comment classer ses photos ? EXIF, IPTC, XMP et autres métadonnées : de la théorie à la pratique ». *Le PC simple* [site web]. Mis à jour le 3 décembre 2007 [En ligne : <http://www.figer.com/publications/collection.htm> (consulté le 12 mai 2013 à 16h17)]. Cet article se propose d'apporter une aide à l'archivage de photographies numériques, et est destiné à des personnes qui n'appartiennent pas forcément aux mondes professionnels de l'informatique ou de l'image.

<sup>84</sup> GEORJON, Emmanuel. « Photos numériques : présentation des métadonnées ». *Emmanuel Georjon. Photo numérique et technologies Web* [site web]. 10 octobre 2008 [En ligne : <http://www.emmanuelgeorjon.com/photos-numeriques-presentation-des-metadonnees-270/> (consulté le 7 décembre 2012 à 18h04)].

<sup>85</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>86</sup> Nous avons trouvé ces termes associés respectivement aux standards XMP et IPTC/IIM dans : CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., pp. 38-39 et 47. (L'auteur cite pour source PECCATTE, Patrick, *Métadonnées : une initiation : Dublin Core, IPTC, EXIF, RDF, XMP*, Soft Experience, mis à jour le 10 mars 2004 [En ligne : <http://peccatte.kerefil.com/Software/Metadata.htm> (consulté le 3 mars 2005)].)

<sup>87</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

pour des champs supplémentaires (non standardisés)<sup>88</sup>. En revanche, toute extension dans le format XMP doit respecter un certain formalisme : pour chaque extension fournie, il faut définir et documenter un schéma RDF. De cette manière, il sera possible d'extraire l'extension et de la stocker quelque part (ce « quelque part » ne posant pas de problème pour autant qu'il ait été documenté). C'est pourquoi le format XMP peut plutôt être qualifié d'« extensible ».

### III. EXPLOITATION DES MÉTADONNÉES

#### 3.1. Les outils de gestion des métadonnées internes aux images

Il existe plusieurs outils pour lire, renseigner et/ou extraire les métadonnées internes d'une image. Nous n'allons pas ici dresser de liste, mais simplement en mentionner quelques-uns, en voyant quelles caractéristiques intéressantes nous pouvons retenir.

Un des outils importants pour la manipulation des métadonnées est ExifTool, utilisé par de nombreux acteurs – comme *Flickr*, par exemple, ou *MédiHAL*<sup>89</sup> (dont nous reparlerons plus loin dans ce mémoire<sup>90</sup>). Ce logiciel regroupe une librairie Perl et une application qui permet de lire et écrire des métadonnées dans plusieurs formats (dont DNG, GIF, JPEG, PDF, PNG, SVG, TIFF, des formats Raw, etc.)<sup>91</sup>. Il supporte les métadonnées Exif, IPTC et XMP, mais aussi d'autres formats tels que les profils ICC, les informations GPS, etc. Il fonctionne avec Unix, Mac OS X et Windows, et est téléchargeable gratuitement<sup>92</sup>.

---

<sup>88</sup> *Ibid.* et CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f., p. 39.

<sup>89</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>90</sup> Voir le deuxième chapitre, consacré aux exemples d'utilisation des métadonnées internes aux images.

<sup>91</sup> *ExifTool by Phil Harvey* [site web] [Accueil] [En ligne : <http://www.sno.phy.queensu.ca/~phil/exiftool/>] (consulté le 6 juin 2013 à 16h14 et 16h42, et le 11 juin 2013 à 17h16).

<sup>92</sup> <http://www.sno.phy.queensu.ca/~phil/exiftool/> (lien vérifié le 11 juin 2013).

Le fait que cet outil soit écrit en Perl a eu comme conséquence, au moins pendant un temps, que seuls les programmeurs en Perl étaient intéressés par cet outil (selon Phil Harvey, le créateur d'ExifTool)<sup>93</sup>.

Patrick Peccatte affirme que l'écriture des métadonnées avec ExifTool (au moins aux formats IPTC, Exif et XMP) est plus laborieuse que la lecture de ces informations avec cet outil ; de plus, le fait qu'ExifTool soit écrit en Perl le rend peu rapide<sup>94</sup>. Néanmoins, il est très utilisé – par exemple, beaucoup de sites comme *Flickr* (service de partage d'images en ligne) utilisent ExifTool pour afficher des métadonnées internes sur leur site<sup>95</sup>.

Nous avons déjà vu que des métadonnées de différents formats (IPTC/IIM, Exif ou XMP) peuvent être incluses dans une même image, et que ces différents types de données sont, dans ce cas, stockés dans des endroits différents du fichier. Or, cette dispersion peut entraîner des discordances entre ces métadonnées lorsque des outils inappropriés sont utilisés<sup>96</sup>. Le cas le plus fréquent se rencontre sur des chaînes de processus<sup>97</sup> : dans le cas d'une gestion de masses considérables d'images, les processus sont industrialisés et réalisés par des automates. Si une photographie possédant des métadonnées IPTC/IIM et XMP est travaillée tout au long d'une chaîne avec des outils qui gèrent ces deux formats, il n'y a aucun problème. Si, par contre, à un moment de la chaîne, un outil qui ne gère pas XMP est utilisé, et qu'ensuite un autre outil qui gère XMP et l'IPTC/IIM est à son tour employé, cela introduit une discordance entre les données XMP et IPTC/IIM : par exemple, les champs qui renseignent sur le titre de l'image n'auront pas le même contenu en IPTC/IIM et en XMP. Donc, pour savoir lequel de ces titres est le bon, il faudra remonter toute la chaîne et voir où la discordance a été introduite, puis la corriger – par exemple en modifiant l'outil. Cependant, la plupart des outils modernes qui gèrent ces métadonnées maintiennent une cohérence optimale entre les données XMP et IPTC/IIM<sup>98</sup>.

---

<sup>93</sup> « An interview with Phil Harvey » [interview de Phil Harvey en 2005] [En ligne : <http://owl.phy.queensu.ca/~phil/exiftool/PhilHarveyInterview.pdf> (consulté le 6 avril 2013, lien vérifié le 11 juin 2013 à 16h06)].

<sup>94</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>95</sup> *Ibid.*

<sup>96</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 17h59)].

<sup>97</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>98</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 17h59)].

Certains outils introduisent d'autres erreurs, comme par exemple plusieurs segments d'un même format de métadonnées<sup>99</sup>.

La gestion et la préservation des métadonnées par les outils de gestion des images sont variables. Voici quelques exemples de problèmes qu'il est possible de rencontrer :

- GD, librairie destinée à la création d'images (dont : PNG, JPEG, GIF, etc.)<sup>100</sup>, ne conserve aucune métadonnée (du moins, jusqu'à juin 2012)<sup>101</sup>.
- ImageMagick, qui permet notamment de créer ou de convertir des images *bitmap*<sup>102</sup> (c'est-à-dire composée par des pixels), ne comprend que les métadonnées IPTC et Exif, donc les images redimensionnées avec ce logiciel perdent toutes les métadonnées XMP qu'elles peuvent contenir<sup>103</sup>.
- Même un logiciel comme Photoshop a pu présenter des problèmes (par exemple, dans certains cas, faire disparaître les métadonnées des images)<sup>104</sup>.

### **3.2. Archivage des métadonnées et indexation**

Les métadonnées internes aux images présentent quelques problèmes quant à la conservation et l'intégrité de l'information qu'elles transmettent. D'une part, il n'existe aucun moyen d'empêcher la modification des métadonnées à l'intérieur des fichiers, y compris pour les informations renseignées automatiquement par le système qui a généré l'image – il est ainsi facile de faire croire, par exemple, qu'une photographie a été prise avec un autre type d'appareil que celui qui a véritablement servi à la créer<sup>105</sup>. De plus, des risques existent, nous l'avons vu, de perdre ces métadonnées en cas d'utilisation d'un outil non approprié.

---

<sup>99</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>100</sup> « About. What is the GD library ? » *GD Library* [site web] [En ligne : <http://libgd.bitbucket.org/pages/about.html> (consulté le 6 juin 2013 à 17h05)]. D'après cette source, GD est surtout utilisé dans le cadre du développement de sites web (bien qu'il puisse être utilisé en dehors du Web).

<sup>101</sup> « Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consulté le 9 mai 2013 à 14h58)], tests réalisés sur le site *Pictify* par David RIECKES, le 19 juin 2012.

<sup>102</sup> Site web officiel d'ImageMagick. [Accueil] [En ligne : <http://www.imagemagick.org/script/index.php> (consulté le 6 juin 2013 à 17h40)]. Cet outil peut lire et écrire des images aux formats JPEG, JPEG 2000, PDF, PNG, Postscript, SVG, TIFF, etc.

<sup>103</sup> Cette information provient des réponses à nos questions envoyées par mail par David RIECKES le 10 mai 2013.

<sup>104</sup> BRÜSCH, Béat (graphiste et illustrateur), dans un commentaire publié le 25 janvier 2011 sur la page « Les photos d'agence de presse sur les sites d'actualité, entre illustration et confusion », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/509> (consulté le 6 juin 2013 à 18h42)]. Cette personne témoigne de la suppression de métadonnées incluses dans une image enregistrée avec la commande « Enregistrer pour le web... » avec Photoshop.

<sup>105</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

Les formats de fichiers peuvent également amener des difficultés. En dehors des principaux formats qui supportent les métadonnées internes (JPEG, TIFF, PSD et DNG<sup>106</sup>), il est possible de distinguer plusieurs cas de figure. Certains formats ne supportent pas – ou mal – ces informations, comme GIF et PNG<sup>107</sup>. Le format Flash, lui, empêche de lire ces métadonnées<sup>108</sup>. Phil Harvey pointe d'autres problèmes liés à certains formats de fichiers tels que TIFF ou JPEG : nous renvoyons vers la page correspondante pour plus de précisions :

« Commentary on Meta Information Formats (or "Why this or that Format Sucks") ». *ExifTool by Phil Harvey* [site web]. 15 novembre 2005, mise à jour le 29 juillet 2010 [En ligne : <http://owl.phy.queensu.ca/~phil/exiftool/commentary.html> (consultée le 14 avril 2013 à 17h07, lien vérifié le 11 juin 2013)].

Les métadonnées internes d'une image, même si elles restent inscrites dans leur fichier, ne suffisent pas pour indexer un document. L'exploitation des informations contenues dans les images numériques à des fins de recherche nécessite d'extraire ces métadonnées des fichiers, et de les enregistrer – en général – dans une base de données<sup>109</sup>. Cette base de données peut être écrite avec différents langages, par exemple en SQL – ce qui est le cas la plupart du temps – ou en XML<sup>110</sup>.

---

<sup>106</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>107</sup> *Ibid.*, et « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 8 mai 2013 à 14h51)].

<sup>108</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>109</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013, et commentaire de PECCATTE, Patrick, publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

<sup>110</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

## DEUXIÈME PARTIE : EXEMPLES D'UTILISATION DES MÉTADONNÉES INTERNES AUX IMAGES

---

Cette partie a pour objectif de présenter des exemples d'utilisation des métadonnées internes aux images par différents types d'acteurs, notamment pour l'indexation des documents iconographiques. Elle cite également des catégories de personnes, d'organisations ou de services en ligne qui pourraient se servir de ces informations, mais ne le font pas, ou peu.

Il ne s'agit pas, là encore, de proposer un panorama complet des usages qui sont faits des données incluses dans les images. L'objectif de ce chapitre est de montrer comment certains acteurs exploitent ces informations et à quelles fins, ou bien au contraire paraissent les délaissier. Notre propos se concentrera beaucoup sur des utilisations dans le cadre du Web, car nos recherches ont majoritairement abouti à des résultats sur ce type d'usages.

Ce développement s'articule en trois parties, qui concernent d'abord les réseaux sociaux et les sites de partage de photographies non spécialisés ; puis les acteurs de la recherche, de la culture et du patrimoine ; enfin, les photographes professionnels et la presse.

### I. RÉSEAUX SOCIAUX ET SERVICES DE PARTAGE DE PHOTOGRAPHIES EN LIGNE

Nous nous intéressons ici à la catégorie des « médias sociaux » (« *social media services* » ou « *social media sites* »)<sup>111</sup> en ligne, un ensemble qui regroupe notamment :

- les réseaux sociaux comme *Facebook*, *Twitter*, etc.
- les services de partage d'images en ligne, tels que *Flickr*, *Ipernity*, etc.

Ces sites n'ont pas prioritairement un objectif patrimonial, scientifique, ou d'information (contrairement, par exemple, à l'archive ouverte *MédiHAL*, qui ne met en ligne que des documents à caractère scientifique, et à laquelle nous nous intéressons plus loin, dans la seconde sous-partie de ce chapitre).

---

<sup>111</sup> « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 7 mai 2013 à 22h14)].

David Riecks<sup>112</sup> constate que ces médias sociaux sont nombreux à ne pas conserver les métadonnées contenues dans les images quand celles-ci sont postées sur leurs sites. Parmi ces services, et parmi ceux qui entravent l'accès à ces informations internes, il distingue plusieurs cas :

- Les métadonnées peuvent être supprimées lors de la mise en ligne de la photographie.
- Elles peuvent être conservées dans la photographie originale, mais enlevées dans toutes les images dérivées de ce fichier (telles que les vignettes, ou d'autres tailles de fichier obtenues à partir de l'original).
- Elles peuvent également être stockées au format Flash, qui rend difficile le téléchargement des fichiers et la visualisation des métadonnées (et qui empêche la visualisation des images sur du matériel mobile tel que iPods, iPads, Blackberries, etc.).
- Dans certains cas, les services peuvent ne conserver que les métadonnées sous la forme spécifiée par le standard IPTC/IIM, ce qui peut supprimer certaines données, comme les informations de contact détaillées<sup>113</sup>.

L'effacement des métadonnées n'est en général pas dû à l'utilisation d'un format qui ne supporterait pas ces informations, puisque la plupart des médias sociaux acceptent la mise en ligne d'images au format JPEG<sup>114</sup>.

Certains services affirment – ou ont affirmé – que cette suppression de métadonnées est effectuée pour diminuer le temps de téléchargement nécessaire à la visualisation des images<sup>115</sup> (par exemple, des services comme Getty Images<sup>116</sup>). La suppression des profils ICC permet également de réduire un peu le volume des fichiers<sup>117</sup>. Plusieurs médias sociaux changent automatiquement la taille des images mises en ligne, pour faire en sorte que les fichiers soient moins volumineux.

---

<sup>112</sup> The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 7 mai 2013 à 22h14)].

<sup>113</sup> *Ibid.* (consulté le 8 mai 2013 à 14h51), et précisions apportées par mail par David RIECKES le 10 mai 2013.

<sup>114</sup> « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 7 mai 2013 à 22h14)].

<sup>115</sup> *Ibid.* (consulté le 7 mai 2013 à 22h14).

<sup>116</sup> Cette information provient des réponses à nos questions envoyées par mail par David RIECKES le 10 mai 2013.

<sup>117</sup> « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 8 mai 2013 à 14h51)].

L'étude « Social Media Photo Metadata use Survey : CV Web Photo Metadata »<sup>118</sup>, menée entre autres par David Riecks, s'intéresse à la préservation des métadonnées dans les images mises en ligne sur les sites des médias sociaux – principalement au format JPEG<sup>119</sup>. Les résultats de cette enquête (toujours en cours) permettent de montrer que le(s) type(s) de métadonnées internes préservées en JPEG, et la proportion des métadonnées conservées, sont variables d'un média social à un autre. Il semble que ce degré de conservation dépende beaucoup des logiciels utilisés et des traitements réalisés pour changer la taille des images ou créer des vignettes. Dans certains cas<sup>120</sup>, seuls certains types de métadonnées sont conservés : par exemple, le logiciel ImageMagick ne comprend que les métadonnées Exif et IPTC/IIM, donc les informations XMP présentes dans les images redimensionnées avec ce logiciel disparaissent. Dans d'autres exemples encore, pour un même format de métadonnées, certains champs sont conservés, et d'autres non<sup>121</sup>. Cependant, dans la plupart des cas, les métadonnées Exif, IPTC/IIM et XMP sont soit toutes préservées, soit toutes supprimées<sup>122</sup>.

## **1.1. Les services de partage de photographies en ligne**

On peut, globalement, distinguer trois catégories parmi ces services<sup>123</sup> :

- ceux qui suppriment la totalité (ou la quasi-totalité) des métadonnées internes des images mises en ligne sur leurs sites
- ceux qui préservent les métadonnées, sans apparemment les utiliser
- ceux qui les préservent et les utilisent pour documenter les photographies sur leurs sites.

---

<sup>118</sup> Les résultats de cette étude sont consultables en ligne à l'adresse suivante : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (lien vérifié le 11 juin 2013 à 22h52). Pour une présentation plus détaillée de cette étude (qui utilise l'outil Jeffrey's Exif Viewer, cité précédemment), voir « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (lien vérifié le 11 juin 2013 à 22h53)].

<sup>119</sup> « The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 8 mai 2013 à 14h51)].

<sup>120</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>121</sup> Voir par exemple le test réalisé le 12 octobre 2012 par David RIECKES pour *Google Plus*, sur une vignette (« Thumbnail (less than 200 pixels) »). Seuls les champs Exif « Artist » et « Copyright » sont conservés dans la vignette générée. (« Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consulté le 9 mai 2013 à 14h58)].)

<sup>122</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>123</sup> Cette catégorisation a été effectuée d'après les résultats de tests datant de 2013 ou 2012, présentés sur la page : « Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consultée le 9 mai 2013 à 14h58)].

Dans la dernière catégorie, on peut citer notamment *Flickr*, qui préserve les métadonnées internes dans le format original des images, et récupère ces informations pour les inscrire dans sa base<sup>124</sup>. Le champ « Caption-Abstract », en IPTC/IIM, vient remplir la description *Flickr*<sup>125</sup>. Vers 2008-2009, et peut-être jusqu'en 2010, *Flickr* propageait les métadonnées internes dans toutes les définitions des images, à partir de la définition originale (haute définition). Cela n'était pas le cas en 2007. Depuis 2010, seule la version originale conserve les métadonnées<sup>126</sup> – si l'utilisateur qui met en ligne l'image possède un compte Pro<sup>127</sup>. David Riecks avance l'hypothèse suivante : ce serait à cause des applications utilisées par *Flickr* que seule l'image originale conserve des métadonnées<sup>128</sup>. Pour Patrick Peccatte, cela peut s'expliquer par des problèmes de performance : *Flickr* fabrique des déclinaisons d'une photographie en plusieurs tailles, et, avant 2010, extrayait les métadonnées de la version originale pour les injecter dans les autres versions : cela avait un coût, et prenait du temps. De plus, Patrick Peccatte pense que ce genre de choses intéresse très peu de monde<sup>129</sup>.

D'autres services de partage de photographies utilisent (ou utilisaient) les métadonnées pour documenter les images sur leurs sites. Par exemple, *Zizics*, un site de partage de photographies généraliste<sup>130</sup>, récupérait – au moins jusqu'en avril 2013 – des informations contenues dans les fichiers, et les affichait sur la page présentant l'image correspondante<sup>131</sup>.

<sup>124</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013, et informations envoyées par David RIECKES le 10 mai 2013.

<sup>125</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013. Cela semble confirmé par des tests que nous avons effectués le 2 juin 2013 sur deux images, dont voici les références :

- Image test fournie sur le site *Controlled Vocabulary*, mise en ligne dans la galerie de Ken Rahaim [En ligne : <http://www.flickr.com/photos/15731269@N00/4143198617/> (lien vérifié le 12 juin 2013 à 0h01)].

- Photographie « p012539.jpg » dans la galerie « PhotosNormandie » [En ligne : <http://www.flickr.com/photos/58897785@N00/393070101/in/photolist-AJA3H-F4KSY-F4KTb-F4KTw-FYtsf-J281h-MNAB6-2ztUTg-33z6ut-4bGZkH-4geKWf-4xNaC4-4xNaGt-4xNaLx-4xSnQo-4xSo5Q-4xSoon-4EK4DN-58zyW5-58NaHA-5rv1F8-5rv1He-5xS12p-5G9kY1-5KHLLu-5TUM5K-5TUTR4-6bByCH-6gpwsK-6gHbgT-6jodf9-6MTqtQ-6ZUeto-79mHr5-7acJ23-7fNecF-7mtC8A-7mDSCX-7uqa2d-dAfcH6-dys6tR-bwrJCF-duERXJ-cupqj9-bHkVq2-bxqe2H-bHUSTZ-dZSjno-aT92Ca-abYqBk-dtx8b6> (lien vérifié le 12 juin 2013 à 0h03)].

<sup>126</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013. Un test effectué le 2 juin 2013 sur la version originale de la photographie « p012539.jpg » dans la galerie « PhotosNormandie » (voir URL dans la note précédente) avec l'outil Jeffrey's Exif Viewer [En ligne : <http://regex.info/exif.cgi>] montre que les métadonnées internes ont été, apparemment, préservées après la mise en ligne de l'image. En revanche, la version « grand format » de cette photographie (de taille inférieure à celle de l'originale), testée avec le même outil, ne contient aucune métadonnée IPTC, Exif ou XMP. (D'autres formats de cette image étaient disponibles, mais nous ne les avons pas testés.)

<sup>127</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>128</sup> *Ibid.*

<sup>129</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>130</sup> *Zizics* [site web]. [Accueil] [En ligne : <http://www.zizics.com/> (consulté le 8 juin 2013 à 12h13)].

<sup>131</sup> Test effectué sur *Zizics* par David RIECKES le 23 avril 2013 (image de type « Large Preview (400+ pixels) », dans l'étude).

(« Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consulté le 9 mai 2013 à 14h58)].)

Les informations qui ont été récupérées par *Zizics* sont toujours présentées à la même place sur la page correspondant à l'image testée [En ligne : <http://www.zizics.com/image/747> (consultée le 8 juin 2013 à 12h28)].

Voici ces informations :

- un bref résumé du contenu de l'image (« headline »)<sup>132</sup>, présenté au-dessus de celle-ci
- une description de l'image
- des descripteurs de l'image, présentés à côté du mot « Tags » sur le site.

La présentation des informations ci-dessus n'est peut-être pas forcément exacte, car nous n'avons pas déterminé à quel format appartenait les champs concernés : en effet, sur la page où l'image testée par David Riecks a été mise en ligne (sur *Zizics*)<sup>133</sup>, les valeurs affichées, extraites des informations internes de l'image, peuvent appartenir à plusieurs formats de métadonnées<sup>134</sup>.

La plupart des services comme *Flickr* offrent la possibilité de voir les métadonnées Exif sur leurs sites<sup>135</sup>. Pour cela, ils utilisent souvent l'outil ExifTool (c'est le cas notamment de *Flickr*)<sup>136</sup>.

Sur *Flickr*, il est possible, dans certains cas, d'afficher les métadonnées internes d'une image grâce à un lien « Afficher les informations Exif », qui permet d'accéder à une page où sont présentées des informations incluses dans le fichier original. Nous présentons, à la page suivante, une capture d'écran montrant un exemple<sup>137</sup>.

---

<sup>132</sup> Nous reprenons ici la définition du contenu du champ « Headline » en IPTC Core présentée dans RIECKES, David. *"IPTC Core" Schema for XMP. Version 1.0. Supplemental documentation. Custom Panels User Guide. Document Revision 13*. 2005 [En ligne : [http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore\\_1.0-doc-CpanelsUserGuide\\_13.pdf](http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore_1.0-doc-CpanelsUserGuide_13.pdf)], p. 9. Cependant, l'image testée par David Riecks ne contient des champs IPTC qu'au format IPTC/IIM (« The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 9 mai 2013 à 15h05)].

<sup>133</sup> Voir note 131.

<sup>134</sup> Les métadonnées internes de l'image concernée, affichées par l'outil Jeffrey's Exif Viewer, peuvent être consultées directement à l'adresse [http://regex.info/exif.cgi?b=3&referer=http%3A//www.controlledvocabulary.com/socialmedia/cv-testbed\\_social-media.jpg&url=http%3A//www.controlledvocabulary.com/socialmedia/cv-testbed\\_social-media.jpg](http://regex.info/exif.cgi?b=3&referer=http%3A//www.controlledvocabulary.com/socialmedia/cv-testbed_social-media.jpg&url=http%3A//www.controlledvocabulary.com/socialmedia/cv-testbed_social-media.jpg) (lien vérifié le 8 juin 2013 à 15h53).

<sup>135</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>136</sup> *Ibid.*

<sup>137</sup> Présentation des métadonnées internes de la photographie « p012539.jpg » dans la galerie « PhotosNormandie » [En ligne : <http://www.flickr.com/photos/58897785@N00/393070101/in/photolist-AJA3H-F4KSY-F4KTb-F4KTw-FYtsf-J281h-MNaB6-2ztUTg-33z6ut-4bGZkH-4geKWf-4xNaC4-4xNaGt-4xNaLx-4xSnQo-4xSo5Q-4xSooN-4EK4DN-58zyW5-58NaHA-5rv1F8-5rv1He-5xS12p-5G9kY1-5KHLLu-5TUM5K-5TUTR4-6bByCH-6gpwsK-6gHbgT-6jodf9-6MTqtQ-6ZUeto-79mHr5-7acJ23-7fNecF-7mtC8A-7mDSCX-7uqa2d-dAfcH6-dys6tR-bwrJcF-duERXJ-cupqj9-bHkVq2-bxqe2H-bHUSTZ-dZSjno-aT92Ca-abYqBk-dtx8b6> (lien vérifié le 12 juin 2013 à 0h03)].

La page présentée a été affichée le 2 juin 2013 à 23h41.

La photographie étudiée est utilisable sous licence Creative Commons « Paternité-Partage selon les conditions initiales » (mention de propriété intellectuelle : Conseil Régional de Basse-Normandie / National Archives USA).

## Photo / Exif



## Que sont les données Exif ?

Les données Exif sont des informations sur les paramètres utilisés par un appareil photo pour prendre une photo ou une vidéo. Ces informations sont incorporées aux fichiers enregistrés par l'appareil photo. Nous les lisons et les affichons ici.

Remarque : Désolé, mais la plupart des données Exif sont disponibles uniquement en anglais actuellement.

## Dates

Prises en 1944  
Ajoutée le 17 février 2007 à 9:05 (PDT)

## Données Exif

Object Name	p012539.jpg
Keywords	"Seconde Guerre Mondiale" "Gi bateau WWII WW2" "Battle of Normandy"
City	OMAHA
Province- State	La Manche / The Channel
Copyright Notice	Conseil Régional de Basse-Normandie / National Archives USA
Caption- Abstract	Une vedette des US Coast Guards remorque un DUKW, ils sont nombreux à être tombés en panne de carburant après avoir tourné en rond pendant plusieurs heures, obligés de faire tourner le moteur pour faire fonctionner les pompes de cale.
X-Resolution	300 dpi
Y-Resolution	300 dpi
Orientation	Horizontal (normal)
Software	Adobe Photoshop CS Windows
Date and Time (Modified)	2004:12:29 18:02:10

*Illustration 3: Affichage de métadonnées internes d'une photographie sur Flickr. Capture d'écran (2 juin 2013).*

(Pour les références et la licence d'utilisation correspondant à la photographie et à la page montrées en Illustration 3 ci-dessus, voir note 137, page précédente.)

Il faut remarquer que ces métadonnées sont présentées comme des « données Exif », alors qu'en réalité les champs appartiennent à plusieurs formats de métadonnées (ainsi, dans l'exemple présenté, les champs IPTC/IIM, XMP et Exif sont présentés sans aucune distinction entre les formats auxquels ils se rapportent) : en témoigne la présence, sur la capture d'écran ci-dessus, du champ IPTC/IIM « Caption-Abstract ».

*Ipernity* propose aussi de voir des métadonnées des images : présentées comme des « Informations de prise de vue », elles se trouvent sur une page à part, accessible à partir de la page qui présente l'image concernée<sup>138</sup>. Les renseignements présentés dans cette rubrique « Informations de prise de vue » peuvent inclure des indications sur ce que montre l'image<sup>139</sup>. Nous n'avons aucune certitude sur la manière dont ces métadonnées sont récupérées par *Ipernity*.

<sup>138</sup> Tests réalisés le 8 juin 2013 sur trois photographies :

« Projet HDR exif auto » [En ligne : <http://www.ipernity.com/doc/grauminay/4017717>].

« Cadrans solaire... il est 10 h (solaire ! Voir exif) » [En ligne : [http://www.ipernity.com/doc/declic\\_67/13695675](http://www.ipernity.com/doc/declic_67/13695675)].

« Closeup » [En ligne : <http://www.ipernity.com/doc/afulki/3048425>].

<sup>139</sup> Voir les pages

[http://www.ipernity.com/doc/declic\\_67/13695675/infos](http://www.ipernity.com/doc/declic_67/13695675/infos) et <http://www.ipernity.com/doc/afulki/3048425/infos> (correspondant aux dernières images citées dans la note précédente), où des descripteurs sont marqués dans le champ « Subject ». (Pages consultées le 8 juin 2013 à 12h05).

Plusieurs sites de partage d'images (comme *Flickr*) utilisent les métadonnées de géolocalisation pour indexer les photographies – le terme « géolocaliser » devant s'entendre comme le fait d'associer à un document numérique des informations géographiques<sup>140</sup>, qui permettent de le rattacher à un lieu. Ces informations peuvent être de différentes natures : elles comprennent les coordonnées géographiques (correspondant à une latitude et à une longitude), mais aussi des noms de lieux ou des indications d'altitude<sup>141</sup>.

Dès le début des années 2000, des travaux ont été menés (par des laboratoires de recherche, majoritairement américains) pour créer des plates-formes de partage d'images qui utiliseraient la position géographique comme un moyen d'accéder à une photographie – et aussi comme un critère de classement<sup>142</sup>.

*Flickr*, par exemple, utilise (ou a utilisé, au moins pendant un temps<sup>143</sup>) les métadonnées Exif pour référencer automatiquement les images qui les incluent sur la carte proposée par le site. Les éléments « géotaggés »<sup>144</sup> sont consultables à partir de cette carte, affichée à l'adresse <http://www.flickr.com/map><sup>145</sup>. Il est possible d'effectuer une recherche par descripteurs en cliquant sur le bouton « Rechercher sur la carte » et en remplissant deux champs proposés par *Flickr* : l'un correspond à des descripteurs tels que « architecture », « urbanisme », ou « forêt », l'autre à la position géographique (voisinage, ville, état, code postal, etc.)<sup>146</sup>. Il est possible de ne remplir que l'un des deux champs<sup>147</sup>. Une fois la requête validée, le site zoome éventuellement sur la partie de la carte correspondante ; il présente alors des photographies qui lui semblent pertinentes – dans une bande située dans la partie inférieure de l'écran –, ainsi que des points roses, sur la carte, qui indiquent des lieux associés à des images répondant à la requête. Il est possible de cliquer sur les points roses pour afficher, en petit format, les photographies qui y sont rattachées. Inversement, si l'on clique sur une image présentée dans le bandeau en bas de l'écran, l'apparence visuelle du point rose correspondant est modifiée, ce qui permet de repérer le lieu associé à la photographie.

<sup>140</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 81.

<sup>141</sup> *Ibid.*, p. 81.

<sup>142</sup> *Ibid.*, p. 82.

<sup>143</sup> *Ibid.*, p. 86. Nous n'avons pas trouvé de source disant que *Flickr* continue à fonctionner de la même manière (ou au contraire que cette pratique s'est arrêtée).

<sup>144</sup> La carte, visualisée le 4 juin 2013, est accompagnée d'un bandeau présentant des images ainsi que le nombre d'« éléments géotaggés ». Dans la FAQ de *Flickr*, il est expliqué que le géotagging « consiste à ajouter des informations géographiques sur des éléments, comme une photo » (rubrique « Que signifie "géotaggé" ? » sur la page « Aide / FAQ / La carte ». *Flickr* [site web] [En ligne : <http://www.flickr.com/help/map/>] (consultée le 12 juin 2013 à 2h11)).

<sup>145</sup> Lien vérifié le 4 juin 2013 à 18h42.

<sup>146</sup> Nous reprenons ici les exemples proposés sur le site. Nous avons testé, pour le champ correspondant à la position géographique, un nom de pays et un code postal (le 4 juin 2013).

<sup>147</sup> Tests effectués le 4 juin 2013.

Voici une capture d'écran de cette interface<sup>148</sup> :

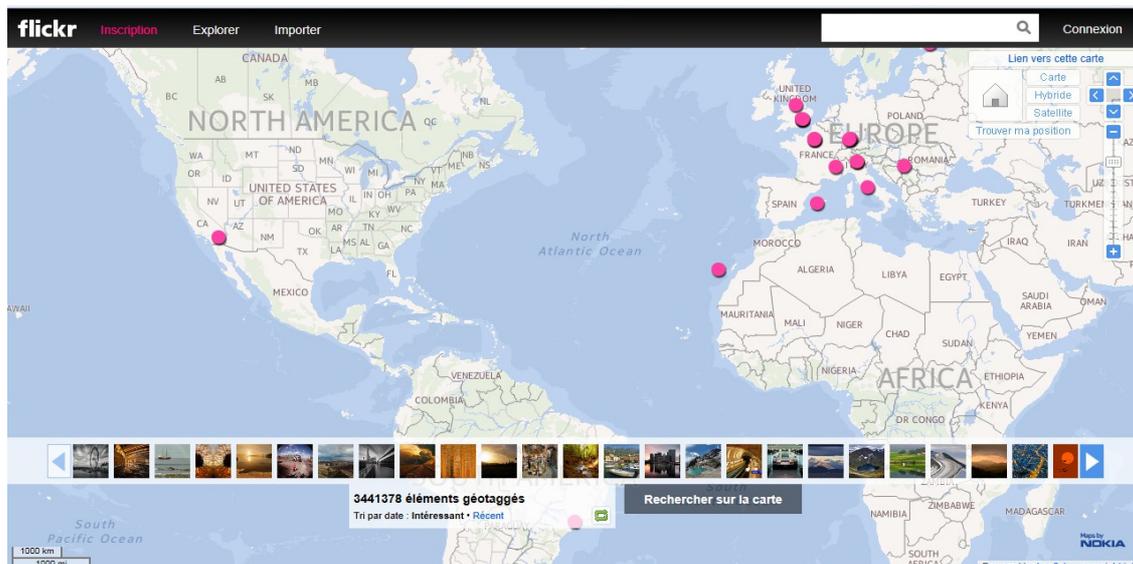


Illustration 4: Interface cartographique de Flickr.  
Capture d'écran (4 juin 2013).

Bien que *Flickr* ait permis d'attribuer des coordonnées géographiques à des photographies qui n'en contenaient pas, les informations de type latitude et longitude n'étaient pas intégrées – au moins jusqu'en 2009 – aux métadonnées Exif de l'image<sup>149</sup>.

Une autre fonctionnalité offerte par *Flickr* nous semble intéressante : la recherche par appareil photo. Sur la page « La recherche par appareil photo »<sup>150</sup>, les marques et les modèles d'appareils photographiques et de téléphones portables les plus utilisés par les membres de la « communauté Flickr » sont présentés, notamment par un tableau dans lequel il est possible de cliquer sur les noms des marques et les noms des modèles (séparément). Le lien correspondant au nom d'un modèle envoie l'internaute sur une page où sont présentées des caractéristiques – notamment techniques – de l'appareil, et une série de photographies « intéressantes » (nous reprenons le terme présenté sur le site) qui ont été prises avec ce modèle<sup>151</sup>.

<sup>148</sup> [En ligne : <http://www.flickr.com/map> (affichée le 4 juin 2013 à 19h09)].

<sup>149</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], pp. 85-86.

<sup>150</sup> « La recherche par appareil photo ». *Flickr* [site web] [En ligne : <http://www.flickr.com/cameras/> (consulté le 7 juin 2013 à 23h00)].

<sup>151</sup> Tests effectués le 7 juin 2013 en cliquant sur les liens correspondant à 4 modèles sur la page « La recherche par appareil photo », et en choisissant ensuite deux photographies pour chacun de ces appareils.

Nous avons cherché à savoir si les données sur l'appareil utilisé provenaient – ou pouvaient provenir – des métadonnées internes : les essais réalisés pour trois modèles<sup>152</sup> montrent que, pour des photographies dont les informations internes sont accessibles sur le site<sup>153</sup>, les noms des appareils utilisés apparaissent parmi ces métadonnées (dans le champ intitulé « Caméra ») et sont cliquables.

Dans les trois cas étudiés, les métadonnées présentées comme « Données Exif » (sur la page « Photo / Exif ») contiennent des dates dans le champ « Date and Time (Original) » qui sont reprises sous la forme d'un lien au-dessus des « Données Exif », dans un champ intitulé « Date de prise de vue ». Ces liens mènent, apparemment<sup>154</sup>, vers les photographies prises le même jour par un même membre de *Flickr*.

## **1.2. Les réseaux sociaux : Facebook et Twitter**

Les réseaux sociaux comme *Facebook* ou *Twitter* permettent également de publier des images en ligne, sans se limiter à cette fonctionnalité. D'après les tests effectués en 2012 dans le cadre de l'étude « Social Media Photo Metadata use Survey : CV Web Photo Metadata »<sup>155</sup>, ni *Facebook* ni *Twitter* ne conservaient, jusqu'alors, les métadonnées contenues dans les photographies<sup>156</sup>.

Pour *Twitter*, l'image testée n'a conservé aucune information.

Sur *Facebook*, un test a été réalisé avec une image de même taille que l'originale avec l'option « High Quality », lors de la mise en ligne<sup>157</sup> : aucune métadonnée IPTC/IIM ou XMP n'est conservée. Pourtant, le profil ICC de l'image, lui, est préservé, et signalé dans le seul champ Exif restant (« Color Space »). Dans le cas d'une vignette<sup>158</sup>, le profil ICC originel est remplacé par le profil ICC de *Facebook*, toutes les

<sup>152</sup> Tests effectués sur *Flickr* le 8 juin 2013 pour trois modèles d'appareil, avec une photographie par modèle.

<sup>153</sup> Ces informations sont affichées sur une page « Photo / Exif », dont un exemple est présenté par la capture d'écran référencée « *Illustration 3* », p. 40 du présent mémoire.

<sup>154</sup> Seule une des trois images avait une date de prise de vue qui correspondait à plusieurs images.

<sup>155</sup> « Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consulté le 9 mai 2013 à 14h58)].

<sup>156</sup> Test effectué sur *Twitter* par David RIECKES le 15 octobre 2012, sur une image de type « Large preview (400 + pixels) ».

Tests effectués sur *Facebook* par David RIECKES sur une image de la même taille que l'originale et sur une vignette (« Thumbnail (less than 200 pixels) »), respectivement les 29 et 25 mars 2012, et par Christy KARPINSKI le 15 septembre 2012, sur une vignette.

<sup>157</sup> Test effectué sur *Facebook* par David RIECKES le 29 mars 2012.

<sup>158</sup> Test effectué sur *Facebook* par David RIECKES le 25 mars 2012 sur une vignette (« Thumbnail (less than 200 pixels) »).

autres métadonnées étant supprimées. David Riecks affirme qu'il ne peut pas y avoir de certitude que *Facebook* supprime des informations volontairement, mais il suppose, puisque ce média sait inclure son propre profil ICC dans les fichiers, que ces suppressions sont volontaires. La mention du profil ICC est notée dans le champ « Color Space » des métadonnées Exif<sup>159</sup>.

Bien que *Facebook* supprime systématiquement les métadonnées d'une image publiée sur son site, il extrait les informations intéressantes du fichier – puis les supprime de ce fichier – pour les utiliser dans la page qui entoure l'image<sup>160</sup>. Ainsi (au moins en mars 2012<sup>161</sup>), les champs IPTC [Core ?] « Title », « Description » et « Copyright Notice » sont utilisés par *Facebook*.

## II. DES ACTEURS DE LA RECHERCHE, DE LA CULTURE ET DU PATRIMOINE

Nous réunissons dans cette sous-partie des informations sur des acteurs qui exercent une activité à caractère culturel, patrimonial ou scientifique (laboratoires, bibliothèques, services d'archives, etc.).

Patrick Peccatte, qui dispense des formations sur les métadonnées internes aux images, affirme que les acteurs culturels (tels que, par exemple, le Musée du Louvre) sont assez intéressés par des renseignements sur ce sujet<sup>162</sup>.

Nous nous proposons d'exposer ici quelques exemples d'utilisation des métadonnées internes aux images par certains de ces organismes ou services.

---

<sup>159</sup> Informations envoyées par David RIECKES le 10 mai 2013.

<sup>160</sup> *Ibid.*

<sup>161</sup> Test effectué sur *Facebook* par David RIECKES le 29 mars 2012 avec une image de la même taille que l'originale. « Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consulté le 9 mai 2013 à 14h58)].

<sup>162</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

Le Musée du Louvre fait partie des organismes qui ont suivi la formation *Les métadonnées des images numériques [IPTC et XMP]* (« Les métadonnées des images numériques fixes ». *Soft Experience* [site web] [En ligne : <http://peccatte.karefil.com/software/IPTCXMPCCours.html> (consulté le 18 décembre à 15h06)]. Cette formation s'adresse aux débutants en la matière, et a notamment pour objectifs :

- de faire connaître les techniques et standards de métadonnées pour les images numériques
- d'apprendre à manipuler (éditer, collecter, transformer) ces métadonnées.

## **2.1. L'archive ouverte *MédiHAL***

*MédiHAL* est une archive ouverte spécifique de *HAL*<sup>163</sup>, réalisée par le CCSD (Centre pour la communication scientifique directe)<sup>164</sup>, en coopération avec le Centre national pour la numérisation de sources visuelles (CN2SV)<sup>165</sup> du CNRS<sup>166</sup>. Ce site permet à des chercheurs, enseignants-chercheurs et personnels d'accompagnement de l'enseignement supérieur et de la recherche, de déposer des images ayant un but ou un contexte de production scientifique (programme de recherche, mission scientifique, etc.).

Les images accueillies par *MédiHAL* peuvent être des photographies numériques, mais aussi des images scannées, des images de synthèse, etc., au format TIFF ou JPEG. Ces documents sont ensuite mis en accès public sur le Web (soit immédiatement, soit après une barrière mobile).

Les disciplines scientifiques concernées appartiennent aux domaines suivants<sup>167</sup> :

- Informatique
- Mathématiques
- Physique
- Planète et Univers
- Sciences de l'environnement (Biodiversité et écologie, Environnement et Société, Ingénierie de l'environnement et Milieux et Changements globaux<sup>168</sup>)
- Sciences de l'Homme et de la Société (Anthropologie, Architecture, Droit, Géographie, Histoire, Linguistique, Littératures, Philosophie, Psychologie, Sciences politiques, Sociologie, etc.<sup>169</sup>)
- Sciences de l'ingénieur
- Sciences du Vivant

D'autres domaines ont été prévus par *MédiHAL*, mais, pour l'instant, aucune image ne leur est associée : il s'agit notamment de la chimie, de l'économie et de la

---

<sup>163</sup> *HAL* est une archive ouverte destinée au dépôt et à la diffusion d'articles scientifiques de « niveau recherche » et de thèses, dans plusieurs disciplines (*HAL* [site web]. [Accueil] [En ligne : <http://hal.archives-ouvertes.fr/> (consulté le 12 juin 2013 à 10h06)]).

<sup>164</sup> Le CCSD, rattaché au CNRS, a pour principale mission de réaliser des archives ouvertes (*CCSD* [site web]. [Accueil] [En ligne : <http://ccsd.cnrs.fr/> (consulté le 12 juin 2013 à 10h12)]).

<sup>165</sup> Le CN2SV, entre autres, anime deux services de diffusion de documents scientifiques numériques ou numérisés, dont *MédiHAL* (*CN2SV. Centre National pour la Numérisation de Sources Visuelles* [site web]. [Accueil] [En ligne : <http://www.cn2sv.cnrs.fr/> (consulté le 12 juin 2013 à 10h26)]).

<sup>166</sup> *MédiHAL* [site web]. [Accueil] [En ligne : <http://medihal.archives-ouvertes.fr/> ] (consulté le 15 décembre 2012 à 19h29)].

<sup>167</sup> Cette liste est tirée de la page qui présente l'interface permettant de rechercher des images par domaine scientifique. « Consulter>Par domaine ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action\\_todo=browse&b\\_type=browse\\_domain](http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action_todo=browse&b_type=browse_domain) (consulté le 6 mai 2013 à 22h18)].

<sup>168</sup> « Consulter>Par domaine ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action\\_todo=browse&b\\_type=browse\\_domain](http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action_todo=browse&b_type=browse_domain) (consulté le 26 mai 2013, lien vérifié le 12 juin 2013 à 10h52)].

<sup>169</sup> *Ibid.*

finance quantitative, des sciences cognitives (informatique, linguistique, neurosciences et psychologie) et des statistiques.<sup>170</sup>

*MédiHAL* s'intéresse à deux standards de métadonnées internes aux images : Exif et l'IPTC Core. Il récupère notamment un certain nombre de ces métadonnées, de manière automatique, pour compléter des notices relatives chacune à une photographie ou à un lot de photographies<sup>171</sup>.

Si une photographie comporte des métadonnées internes, les informations concernant le titre, l'auteur, la date de prise de vue, les coordonnées GPS, les « mots-clés », et la description du contenu, sont extraites pour compléter la notice correspondante.

Dans le cas du dépôt d'un lot de photographies, si toutes les images ne contiennent pas de métadonnées, il est possible de choisir une photographie de référence, dont les métadonnées vont s'appliquer à toutes les images du lot qui n'en contiennent pas. Les photographies qui ne sont pas choisies comme référence, mais qui possédaient déjà des métadonnées propres, conservent celles-ci à l'issue de ce traitement.

Pour une image isolée comme pour un lot de photographies, il est toujours possible de compléter et modifier les notices après l'extraction automatique des métadonnées. Les champs à compléter sont notamment la licence d'exploitation, le domaine scientifique auquel est rattachée l'image, et éventuellement le laboratoire auquel appartient l'auteur de la photographie.

Les coordonnées GPS inscrites dans les fichiers images font l'objet d'un traitement spécifique : elles sont récupérées automatiquement pour positionner l'image sur une représentation cartographique fournie par Google Maps. *MédiHAL* utilise également le référentiel GeoNames pour associer les coordonnées à des noms de lieux : ce service met à disposition une base de données géographique sous licence Creative Commons<sup>172</sup>.

Voici un exemple de photographie accompagnée de sa notice<sup>173</sup> :

---

<sup>170</sup> « Consulter>Par domaine ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action\\_todo=browse&b\\_type=browse\\_domain](http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action_todo=browse&b_type=browse_domain) (consulté le 26 mai 2013, lien vérifié le 12 juin 2013 à 10h52)].

<sup>171</sup> « Aide en ligne ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?action\\_todo=faq#videos](http://medihal.archives-ouvertes.fr/index.php?action_todo=faq#videos) ] (consulté le 15 décembre 2012)]. Complété par des vérifications sur le site de *MédiHAL*, effectuées le 20 mai 2013.

<sup>172</sup> « About GeoNames ». *GeoNames* [site web] [En ligne : <http://www.geonames.org/about.html> (consulté le 12 juin 2013 à 11h42)].

<sup>173</sup> Exemple extrait de *MédiHAL* le 26 mai 2013 à 16h01 (l'adresse correspondante n'a pas été indiquée, car elle renvoie à l'accueil du site (lien vérifié le 12 juin 2013)).



**MédiHAL**  
Archive ouverte de photographies et d'images scientifiques

Accueil Déposer Consulter Rechercher S'abonner

version française RSS

### République de Djibouti : faille du Ghoubet

[Voir l'ensemble des métadonnées]

Stéphane Pouyllau  
11/11/2009 Paternité - Pas d'utilisation commerciale - Pas de modification

Domaine : Planète et Univers/Sciences de la Terre/Tectonique  
Sciences de l'environnement/Milieux et Changements globaux  
Sciences de l'Homme et Société/Géographie

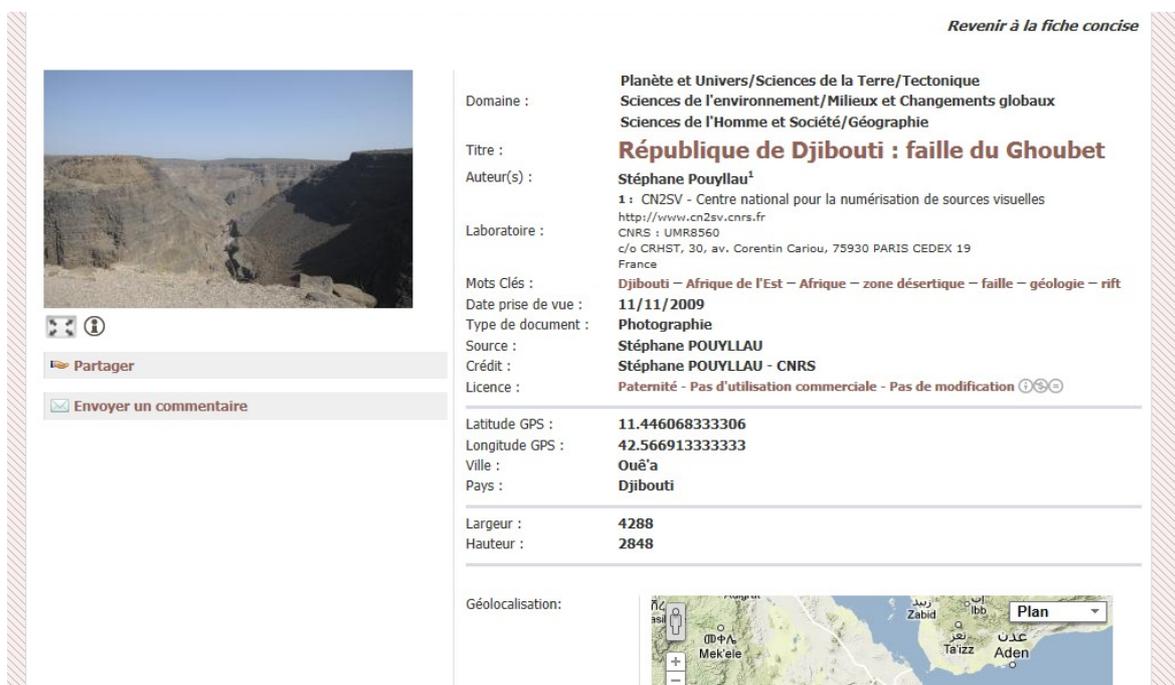
Géolocalisation:



Mots Clés : Djibouti – Afrique de l'Est – Afrique – zone désertique – faille – géologie – rift

Illustration 5: Exemple de notice sur MédiHAL, avec le positionnement géographique de la photographie sur une carte. Capture d'écran (26 mai 2013).

Il est possible d'afficher une notice plus complète en cliquant sur « Voir l'ensemble des métadonnées », à droite. Voici, pour notre exemple, le résultat<sup>174</sup> :



Revenir à la fiche concise

Domaine : Planète et Univers/Sciences de la Terre/Tectonique  
Sciences de l'environnement/Milieux et Changements globaux  
Sciences de l'Homme et Société/Géographie

Titre : **République de Djibouti : faille du Ghoubet**

Auteur(s) : Stéphane Pouyllau<sup>1</sup>

Laboratoire : 1 : CNZSV - Centre national pour la numérisation de sources visuelles  
http://www.cnzsv.cnrs.fr  
CNRS : UMR8560  
c/o CRHST, 30, av. Corentin Cariou, 75930 PARIS CEDEX 19  
France

Mots Clés : Djibouti – Afrique de l'Est – Afrique – zone désertique – faille – géologie – rift

Date prise de vue : 11/11/2009

Type de document : Photographie

Source : Stéphane POUYLLAU

Crédit : Stéphane POUYLLAU - CNRS

Licence : Paternité - Pas d'utilisation commerciale - Pas de modification

Latitude GPS : 11.446068333306  
Longitude GPS : 42.566913333333  
Ville : Oué'a  
Pays : Djibouti

Largeur : 4288  
Hauteur : 2848

Géolocalisation:



Illustration 6: Exemple de notice sur MédiHAL (2) : notice détaillée. Photographie identique à celle de l'Illustration 5. Capture d'écran (26 mai 2013).

<sup>174</sup> Exemple consulté le 26 mai 2013 à 16h05. L'adresse correspondante n'a pas été indiquée, car elle revoie à l'accueil du site (lien vérifié le 12 juin 2013).

L'utilisation des métadonnées par *MédiHAL* ne se limite pas au remplissage automatique de notices, ni au positionnement géographique des images sur une carte. D'autres informations que celles mentionnées précédemment sont utilisées, par exemple :

- pour les métadonnées Exif, la taille du fichier, les dimensions et la résolution de l'image, la date et l'heure de la prise de vue ou de la numérisation, le temps d'exposition, etc.
- pour les informations en IPTC, le code du pays (« Country/Primary Location Code »)<sup>175</sup> est extrait, en plus d'autres renseignements.

Nous n'avons malheureusement pas plus de précisions sur l'usage que *MédiHAL* fait de ces informations qui, *a priori*, ne sont pas présentées dans les notices détaillées mises en ligne<sup>176</sup>.

### **2.3. L'ECPAD (Établissement de Communication et de Production Audiovisuelle de la Défense)**

L'ECPAD assure plusieurs missions, dont nous ne présenterons que les principales. D'une part, il s'occupe de la collecte – par des équipes de reportage – d'images en lien avec l'actualité de la défense, et diffuse ces informations, notamment aux agences de presse et à des chaînes de télévision. D'autre part, il reçoit et conserve les archives audiovisuelles et photographiques de la défense<sup>177</sup>.

Depuis février 2012, mandaté par la DICOd (Délégation à l'information et à la communication de la défense)<sup>178</sup>, l'ECPAD propose, sur son site web, une banque

---

<sup>175</sup> Bien que *MédiHAL* (d'après les informations notées sur son site) s'appuie sur les standards Exif et IPTC Core, l'intitulé du champ donné sur le site (« Country/Primary Location Code ») correspond, apparemment, à un champ IPTC/IIM, transposé dans le champ « Country Code » en IPTC Core.

Le champ « Country Code », en IPTC Core 1.1., concerne (pour les images) le pays dans lequel se trouve le lieu montré sur une photographie. Le code correspondant à ce pays devrait être conforme à la norme ISO 3166 (IPTC. *IPTC Standard Photo Metadata. Document Revision 1*. (Juillet 2010) International Press Telecommunications Council, 2010 [En ligne : [http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007\\_1.pdf](http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007_1.pdf) (lien vérifié le 12 juin 2013 à 12h46)], p. 13).

<sup>176</sup> Tests effectués le 20 mai 2013 sur deux photographies qui, peut-être, dans leur version originale, contenaient des métadonnées internes : « République de Djibouti : paysage de l'Ardoukôba » et « République de Djibouti : faille du Ghoubet », par Stéphane POUYLLAU (consultées le 20 mai 2013 à 19h34). Nous n'indiquons pas les adresses correspondant à ces images, car elles renvoyaient toutes deux à la page d'accueil du site.

Nous pensons que ces images ont peut-être contenu des métadonnées, car leur auteur, M. Stéphane Pouyllau, est le co-concepteur de *MédiHAL* (voir la présentation qu'il fait de lui-même sur la page <http://www.stephanepouyllau.org/> (consultée le 12 juin 2013 à 12h37).

<sup>177</sup> « Présentation ». *ECPAD* [site web] [En ligne : <http://www.ecpad.fr/lecpad/presentation> (consulté le 1<sup>er</sup> juin 2013 à 17h38)].

<sup>178</sup> La DICOd a principalement pour missions de gérer les organismes chargés d'information et de communication au sein du ministère de la Défense, de mener la communication des armées dans les médias, et d'assurer la cohérence des « actions d'information et de relations publiques sur la politique de défense ». (« La Délégation à l'information et à la

d'images baptisée « Armées Photos Actu », qui rassemble des photographies en lien avec l'actualité de la défense<sup>179</sup> – il est également possible d'y trouver des images relatives à des événements militaires assez anciens, par exemple de la Seconde Guerre mondiale. Les photographies présentées dans cette rubrique sont censées toutes posséder des métadonnées au format IPTC Core<sup>180</sup>, renseignées par les photographes à la demande de l'ECPAD<sup>181</sup>.

En 2010, le capitaine Christophe Cazorla a conduit un groupe de travail qui a permis la rédaction du Guide de recommandations pour l'audiovisuel de la Défense<sup>182</sup>. Avec ces recommandations, l'ECPAD a cherché à faciliter l'archivage électronique automatisé de documents audiovisuels. En effet, cet établissement ne pouvant pas saisir toutes les informations relatives aux images et aux vidéos qui lui sont versées, il a choisi de demander aux producteurs de renseigner ces métadonnées, pour que celles-ci soient ensuite automatiquement récupérées et importées dans des champs documentaires (nous n'avons pas plus de précisions à apporter sur la nature de ces champs). Ces procédures concernent plusieurs entités au sein de la défense.

Le Guide de recommandations pour l'audiovisuel numérique des OICD (organismes d'information et de communication de la défense)<sup>183</sup> demande qu'au minimum soient remplis les champs IPTC suivants<sup>184</sup> :

- Le nom et le prénom du photographe
- L'unité de production et l'adresse de l'unité de production
- Le contexte du reportage
- Le contexte du dossier (ou groupe photo)
- La légende
- La ville, la région et le pays

---

communication de la défense (DICOd) ». Site web du Ministère de la Défense. Mis à jour le 13 mars 2013 [En ligne : <http://www.defense.gouv.fr/portail-defense/ministere/organisation/organisation-du-ministere-de-la-defense/organismes-dependant-du-ministre/la-delegation-a-l-information-et-a-la-communication-de-la-defense-dicod> (consulté le 12 juin 2013 à 15h26)].

<sup>179</sup> « L'ECPAD met en ligne les photos d'actualité du ministère de la défense ». ECPAD [site web]. 23 février 2012 [En ligne : <http://www.ecpad.fr/l%E2%80%99ecpad-met-en-ligne-les-photos-d%E2%80%99actualite-du-ministere-de-la-defense> (consulté pour la dernière fois le 2 juin 2013 à 15h39)].

<sup>180</sup> *Ibid.* (consulté pour la dernière fois le 2 juin 2013 à 15h39) et informations envoyées par Christophe CAZORLA, adjoint au chef du pôle archives de l'ECPAD, le 11 janvier 2013.

<sup>181</sup> Informations envoyées par Christophe CAZORLA le 11 janvier 2013.

<sup>182</sup> *Ibid.* La constitution de ce groupe de travail avait été demandée par la DICOd.

<sup>183</sup> Ministère de la Défense et des Anciens Combattants / ECPAD. *Guide de recommandations des formats de captation, indexation, diffusion, archivage de l'audiovisuel numérique de la défense*. [V2011\_10\_07]. Communiqué par mail par Christophe Cazorla, le 11 janvier 2013. Ce document a été envoyé au format Word, en plusieurs fichiers, et l'annexe 3 est manquante.

<sup>184</sup> *Ibid.*, dans la partie « 1. Recommandations pour la photographie », dans la sous-partie « 1.3. Indexation ».

- La date (inscrite automatiquement dans les métadonnées de prise de vue)
- Le copyright
- Les conditions d'utilisation, qui permettent « d'inscrire la classification ou la restriction des documents ainsi que la durée de cette restriction ».

La majorité des photographes professionnels de la défense utilise le logiciel Lightroom pour légender les images avec des métadonnées internes<sup>185</sup>. Cet outil est proposé par Adobe. Le Guide de recommandations pour l'audiovisuel numérique des OICD précise les champs IPTC à remplir dans Lightroom 3, et la manière de les remplir<sup>186</sup>.

Seuls les documents numériques natifs, depuis fin 2010, sont censés avoir leurs métadonnées renseignées selon ces recommandations<sup>187</sup>. Pour l'instant, il n'est pas demandé aux photographes de renseigner les champs Exif ou d'autres champs XMP. Le capitaine Christophe Cazorla affirme que, dans un futur proche, les métadonnées des photographies « argentiques » numérisées devraient être importées en XMP – pour l'instant, elles sont consignées au format XML<sup>188</sup>.

L'ECPAD met à disposition les photographies au format JPEG, pour consultation, et les archive au format TIFF<sup>189</sup>.

Pour l'instant (selon des tests que nous avons menés en janvier et en juin 2013 sur le site web de l'ECPAD), la rubrique « Armées Photos Actu » semble utiliser très peu les métadonnées internes pour renseigner l'internaute sur les photographies présentées<sup>190</sup>. Les images sont rassemblées en reportages, et accompagnées d'un commentaire correspondant à l'ensemble d'un reportage. Les seules métadonnées accessibles directement en ligne, et qui portent sur des photographies individuelles, sont présentées sous la forme d'infobulles<sup>191</sup> qui ne s'affichent à l'écran que quelques secondes.

<sup>185</sup> Informations envoyées par Christophe CAZORLA le 11 janvier 2013.

<sup>186</sup> Ministère de la Défense et des Anciens Combattants / ECPAD. *Guide de recommandations des formats de captation, indexation, diffusion, archivage de l'audiovisuel numérique de la défense*. [V2011\_10\_07], dans la partie « 1. Recommandations pour la photographie », dans la sous-partie « 1.3. Indexation ».

<sup>187</sup> Informations envoyées par Christophe CAZORLA le 11 janvier 2013.

<sup>188</sup> *Ibid.*

<sup>189</sup> Les fichiers Raw sont également archivés (informations envoyées par Christophe CAZORLA le 11 janvier 2013).

<sup>190</sup> « Armées Photos Actu ». ECPAD [site web] [En ligne : <http://www.ecpad.fr/categorie/actu> (lien vérifié le 1<sup>er</sup> juin 2013)]. Tests effectués les 4 et 6 janvier 2013 et le 2 juin 2013.

<sup>191</sup> Une infobulle est une petite zone de texte qui apparaît parfois sur des sites web quand le pointeur de la souris est placé sur un objet particulier (lien, image, etc.) : elle permet d'apporter des renseignements sur cet objet.

Les tests effectués ont mis en évidence plusieurs cas de figure (qui ne suffisent sûrement pas à caractériser de manière exhaustive les pratiques de l'ECPAD concernant son site web) :

- Certaines photographies ne comportent pas les métadonnées demandées par le Guide de recommandations pour l'audiovisuel numérique des OICD. Dans le cas que nous avons étudié<sup>192</sup>, la légende contenue dans les infobulles semble notamment être complétée avec le début du commentaire associé à l'ensemble du reportage (au-dessus des photographies).
- Certaines photographies comportent des informations internes qui ne sont pas fournies automatiquement par un appareil, par exemple dans le champ « Rights » – qui présente des renseignements sur la propriété intellectuelle, en XMP – ; mais ces métadonnées ne sont pas reprises telles quelles dans les infobulles. Dans l'exemple testé<sup>193</sup>, seul le titre du reportage a été repris pour légender les photographies dans leurs infobulles.
- D'autres photographies contiennent des métadonnées qui, apparemment, sont reprises sans modification notable dans les infobulles : ainsi, dans le reportage que nous avons étudié<sup>194</sup>, le champ XMP « Description » ou IPTC (IIM?) « Caption-Abstract » a été repris pour légender les photographies concernées.

Pour l'internaute désireux d'obtenir des renseignements précis sur une photographie, l'indexation (notamment à partir de métadonnées internes) est donc encore très limitée. Cependant, le capitaine Christophe Cazorla nous a indiqué<sup>195</sup> qu'une indexation automatique des photographies à partir de leurs métadonnées internes – sous une forme autre qu'une infobulle – était envisagée pour le futur système de consultation sur Internet des archives audiovisuelles de la Défense (indépendant du site de l'ECPAD).

---

<sup>192</sup> « Serval : Inspection du chef du CPCO sur le théâtre malien ». *ECPAD* [site web]. 27 mai 2013 [En ligne : <http://www.ecpad.fr/serval-inspection-du-chef-du-cpc0-sur-le-theatre-malien>]. Photographies testées le 2 juin 2013 avec Jeffrey's Exif Viewer [En ligne : <http://regex.info/exif.cgi>].

<sup>193</sup> « Point de situation sur les opérations de la force Serval depuis le 23 mai 18h jusqu'au 30 mai 18h ». *ECPAD* [site web]. 31 mai 2013 [En ligne : <http://www.ecpad.fr/point-de-situation-sur-les-operations-de-la-force-serval-depuis-le-23-mai-18-h-jusquau-30-mai-18h>]. Deux photographies présentées sur cette page ont été testées avec Jeffrey's Exif Viewer, le 2 juin 2013.

<sup>194</sup> « Bataille d'El Alamein, 1942 ». *ECPAD* [site web]. 18 octobre 2012 [En ligne : <http://www.ecpad.fr/bataille-del-alamein-1942>]. Deux photographies présentées sur cette page ont été testées avec Jeffrey's Exif Viewer, le 2 juin 2013.

<sup>195</sup> Informations envoyées par Christophe CAZORLA, adjoint au chef du pôle archives de l'ECPAD, le 11 janvier 2013.

## 2.4. La Bibliothèque municipale de Lyon

La Bibliothèque municipale de Lyon utilise les métadonnées internes pour attacher des informations aux photographies publiées dans la base *Photographes en Rhône-Alpes*<sup>196</sup>. Cette base<sup>197</sup> rassemble des images (nativement numériques ou numérisées) prises par des photographes amateurs ou professionnels, et qui concernent la région Rhône-Alpes<sup>198</sup>. L'enrichissement de cette banque d'images repose beaucoup sur les contributions de personnes qui déposent leurs photographies à la Bibliothèque (tout en restant libres de les exploiter comme ils l'entendent).

Les métadonnées internes déjà présentes dans les photographies reçues par la Bibliothèque (au moins pour la base *Photographes en Rhône-Alpes*) sont récupérées, et les informations manquantes sont créées par la Bibliothèque<sup>199</sup>. Voici le processus de documentation des photographies versées dans cette banque d'images (tel qu'il était au moins en décembre 2012)<sup>200</sup> :

- Des métadonnées IPTC sont saisies pour chaque photographie ;
- Ces métadonnées sont ensuite extraites vers Excel, puis complétées avec ce logiciel ;
- Le fichier Excel est ensuite converti dans des formats XML : METS et MODS.

METS<sup>201</sup> est un format qui permet de structurer des métadonnées diverses relatives à un document numérique. Il a principalement pour fonctions de permettre l'accès au document, sa conservation et l'échange de ressources numériques. Il peut inclure des informations exprimées dans d'autres formats – par exemple, pour les données bibliographiques, en MODS, en Dublin Core, en EAD, etc.

---

<sup>196</sup> Informations communiquées par mail par Anne MEYER, responsable du service Documentation Lyon et Rhône-Alpes à la Bibliothèque municipale de Lyon, le 27 décembre 2012. La bibliothèque utilise peut-être les métadonnées internes dans d'autres banques d'images, ou dans d'autres services, mais nous n'avons actuellement trouvé aucune information allant dans ce sens.

<sup>197</sup> La base *Photographes en Rhône-Alpes* est accessible sur le Web à l'adresse <http://collections.bm-lyon.fr/photo-rhone-alpes> (lien vérifié le 2 juin 2013).

<sup>198</sup> « A propos du projet ». *Photographes en Rhône-Alpes* [site web]. Bibliothèque municipale de Lyon [En ligne : <http://collections.bm-lyon.fr/photo-rhone-alpes/projet> (consulté le 2 juin 2013 à 16h08)].

<sup>199</sup> Informations communiquées par Anne MEYER le 27 décembre 2012.

<sup>200</sup> *Ibid.*

<sup>201</sup> « METS : Metadata Encoding and Transmission Standard ». *BnF* [site web]. 2 août 2012 [En ligne : [http://www.bnf.fr/fr/professionnels/formats\\_catalogage/a\\_f\\_mets.html](http://www.bnf.fr/fr/professionnels/formats_catalogage/a_f_mets.html) (consulté le 2 juin 2013 à 16h30)].

CLAERR, Thierry (dir.), WESTEEL, Isabelle (dir.). *Manuel de la numérisation*. Paris : Cercle de la librairie, 2012 (version e-book). Sous-partie « 2.2. Un format global de gestion des objets numériques : METS », dans la partie « VI. La gestion des collections numériques : vers une numérisation « durable » » (p. 286).

MODS<sup>202</sup> est un modèle destiné au traitement de données bibliographiques relatives à des documents de tout type et sur tous les supports. Il permet de décrire ces documents, en structurant les informations de description en XML ; il ne donne, par contre, aucune règle de description. MODS permet de structurer des données bibliographiques pour la création de catalogues ou de portails collectifs.

Les photographies sont ensuite versées dans la base *Photographies en Rhône-Alpes*.

Sur le site de *Photographies en Rhône-Alpes*, un fichier XML contenant des métadonnées METS et MODS des images est rendu accessible, à côté de la photographie correspondante, par un bouton « source XML »<sup>203</sup>. La page où est affichée l'image présente directement à l'internaute des informations telles que les « mots-clés », l'auteur ou les auteurs, le(s) donateur(s), des dates, etc. En revanche, les photographies testées ne possèdent pas dans leurs métadonnées internes toutes ces informations (seule une de ces photographies incluait, à l'intérieur du fichier, la mention de l'auteur de la photographie ; aucune ne contenait les informations présentées sur le site comme « mots-clés » ou comme lieux).

---

<sup>202</sup> « MODS : Metadata Object Description Schema ». *BnF* [site web]. 2 août 2012 [En ligne : [http://www.bnf.fr/fr/professionnels/f\\_mods/s\\_mods\\_presentation.html](http://www.bnf.fr/fr/professionnels/f_mods/s_mods_presentation.html) (consulté le 29 décembre 2012 à 18h26)].

<sup>203</sup> Tests effectués sur le site de *Photographies en Rhône-Alpes* le 2 juin 2013. Trois photographies ont été testées :  
- « [Panorama de la ville de Lyon] » [En ligne : [http://collections.bm-lyon.fr/BML\\_01ICO00101P0804\\_LOT001\\_7293?&query\[\]=source:"Vos photos"&hitStart=15&hitTotal=8375&hitPageSize=16](http://collections.bm-lyon.fr/BML_01ICO00101P0804_LOT001_7293?&query[]=source:) ]  
- « [Un père et son nouveau né] » [En ligne : [http://collections.bm-lyon.fr/BML\\_01ICO001014e5de80a89a87?&query\[\]=source:"Marcelle Vallet \(1907-2000\)"&query\[\]=serie\\_s:"Le bidonville du Chaâba"&hitStart=17&hitTotal=56&hitPageSize=16](http://collections.bm-lyon.fr/BML_01ICO001014e5de80a89a87?&query[]=source:) ]  
- « Cygnes sur le Rhône au Confluent » [En ligne : [http://collections.bm-lyon.fr/BML\\_01ICO0010150f70e9a4c44f](http://collections.bm-lyon.fr/BML_01ICO0010150f70e9a4c44f) ].

### III. LA PRESSE, L'ÉDITION ET LES PHOTOGRAPHES PROFESSIONNELS

Les deux secteurs où les métadonnées internes aux images sont les plus utilisées, actuellement, sont la presse et l'édition<sup>204</sup> – l'utilisation de ces informations étant moins répandue et moins cruciale dans ce dernier domaine<sup>205</sup>.

Il faut rappeler ici que le format IPTC/IIM a acquis une grande importance dans le secteur de la presse – à la suite de son intégration partielle dans Photoshop –, et que ce standard a été élaboré par un regroupement d'organisations de ce domaine, l'International Press Telecommunications Council, qui comprend aujourd'hui certaines des principales agences de presse mondiales, telles que Reuters, Associated Press (AP) et l'Agence France-Presse (AFP)<sup>206</sup>. Cependant, il est possible que des images présentées sur les sites de ces agences ne contiennent pas ces informations<sup>207</sup>. Il est par contre fondamental que ces données soient présentes pour l'échange, l'import-export d'images<sup>208</sup>.

Il faut noter l'attachement de la presse et de l'édition au standard IPTC/IIM, et ce malgré le fait que ce format a été remplacé par l'IPTC Core<sup>209</sup>. Comme tous les outils professionnels gèrent à la fois l'IIM et XMP, ces acteurs peuvent continuer à utiliser l'IIM avec lequel ils se sont habitués à travailler.

Les photographes qui travaillent avec la presse utilisent de plus en plus les informations internes aux images – ils sont au moins obligés d'y être sensibilisés (toujours selon Patrick Peccatte). Il existe deux possibilités : soit ces photographes travaillent seuls, et donc le remplissage des champs de ces métadonnées leur incombe ; soit ils travaillent dans un groupe, un organisme, ce qui peut leur donner l'opportunité de pouvoir faire réaliser cette tâche par d'autres personnes<sup>210</sup>.

Il est de toute façon essentiel pour les agences photos d'être capables de fournir à leurs clients des images accompagnées de légendes<sup>211</sup>.

---

<sup>204</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>205</sup> *Ibid.*

<sup>206</sup> Voir la présentation de l'IPTC (International Press Telecommunications Council) dans le présent mémoire, page 21.

Nous avons trouvé un article, publié en 2010, qui indique que Reuters, Associated Press et l'Agence France-Presse sont les trois premières agences d'information mondiales, sans que nous sachions ce qui justifie cette affirmation :

LAVILLE, Camille. « L'AFP : une agence d'avenir ? ». *InaGlobal* [site web]. 8 septembre 2010, mis à jour le 11 octobre 2010 [En ligne : <http://www.inaglobal.fr/presse/article/lafp-une-agence-davenir> (consulté le 12 juin 2013 à 22h18)].

<sup>207</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>208</sup> *Ibid.*

<sup>209</sup> *Ibid.*

<sup>210</sup> *Ibid.*

<sup>211</sup> *Ibid.*

Les métadonnées contenues dans les photographies de presse peuvent avoir une grande importance pour la constitution de contenus dans une publication. Patrick Peccatte (qui a travaillé pour des éditeurs de presse<sup>212</sup>) témoigne ainsi de l'utilisation des informations IPTC par un journal :

« J'ai vu jadis dans un journal des colonnes de brèves illustrées créées quasi-automatiquement à partir de photos d'agences sélectionnées et composées avec leurs légendes IPTC (je dis "quasi" car le rédacteur devait souvent reprendre les légendes...).<sup>213</sup> »

Une telle utilisation des métadonnées des photographies se fait toujours, selon lui, très couramment<sup>214</sup> : les colonnes de brèves sont souvent fabriquées en éditant des descriptions, avec des aménagements. Cette pratique a cours avec les brèves illustrées, et beaucoup avec les dépêches d'agences<sup>215</sup>.

---

<sup>212</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>213</sup> PECCATTE, Patrick, dans un commentaire publié le 25 juillet 2010 sur la page « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 17 novembre à 14h58)].

<sup>214</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>215</sup> *Ibid.*



## **TROISIÈME PARTIE : ÉLÉMENTS SUR LES CHAMPS D'APPLICATION ET SUR LES POSSIBILITÉS D'ÉVOLUTIONS DES MÉTADONNÉES INTERNES AUX IMAGES**

---

Ce dernier chapitre cherche à donner des éléments sur les possibilités d'évolution et d'utilisation des métadonnées internes aux images. Il s'agit à la fois de résumer certaines applications actuelles de ces informations, d'en proposer d'autres, et de voir comment ces métadonnées pourraient évoluer techniquement, tout en abordant les questions autour de l'intérêt ou du désintérêt que suscitent ces informations chez les acteurs susceptibles de les utiliser.

### **I. LES CHAMPS D'APPLICATION POSSIBLES DES MÉTADONNÉES INTERNES AUX IMAGES**

Cette partie s'appuie beaucoup sur les explications et exemples présentés dans les deux premiers chapitres de ce mémoire, afin d'exposer ou d'imaginer certains domaines de compétence des métadonnées internes aux images fixes – c'est-à-dire de répondre (en partie) à cette simple question : à quoi peuvent servir ces métadonnées ?

Les données enregistrées dans une image numérique peuvent renseigner à la fois sur l'image elle-même, sans considération de son support, et sur le fichier qui la transporte. Reste à déterminer quelle peut être l'utilité d'informations incluses dans les fichiers (sous une forme standardisée).

#### **1.1. L'association étroite entre l'image et ses métadonnées : transport et préservation des informations**

Cette sous-partie n'aborde pas la question de l'accès aux images grâce à un processus d'indexation. Cependant, il nous paraissait utile d'expliquer pourquoi il peut être intéressant de travailler avec les métadonnées internes pour conserver et transporter des informations : il s'agit de montrer dans quel contexte ces données

peuvent être récupérées pour indexer des documents. C'est ce que nous proposons de faire dans ce développement.

Le but essentiel des métadonnées internes est d'éviter de perdre l'association entre les informations sur l'image et l'image elle-même<sup>216</sup>.

Lorsque, dans un système informatisé, les métadonnées d'un document sont distinctes du fichier dans lequel est enregistré ce document, alors elles sont conservées dans une base de données : dans ce cas, elles doivent être exportées puis associées à la bonne ressource chaque fois que celle-ci est consultée<sup>217</sup>. Il existe un risque que ces informations soient perdues lors d'un transfert<sup>218</sup>. Ce problème semble résolu avec des données incluses dans le fichier, puisqu'elles suivent systématiquement (en théorie) les images qu'elles documentent.

Cette association étroite entre l'image et ses métadonnées semble donc dispenser de recourir à d'autres processus de documentation des images. Nous verrons plus loin qu'il n'en est rien.

La mise en place d'un ou de standard(s)<sup>219</sup> se justifie par la nécessité que tous les acteurs concernés par les images numériques s'entendent au moins sur des aspects techniques qui leur permettent de lire et d'écrire des métadonnées. La définition de champs dédiés à des catégories de données précises permet de structurer l'information pour mieux la retrouver en fonction de besoins documentaires spécifiques. Cela peut aussi, peut-être, inciter les personnes qui manipulent ces métadonnées à fournir des informations plus précises et complètes, en leur proposant un canevas, une sorte de formulaire qu'elles n'auraient plus qu'à compléter : ainsi, cela permettrait aux utilisateurs des images de retrouver toujours à peu près les mêmes types d'informations (en fonction, toujours, de leurs besoins documentaires), y compris en l'absence de règles de documentation explicites imposées aux personnes ou aux organisations qui créent ou publient ces images. Ce dernier avantage n'est qu'une hypothèse de notre part, et d'ailleurs ne se vérifie pas dans tous les cas<sup>220</sup>.

---

<sup>216</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>217</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], pp. 40-41.

<sup>218</sup> *Ibid*, pp. 40-41.

<sup>219</sup> Le mot est entendu au sens (informatique) de « spécification documentée permettant de développer des produits interopérables » (PECCATTE, Patrick, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].).

<sup>220</sup> Voir notamment l'exemple de l'ECPAD, pp. 48-51 de ce mémoire.

Plusieurs contextes d'utilisation des images rendent les métadonnées internes utiles, voire nécessaires.

Pour Patrick Peccatte, avoir la description de l'image en même temps que l'image elle-même est fondamental pour toute utilisation industrielle, en particulier pour les agences photos<sup>221</sup>.

Les services ou organisations qui mettent à disposition des photographies sur Internet, et permettent aux utilisateurs de les télécharger, ont certainement intérêt à utiliser les métadonnées internes pour documenter ces images. Il faut d'abord considérer, en effet, que l'utilisateur qui télécharge une photographie peut souhaiter obtenir en même temps les informations relatives à cette image : or, la notice présentée sur la page internet incluant la photographie ne suffit pas si cette personne veut enregistrer l'image en dehors du site. Inclure des métadonnées à l'intérieur du document semble pouvoir éviter de télécharger l'image et ses métadonnées dans des fichiers séparés ; cependant, il faut que l'utilisateur puisse accéder à ces métadonnées, or il n'est pas certain qu'il connaisse les outils pour lire ou extraire ces informations, ou qu'il en soit équipé<sup>222</sup>. Mettre à disposition, sur le site internet, une notice téléchargeable présentant les informations contenues dans la photographie, peut être une solution plus efficace pour certains utilisateurs.

Dans ce contexte de mise à disposition d'images téléchargeables sur Internet, les métadonnées peuvent servir à apporter des preuves dans le cas, par exemple, d'une incertitude relative à l'image ou à son support, ou dans le cas d'une utilisation malhonnête. Les informations contenues dans le fichier peuvent servir à éviter qu'une photographie passe pour ce qu'elle n'est pas. La description de l'image, par exemple, permet ainsi d'identifier ce qui a été photographié. De même, les mentions de propriété intellectuelle et les conditions d'utilisation restent attachées au fichier, et permettent de rappeler ce qu'il est possible de faire légalement avec ce document. Cependant, tout cela suppose que l'image ne soit pas convertie dans un format ne supportant pas les métadonnées, ou ne soit pas manipulée par un outil qui ne les conserve pas, ou encore

<sup>221</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>222</sup> Si l'on en croit André GUNTHERT, la gestion des métadonnées (en tout cas, jusqu'à début 2012) pourrait ne faire l'objet que d'une « faible préoccupation, y compris du côté des professionnels » (GUNTHERT, André, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)]).

D'autres personnes ont évoqué le fait que les métadonnées internes des images sont [étaient ?] peu connues et/ou peu utilisées (par exemple, BÉAT BRÛSCH (graphiste et illustrateur) dans un commentaire publié le 25 janvier 2011 sur la page « Les photos d'agence de presse sur les sites d'actualité, entre illustration et confusion », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/509> (consulté le 7 juin 2013 à 11h08)]). Patrick Peccatte affirme que peu de monde est intéressé par le fait d'avoir, sur *Flickr*, toutes les tailles d'une même image renseignées avec les métadonnées de la photographie originale (Entretien avec Patrick PECCATTE le 23 janvier 2013.)

que les informations qu'elle contient ne soient pas volontairement modifiées, ce qui est toujours possible<sup>223</sup>.

Le fait de documenter une photographie de l'intérieur améliore – ou sécurise – sa qualité patrimoniale, c'est-à-dire surtout sa valeur historique ou de témoignage. Il est évident qu'une image ne naît pas forcément à des fins de transmission, et qu'elle peut n'avoir, au moins au départ, qu'une visée illustrative ou artistique. Mais, pour reprendre un vocabulaire employé dans le monde des archives, un document (créé par un ou des être(s) humain(s)<sup>224</sup>) peut avoir une valeur primaire – qui est celle de la fonction pour laquelle il a été créé – et une valeur secondaire, qui n'est pas (forcément) prévue à l'origine. Cette valeur secondaire peut être, par exemple, une valeur de source historique. Ainsi, toute photographie peut témoigner d'un événement, de l'existence de quelque chose, de pratiques, etc.

Associer à une image numérique une sorte d'étiquette peut aider à accroître cette valeur de témoignage ou de souvenir. La préservation de ces informations internes est d'autant plus facilitée que les formats de métadonnées utilisés ont fait l'objet de spécifications en accès public.

Du côté des photographes amateurs, ou des professionnels qui ne travaillent pas dans le domaine de la recherche ou du patrimoine, cet aspect n'est probablement pas perçu comme prioritaire – d'autant plus que documenter une image prend du temps<sup>225</sup> – ; mais les acteurs chargés de la conservation et de la transmission d'informations à caractère patrimonial ou historique (tels que les musées, les services d'archives, les bibliothèques, etc.) devraient, selon nous, s'intéresser à cette démarche. Ainsi, nous pensons que ces institutions pourraient, autant que possible, renseigner les métadonnées des photographies, ou un certain nombre de ces métadonnées, en fonction des informations qu'ils ont à leur disposition. Un service d'archives n'a certes pas vocation à modifier un document dont il a la responsabilité, mais il est possible d'utiliser les métadonnées internes au moins pour des copies de documents (effectuées notamment pour renforcer la sécurité de la préservation ou pour faciliter l'accès à certaines ressources). Cette

---

<sup>223</sup> Voir pp. 32-34 de ce mémoire.

<sup>224</sup> Nous tenons à rappeler que nous ne cherchons pas ici à prendre position dans le débat sur la signification du terme « document » en sciences de l'information. Toutefois, à cet endroit de notre développement, cette précision sur l'origine du document paraît nécessaire pour notre propos.

<sup>225</sup> DEHESDIN, Thierry, (photographe), dans un commentaire publié le 23 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

pratique pourrait s'appliquer aussi bien à des copies de documents déjà numériques qu'à des fichiers issus d'une numérisation. Cependant, encore une fois, il faut garder à l'esprit les dangers encourus en cas d'utilisation d'outils non adaptés, ce qui pourrait arriver, par exemple, lors de migrations de formats de fichiers<sup>226</sup>. Nous pensons que la duplication des métadonnées, c'est-à-dire leur enregistrement à la fois dans l'image elle-même et dans un autre fichier (par exemple un outil de recherche en XML) peut être une solution à étudier pour sécuriser au mieux l'information relative aux photographies conservées.

Certains établissements ont déjà adopté les métadonnées internes pour l'archivage, comme l'ECPAD : dans ce cas précis, les métadonnées, renseignées par les photographes, sont extraites des fichiers<sup>227</sup>.

## **1.2. Indexation et renseignement des métadonnées**

L'un des problèmes des métadonnées internes aux images est le temps nécessaire au remplissage des champs. Amélie Segonds écrit que « le nombre massif et croissant de documents » – notamment de photographies – mis en ligne sur le Web interdit une saisie manuelle de descripteurs<sup>228</sup>. Certes, l'exemple de *MédiHAL*<sup>229</sup> montre qu'il est possible d'appliquer automatiquement des métadonnées à plusieurs images (à partir d'un seul fichier) ; cependant, cela ne fait gagner du temps que dans le cas de documents proches par leur sujet, leur origine, ou d'autres caractéristiques, suivant ce que l'on souhaite inclure comme types d'informations.

Le projet PhotosNormandie montre peut-être comment il est possible de faire participer des personnes qui ne sont pas forcément des professionnels de l'image, de l'informatique ou de la documentation, à l'indexation de photographies et au remplissage de leurs métadonnées. Ce projet<sup>230</sup> a pour objectif d'améliorer les légendes de photographies de la bataille de Normandie (été 1944), en mettant en ligne ces images sur *Flickr*, et en prenant en compte des propositions d'amélioration faites par des membres

---

<sup>226</sup> Effectuer une migration de format signifie changer le format d'enregistrement d'un document, par exemple pour le conserver dans un format plus moderne. Ce type d'opération peut se révéler nécessaire quand le format dans lequel est enregistré l'image vient à n'être plus utilisé, ou n'est plus supporté par des logiciels.

<sup>227</sup> Voir pp. 48-51 de ce mémoire.

<sup>228</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 42.

<sup>229</sup> Voir p. 46 de ce mémoire.

<sup>230</sup> Les informations sur ce projet proviennent de la page de présentation de « PhotosNormandie » sur *Flickr* [site web] [En ligne : <http://www.flickr.com/people/photosnormandie/> (consultée le 8 juin 2013 à 23h24)].

de *Flickr* : les modifications proposées qui sont validées sont intégrées aux métadonnées IPTC/IIM des images correspondantes. Plus de 3 000 photographies sont concernées. Selon la page de présentation du projet<sup>231</sup>, plus de 7 000 descriptions d'images ont été complétées ou corrigées (certaines étant mises à jour plusieurs fois) depuis le 29 janvier 2007.

Il est envisageable de renseigner automatiquement les informations internes d'images grâce aux contributions de personnes, sans que celles-ci soient obligatoirement familières des standards de métadonnées. Il est possible d'imaginer un système où des internautes feraient, comme pour le projet PhotosNormandie, des propositions pour compléter ou modifier les informations qui documentent les images mises en ligne – cela pourrait se faire sous plusieurs formes : commentaires, ajouts de tags, remplissage de champs pré-définis, etc. Ces contributions seraient ensuite intégrées (éventuellement après avoir été validées), de manière automatique, à l'intérieur des photographies. Cela permettrait ainsi d'automatiser le remplissage de certains champs de métadonnées, sans que les contributeurs soient forcément obligés de connaître les métadonnées internes aux images et les pratiques techniques qui leur sont associées. Picasa, logiciel permettant notamment la mise en ligne de photographies, inscrivait des coordonnées cartographiques dans les données Exif d'une image quand celle-ci était positionnée manuellement dans Google Earth (au moins jusqu'en 2009)<sup>232</sup>.

Les services de partage de photographies sur Internet ont tout intérêt à récupérer les informations déjà contenues dans les photographies, car cela évite aux internautes qui souhaitent mettre en ligne des images de renseigner ces informations manuellement.

Patrick Peccatte souligne que les données concernant le modèle d'appareil, la date et les coordonnées GPS peuvent être utilisées par le grand public pour le classement et la recherche, et cite *Flickr* comme exemple<sup>233</sup>.

L'un des champs d'application importants des métadonnées internes est celui du positionnement et de la contextualisation géographiques du document, grâce

---

<sup>231</sup> Page de présentation de « PhotosNormandie ». *Flickr* [site web] [En ligne : <http://www.flickr.com/people/photosnormandie/> (consultée le 8 juin 2013 à 23h24)].

<sup>232</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], pp. 87-88.

<sup>233</sup> PECCATTE, Patrick, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

notamment aux informations de géolocalisation. Pour Patrick Peccatte, les demandes de possibilités d'association entre des images et une représentation cartographique sont amenées à se multiplier<sup>234</sup>. Ces pratiques sont liées aux possibilités techniques offertes, par exemple, par les portables, notamment par l'iPhone dont le succès est croissant<sup>235</sup>.

Les coordonnées GPS inscrites dans les métadonnées internes des photographies peuvent, nous l'avons vu, être récupérées pour positionner les images correspondantes sur une carte<sup>236</sup>.

Selon Amélie Segonds<sup>237</sup>, la géolocalisation est devenue une pratique répandue sur les sites de partage d'images. Les sites qui offrent la possibilité d'associer des informations géographiques aux photographies pour le positionnement et la recherche sont devenus nombreux, et, depuis le début des années 2000<sup>238</sup> (et jusqu'en 2009 au moins), le nombre de personnes qui utilisent la géolocalisation pour indexer et classer des ressources iconographiques a beaucoup augmenté<sup>239</sup>.

Certains sites spécialisés peuvent profiter de ces données, par exemple des services en ligne destinés à fournir conseils ou témoignages pour des voyages touristiques.

L'information géographique renseignée automatiquement par les appareils photographiques, les téléphones mobiles, etc., présente au moins un risque : il est possible de subir une confusion entre le lieu de prise de vue et l'endroit montré sur l'image générée. Si, par exemple, une personne prend une photographie du Mont Blanc depuis Lyon, et que son appareil est équipé d'un GPS, les coordonnées indiquées dans l'image créée correspondront à Lyon.

Au-delà des services de partage d'images, des entreprises pourraient trouver intéressant de proposer des publicités ciblées en fonction de données de géolocalisation, et non plus seulement en s'adaptant aux recherches des internautes<sup>240</sup>. Il est certainement possible d'analyser les informations contenues dans les photographies mises en ligne par une personne, ou consultées par cette personne, pour proposer des contenus –

<sup>234</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>235</sup> *Ibid.*

<sup>236</sup> Voir notamment les exemples de *Flickr* et de *MédiHAL*, pp. 41-42 et 46-47 de ce mémoire.

<sup>237</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 83.

<sup>238</sup> Période dès laquelle des laboratoires (notamment américains) ont mené des travaux pour créer des plates-formes de partage d'images où la position géographique serait un critère de classement de la photographie et un moyen d'y accéder. SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 82.

<sup>239</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHER. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 83.

<sup>240</sup> *Ibid.*, p. 100.

notamment publicitaires – portant sur des services en relation avec la position indiquée<sup>241</sup>.

Plus généralement, des personnes ou des organisations (entreprises, établissements publics, collectivités, etc.) qui souhaitent vendre ou délivrer gratuitement des services, peuvent certainement trouver une manière de cibler leur public à partir de métadonnées internes, sans se limiter aux informations géographiques. Ces services pourraient être de différentes natures : publicités pour des destinations touristiques, adresses de restaurants à proximité d'un lieu déterminé, information sur un monument, etc.<sup>242</sup>

## II. QUELQUES ÉVOLUTIONS POSSIBLES DANS LE DOMAINE DES MÉTADONNÉES INTERNES AUX IMAGES

### **2.1. La place des métadonnées internes dans les pratiques relatives aux images : entre désintérêt et pédagogie**

En février 2012, André Gunthert affirme qu'« aucune métadonnée n'est véritablement mobilisée de manière active pour la recherche et le classement des images dans les usages grand public » – mis à part les données de datation des fichiers –, et, plus généralement, semble penser que la tendance est à une « faible préoccupation, y compris du côté des professionnels, pour la gestion des métadonnées »<sup>243</sup>. Amélie Segonds note, en 2009, que les systèmes d'indexation tels que les métadonnées internes aux images – appliqués au Web – sont peu connus du grand public, du fait du « savoir faire » et de la maîtrise des « outils informatiques » qu'il supposent<sup>244</sup>. Patrick Peccatte, quant à lui, affirme que peu de monde est intéressé par la préservation des métadonnées dans les images postées

---

<sup>241</sup> Nous n'avons pas vérifié si ces quelques utilisations que nous proposons, inspirées par Amélie SEGONDS (voir son mémoire, notamment pp. 100-103, dont la référence est donnée dans la note 244), existent déjà ou peuvent s'appuyer sur des systèmes déjà existants.

<sup>242</sup> Même remarque que dans la note précédente.

<sup>243</sup> GUNTHERT, André, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118>] (consulté le 15 novembre 2012 à 22h43)].

<sup>244</sup> SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHERT. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722>], p. 42.

sur *Flickr* (du moins en ce qui concerne les versions dérivées d'une photographie)<sup>245</sup>. Nous avons vu aussi que, globalement, les réseaux sociaux *Facebook* et *Twitter* ne cherchaient pas, au moins jusqu'en 2012, à conserver les informations présentes dans les images mises en ligne sur leurs sites<sup>246</sup>.

En ce qui concerne les photographes professionnels, Thierry Dehesdin, photographe indépendant, affirme que le remplissage des champs IPTC, s'il est précis et détaillé, demande un temps bien plus long que celui nécessaire à la prise de vue, et cela entraîne un coût supplémentaire que les clients ne sont presque jamais prêts à payer. De plus, ce travail est un « exercice pour lequel peu de photographes possèdent une attirance naturelle »<sup>247</sup>.

Ces discours ne suggèrent pas un avenir rayonnant pour les métadonnées internes aux images. Pourtant, ces informations sont utilisées, y compris pour l'indexation des documents iconographiques et dans des contextes qui peuvent concerner le grand public.

Des organisations à mission patrimoniale, comme l'ECPAD ou la Bibliothèque municipale de Lyon<sup>248</sup>, ont décidé de renseigner ou de faire renseigner les métadonnées des photographies, et se servent de ces informations – ou projettent de s'en servir – pour documenter ces images sur des sites web.

Par ailleurs, des sites de partage de photographies généralistes récupèrent les informations contenues dans les images pour fournir un accès à ces documents, que ce soit par le biais d'une recherche par mots ou d'une navigation sur une représentation cartographique. Certains de ces sites offrent même un accès facilité aux métadonnées des images, comme c'est le cas pour *Flickr* ou *Ipernity*<sup>249</sup>, par exemple.

Certes, il faut peut-être relativiser l'importance de ces utilisations. Sur *Flickr*, par exemple, seuls 3 534 446 éléments sont « géotaggés »<sup>250</sup> (le 8 juin 2013 à 17h33), pour un volume de plusieurs milliards de documents iconographiques dans tout le site<sup>251</sup>. De plus, nous ne sommes pas à même de fournir ici des estimations du nombre de personnes qui utilisent les extractions de métadonnées pour des recherches de contenus. Néanmoins, il faut remarquer que les applications telles que l'indexation cartographique

<sup>245</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>246</sup> Voir pp. 43-44 de ce mémoire.

<sup>247</sup> DEHESDIN, Thierry, dans un commentaire publié le 25 juillet 2010 sur la page « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté le 7 juin 2013 à 18h30)]. Thierry Dehesdin travaille surtout pour l'industrie et la publicité (voir la page de son blog : <http://blog.dehesdin.com/about/> (consultée le 13 juin 2013 à 16h11)).

<sup>248</sup> Voir pp. 48-51 et 52-53 de ce mémoire.

<sup>249</sup> Voir pp. 39-40 de ce mémoire.

<sup>250</sup> Interface cartographique de *Flickr* [En ligne : <http://www.flickr.com/map> (consultée le 8 juin 2013 à 17h33)].

<sup>251</sup> En février 2012, Patrick PECCATTE estime à environ 5 milliards le nombre de photographies mises en ligne sur *Flickr* (PECCATTE, Patrick. *Le projet The Commons sur la plate-forme Flickr* [diaporama]. Paris : Séminaire Muséologie 2.0 : Institut de Recherche et d'Innovation. 14 février 2012 [Disponible en ligne sur la page : <http://culturevisuelle.org/dejavu/1116> (consulté le 8 juin 2013 à 17h38)], *slide 2*).

ou la présentation des métadonnées des images ne sont pas isolées (au sens où elles ne sont pas proposées par un seul service, mais par un certain nombre d'entre eux).

Pour que des photographes (amateurs ou professionnels) renseignent les métadonnées internes de leurs images, il faut certainement – sans que cela se limite, bien entendu, à ces conditions – leur présenter des explications sur ces données, et sur le rôle qu'elles tiennent dans le fonctionnement d'applications intéressantes. *Flickr* et *Ipernity* se sont avancés sur cette voie, en adoptant une démarche pédagogique envers leurs utilisateurs. En effet, ces deux sites donnent accès à des explications sur les métadonnées internes :

- *Flickr* propose, sur les pages affichant les métadonnées des images<sup>252</sup>, un petit paragraphe intitulé « Que sont les données Exif ? », dans lequel il est expliqué que ces données sont des informations sur les « paramètres utilisés par un appareil photo pour prendre une photo ou une vidéo », et que ces informations sont « incorporées aux fichiers enregistrés par l'appareil photo »<sup>253</sup>.
- *Ipernity* affiche, sur la même page que les métadonnées<sup>254</sup>, un lien vers la page de *Wikipédia* consacrée au format Exif<sup>255</sup>.

Ces explications ont leurs limites, car elles ne portent que sur les métadonnées Exif – ce qui est d'autant plus regrettable que les métadonnées affichées, au moins par *Flickr*, ne sont pas forcément enregistrées au format Exif<sup>256</sup>.

Début 2012, André Gunthert disait avoir « l'impression générale d'un caractère hyperspécialisé et d'une insécurité des formats de métadonnées, globalement encore non stabilisés », ce à quoi Patrick Peccatte répondait que l'univers des métadonnées [internes aux images ?] n'était « pas moins stabilisé que bien d'autres technologies informatiques plus connues », par exemple HTML ou des formats de traitement de texte, « sur une période de temps très comparable »<sup>257</sup>.

---

<sup>252</sup> Voir pp. 39-40 de ce mémoire.

<sup>253</sup> Tests effectués sur trois photographies : les pages consultées ont pour adresses :

[http://www.flickr.com/photos/eric\\_fischer/7924450222/meta/](http://www.flickr.com/photos/eric_fischer/7924450222/meta/)  
<http://www.flickr.com/photos/sytenes/5813836162/meta/>  
[http://www.flickr.com/photos/saffron\\_blaze/8936275752/meta/](http://www.flickr.com/photos/saffron_blaze/8936275752/meta/)  
(consultées le 8 juin 2013 à 18h10).

<sup>254</sup> Voir p. 40 de ce mémoire.

<sup>255</sup> Tests effectués sur trois photographies : les pages consultées ont pour adresses :

<http://www.ipernity.com/doc/grauminay/4017717/infos>  
[http://www.ipernity.com/doc/declic\\_67/13695675/infos](http://www.ipernity.com/doc/declic_67/13695675/infos)  
<http://www.ipernity.com/doc/afulki/3048425/infos>  
(consultées le 8 juin 2013 à 18h21).

<sup>256</sup> Voir pp. 39-40 de ce mémoire.

<sup>257</sup> GUNTHERT, André, et PECCATTE, Patrick, dans des commentaires publiés le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consultés le 15 novembre 2012 à 22h43)].

De notre point de vue, les métadonnées internes des images peuvent pâtir d'une complexité due à l'existence de plusieurs standards dont les champs de compétence peuvent se chevaucher. La convergence vers un seul standard prédominant, qui répondrait à la plupart des besoins documentaires des photographes et des chercheurs d'images, viendrait peut-être simplifier cet univers des métadonnées, et inciter les acteurs concernés à manipuler ces informations.

## **2.2. Regards sur les possibilités d'évolution technique des métadonnées internes aux images**

Cette dernière sous-partie a pour objectif de présenter quelques possibles évolutions techniques des métadonnées internes aux documents iconographiques numériques, telles que les décrivent des spécialistes de l'image ou des informations internes aux images.

Pour André Gunthert et Patrick Peccatte, les pratiques documentaires relatives aux images fixes et aux vidéos pourraient converger dans l'avenir<sup>258</sup>. André Gunthert écrit ainsi que ces pratiques ont « vocation à se rapprocher ».

Les standards dans le domaine de l'image animée sont beaucoup moins avancés et répandus que dans le secteur de l'image fixe<sup>259</sup>. Adobe, qui n'a pas dans ce domaine de position dominante comme dans le secteur de l'image fixe, a développé des formats basés sur XMP<sup>260</sup>.

Patrick Peccatte affirme que les utilisateurs éventuels ne peuvent qu'espérer une convergence du formalisme entre les deux univers, mais rappelle que les problématiques documentaires concernant la vidéo sont plus complexes que celles de l'image fixe – car il faut à la fois décrire l'ensemble d'un film et les scènes qui le composent.

---

<sup>258</sup> GUNTHERT, André, dans un commentaire publié le 21 février 2012 sur la page « La France forte et ses métadonnées », sur le site web *Déjà vu. Carnet de recherche visuel* [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 15 novembre 2012 à 22h43)].

Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>259</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>260</sup> *Ibid.*

Patrick Peccatte pense également que, dans le domaine des métadonnées telles que XMP, IPTC et Exif, la tendance est de se diriger vers une domination de XMP<sup>261</sup>. De plus en plus, les outils modernes de post-traitement gèrent XMP ; il semble donc naturel, selon lui, de prévoir une évolution où l'IPTC/IIM serait remplacé, et où le format Exif, sans être abandonné, serait au moins reproduit en XMP. Début 2012, il affirmait que le standard Exif 2.3 metadata for XMP<sup>262</sup> pouvait accélérer la « migration » vers XMP<sup>263</sup>, en notant toutefois qu'une conversion totale des métadonnées descriptives en XMP était ralentie par le « poids des habitudes » et le « manque d'intérêt des professionnels » pour les possibilités de ce format<sup>264</sup>. XMP est basé sur XML, qui est, pour Patrick Peccatte, une technologie arrivée à maturité, dont il n'est plus concevable de se passer en informatique : XML présente l'avantage d'être simple et relativement souple du point de vue de sa structuration, tout en permettant de réaliser des choses très complexes.

Phil Harvey, créateur de l'outil ExifTool<sup>265</sup>, se montre plus critique vis-à-vis de XMP<sup>266</sup>. Pour lui, le fait que ce format repose sur XML est beaucoup plus compliqué que nécessaire. Il écrit que XMP se révèle difficile à implémenter correctement dans son intégralité, la complexité de la spécification faisant que les outils d'édition de métadonnées ne le supportent pas bien. De plus, ce format n'accepte que des informations textuelles, ce qui le rend inadéquat pour la conservation d'informations binaires – comme une image en résolution réduite, notamment<sup>267</sup>.

Patrick Peccatte envisage d'autres évolutions pour les métadonnées internes aux images<sup>268</sup>. La reconnaissance faciale pourrait, selon lui, prendre de l'importance, en utilisant ces données. Adobe a déjà commencé à vouloir exprimer en XMP les informations nécessaires à la reconnaissance des visages<sup>269</sup>.

---

<sup>261</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>262</sup> Voir *E. Les métadonnées Exif 2.3 metadata for XMP*, p. 28 de ce mémoire.

<sup>263</sup> PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté le 2 décembre 2012 à 15h10)].

<sup>264</sup> *Ibid.*

<sup>265</sup> Voir pp. 31-32 de ce mémoire pour une brève présentation d'ExifTool.

<sup>266</sup> « Commentary on Meta Information Formats (or "Why this or that Format Sucks") ». *ExifTool by Phil Harvey* [site web]. Page créée le 15 novembre 2005, mise à jour le 29 juillet 2010 [En ligne : <http://owl.phy.queensu.ca/~phil/exiftool/commentary.html> (consulté le 14 avril 2013 à 17h07)].

<sup>267</sup> *Ibid.* (consulté le 14 avril 2013 à 17h07), et réponse à une de nos questions envoyée par mail par Phil HARVEY le 10 juin 2013.

<sup>268</sup> Entretien avec Patrick PECCATTE le 23 janvier 2013.

<sup>269</sup> *Ibid.*

Une autre évolution pourrait être<sup>270</sup> de faire supporter à XMP des arborescences de « mots-clés ». Actuellement, ce format ne permet pas de structurer les mots-clés en plusieurs niveaux. Cette nouvelle possibilité, très demandée, pourrait intéresser notamment des acteurs qui ont besoin de travailler avec des métadonnées très riches : il s'agit en général d'acteurs professionnels, par exemple dans les secteurs culturels. Ce type de technologie existe déjà, mais elle est propriétaire.

---

<sup>270</sup> Toujours selon Patrick PECCATTE (Entretien du 23 janvier 2013).



## CONCLUSION

---

Le travail de recherche exposé dans ce mémoire, bien qu'inégalement approfondi selon les champs explorés, a permis de mettre en évidence, notamment grâce à l'étude d'exemples concrets, le rôle que peuvent tenir les métadonnées internes dans l'indexation d'images numériques.

Ces données, associées étroitement à la ressource qu'elles documentent, permettent de transporter et de conserver les informations sur une image en même temps et au même endroit que l'image elle-même – à condition de veiller à utiliser des outils qui gèrent les formats des métadonnées inscrites dans le document iconographique que l'on veut manipuler. Ces avantages nous poussent à dire que les métadonnées internes sont des éléments qu'il est intéressant de prendre en compte dès qu'il s'agit de gérer des images destinées à être archivées ou partagées. Dans ce cadre d'utilisation, indexer des images à partir de ces données est tout à fait possible, et ce de manière automatisée : les informations internes peuvent notamment être extraites du fichier, ou servir à positionner sur une carte un document géolocalisé, ce qui permet d'accéder aux images en les recherchant selon plusieurs caractéristiques.

Les exemples d'utilisation que nous avons proposés dans la deuxième partie concernent très majoritairement des services en ligne, y compris pour les acteurs du patrimoine, de la culture et de la recherche que nous avons présentés. Il serait donc pertinent de prolonger ces recherches en s'intéressant aux usages hors Internet : il faudrait examiner, entres autres, les pratiques des photographes – professionnels et amateurs –, celles du secteur de l'édition, et voir si des établissements, des organismes, des institutions, des collectivités ou des entreprises se servent des métadonnées internes dans d'autres contextes que celui d'un réseau comme le Web.

Le contexte technique des métadonnées pourrait aussi être plus précisément étudié. Il pourrait être intéressant d'aller plus loin que ce que nous avons ébauché, en dressant un panorama plus complet des types d'outils de gestion de ces données. Le fonctionnement des processus d'enregistrement de ces données peut aussi être approfondi, ainsi que la problématique de la sécurité des informations dans les fichiers – c'est-à-dire surtout du maintien de leur intégrité.

La pertinence des champs documentaires proposés par les principaux standards cités, dans une optique d'indexation des images, pourrait aussi être un sujet d'étude : les informations à renseigner correspondent-elles aux besoins des utilisateurs ?

Enfin, il reste à aborder la question de l'évolution de ces métadonnées et de leurs usages. Le paysage des informations internes des images fixes est-il amené à être dominé par un seul standard – et, si oui, quelles conséquences pourraient en découler ? Les pratiques documentaires dans ce domaine vont-elles converger vers celles qui concernent la vidéo ? Des recherches prospectives sont réalisables, qu'elles soient orientées vers les modifications des formats ou vers les usages à venir des métadonnées internes aux images.

# Sources

## I/ Informations sur des logiciels

- ExifTool :

Page d'accueil du site *ExifTool by Phil Harvey* [site web] [En ligne : <http://www.sno.phy.queensu.ca/~phil/exiftool/> (consultée pour la dernière fois le 11 juin 2013 à 17h16)].

« An interview with Phil Harvey » [interview de Phil Harvey en 2005] [En ligne : <http://owl.phy.queensu.ca/~phil/exiftool/PhilHarveyInterview.pdf> (consultée le 6 avril 2013, lien vérifié le 11 juin 2013 à 16h06)].

- GD :

« About. What is the GD library ? » *GD Library* [site web] [En ligne : <http://libgd.bitbucket.org/pages/about.html> (consultée le 6 juin 2013 à 17h05)].

- ImageMagick :

Page d'accueil du site web officiel d'ImageMagick.  
[En ligne : <http://www.imagemagick.org/script/index.php> (consultée le 6 juin 2013 à 17h40)].

## II/ Sources sur des utilisateurs et des utilisations des métadonnées internes aux images

### 1. Page de présentation de la formation *Les métadonnées des images numériques [IPTC et XMP]*

« Les métadonnées des images numériques fixes ». *Soft Experience* [site web] [En ligne : <http://peccatte.karefil.com/software/IPTCXMP Cours.html> (consultée pour la dernière fois le 18 décembre 2012 à 15h06)].

### 2. Étude sur la préservation des métadonnées internes des images dans des médias sociaux sur le Web

« Social Media Photo Metadata use Survey : CV Web Photo Metadata » [En ligne : <https://spreadsheets.google.com/pub?key=tceeIYNw8ZDC0N52UgRcgnA&single=true&gid=0&output=html> (consultée le 9 mai 2013 à 14h58)].

### 3. Sites web de partage de photographies

Pour ces sites, nous ne citons pas toutes les pages consultées. Le lecteur peut trouver plus de détails dans le développement du mémoire.

- Sur *Flickr* (<http://www.flickr.com/>), nous avons notamment consulté les pages suivantes :

- « Aide / FAQ / La carte ». *Flickr* [site web] [En ligne : <http://www.flickr.com/help/map/> (consultée le 12 juin 2013 à 2h11)].

- Interface cartographique de *Flickr* [En ligne : <http://www.flickr.com/map> (consultée le 8 juin 2013 à 17h33)].

- « La recherche par appareil photo ». *Flickr* [site web] [En ligne : <http://www.flickr.com/cameras/> (consultée le 7 juin 2013 à 23h00)].

- Page de présentation de « PhotosNormandie ». *Flickr* [site web] [En ligne : <http://www.flickr.com/people/photosnormandie/> (consultée le 8 juin 2013 à 23h24)].

- Sur *Ipernity* (<http://www.ipernity.com/>), plusieurs pages ont été consultées (le 8 juin 2013), dont le détail est donné dans le développement de ce mémoire.

- Sur *Zizics* (<http://www.zizics.com/>), nous avons notamment consulté la page d'accueil :

- Page d'accueil de *Zizics* [site web] [En ligne : <http://www.zizics.com/> (consultée le 8 juin 2013 à 12h13)].

#### 4. Organismes utilisant des métadonnées internes pour des images fixes

- La Bibliothèque municipale de Lyon :

Des informations nous ont été communiquées par mail par Mme Anne MEYER, responsable du service Documentation Lyon et Rhône-Alpes à la Bibliothèque municipale de Lyon, le 27 décembre 2012.

Sur le site web de *Photographes en Rhône-Alpes* (banque d'images de la Bibliothèque municipale de Lyon [En ligne : <http://collections.bm-lyon.fr/photo-rhone-alpes/>]), nous avons notamment consulté la page suivante :

- « A propos du projet ». *Photographes en Rhône-Alpes* [site web]. Bibliothèque municipale de Lyon [En ligne : <http://collections.bm-lyon.fr/photo-rhone-alpes/projet> (consultée le 2 juin 2013 à 16h08)].

D'autres pages ont été consultées : le détail est donné dans le développement du mémoire.

- L'ECPAD (Établissement de Communication et de Production Audiovisuelle de la Défense) :

Des informations et des documents nous ont été communiqués par mail par le Capitaine Christophe CAZORLA, adjoint au chef du pôle archives de l'ECPAD, le 11 janvier 2013.

Parmi ces documents, nous avons consulté la ressource suivante : Ministère de la Défense et des Anciens Combattants / ECPAD. *Guide de recommandations des formats de captation, indexation, diffusion, archivage de l'audiovisuel numérique de la défense*. [V2011\_10\_07]. (Ce document a été envoyé au format Word, en plusieurs fichiers, et l'annexe 3 est manquante).

Sur le site web de l'ECPAD (<http://www.ecpad.fr/>), nous avons notamment consulté les pages suivantes :

- « L'ECPAD met en ligne les photos d'actualité du ministère de la défense ». *ECPAD* [site web]. 23 février 2012 [En ligne : <http://www.ecpad.fr/1%E2%80%99ecpad-met-en-ligne-les-photos-d%E2%80%99actualite-du-ministere-de-la-defense> (consultée pour la dernière fois le 2 juin 2013 à 15h39)].

- « Présentation ». *ECPAD* [site web]

[En ligne : <http://www.ecpad.fr/lecpad/presentation> (consultée le 1<sup>er</sup> juin 2013 à 17h38)].

Des tests ont également été effectués dans la rubrique « Armées Photos Actu » du site :

- « Armées Photos Actu ». *ECPAD* [site web] [En ligne : <http://www.ecpad.fr/categorie/actu> (lien vérifié le 1<sup>er</sup> juin 2013)].

Les derniers tests datent du 2 juin 2013. Des détails supplémentaires sont donnés dans le développement du mémoire.

Nous avons aussi consulté la page suivante sur le site du ministère de la Défense :

- « La Délégation à l'information et à la communication de la défense (DICOd) ». Site web du Ministère de la Défense. Mise à jour le 13 mars 2013 [En ligne : <http://www.defense.gouv.fr/portail-defense/ministere/organisation/organisation-du-ministere-de-la-defense/organismes-dependant-du-ministre/la-delegation-a-l-information-et-a-la-communication-de-la-defense-dicod> (consultée le 12 juin 2013 à 15h26)].

- *MédiHAL* (archive ouverte d'images scientifiques) :

Sur le site web de *MédiHAL* (<http://medihal.archives-ouvertes.fr/>), nous avons notamment consulté les pages suivantes :

- Page d'accueil de *MédiHAL* [site web] [En ligne : <http://medihal.archives-ouvertes.fr/> (consultée pour la dernière fois le 15 décembre 2012 à 19h29)].

- « Aide en ligne ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?action\\_todo=faq#videos](http://medihal.archives-ouvertes.fr/index.php?action_todo=faq#videos) (consultée le 15 décembre 2012)].

- « Consulter>Par domaine ». *MédiHAL* [site web] [En ligne : [http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action\\_todo=browse&b\\_type=browse\\_domain](http://medihal.archives-ouvertes.fr/index.php?halsid=dp9aur36mr71hq0gk0vf6oi282&action_todo=browse&b_type=browse_domain) (consultée pour la dernière fois le 26 mai 2013)].

D'autres pages ont également été consultées : des détails supplémentaires sont donnés dans le développement du mémoire.

Pour nous documenter sur *MédiHAL*, nous avons également consulté les ressources suivantes :

- Page d'accueil de *CCSD* [site web] [En ligne : <http://ccsd.cnrs.fr/> (consultée le 12 juin 2013 à 10h12)].

- Page d'accueil de *CN2SV. Centre National pour la Numérisation de Sources Visuelles* [site web] [En ligne : <http://www.cn2sv.cnrs.fr/> (consultée le 12 juin 2013 à 10h26)].

- « About GeoNames ». *GeoNames* [site web] [En ligne : <http://www.geonames.org/about.html> (consultée le 12 juin 2013 à 11h42)].

### III/ Spécifications et sites web d'organisations actives dans le développement de formats de métadonnées internes aux images fixes

#### 1. IPTC (International Press Telecommunications Council), pour les métadonnées IPTC/IIM, IPTC Core et IPTC Extension

« IPTC Core & Extension = the IPTC Photo Metadata Standard ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0099> (consultée pour la dernière fois le 8 juin 2013 à 16h18)].

« IIM Information Interchange Model ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/cms/site/index.html?channel=CH0134> (consultée le 2 décembre 2012 à 15h28)].

« IIM Information Interchange Model. IIM for photos : a good solution for more than a decade ». *International Press Telecommunications Council* [site web] [En ligne : [http://www.iptc.org/site/Photo\\_Metadata/IIM/](http://www.iptc.org/site/Photo_Metadata/IIM/) (consultée pour la dernière fois le 10 juin 2013 à 22h12)].

« The IPTC... as a membership organisation ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/site/Home/About/> (consultée le 11 mai 2013 à 14h39)].

« IPTC Members ». *International Press Telecommunications Council* [site web] [En ligne : <http://www.iptc.org/site/Home/Members/> (consultée le 11 mai 2013 à 15h)].

IPTC. *IPTC Standard Photo Metadata. Document Revision 1.* (Juillet 2010) International Press Telecommunications Council, 2010 [En ligne : [http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007\\_1.pdf](http://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata-201007_1.pdf) (lien vérifié le 12 juin 2013 à 12h46)].

RIECKS, David. *"IPTC Core" Schema for XMP. Version 1.0. Supplemental documentation. Custom Panels User Guide. Document Revision 13.* 2005

[En ligne :

[http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore\\_1.0-doc-CpanelsUserGuide\\_13.pdf](http://www.iptc.org/std/Iptc4xmpCore/1.0/documentation/Iptc4xmpCore_1.0-doc-CpanelsUserGuide_13.pdf) (lien vérifié le 13 juin 2013 à 23h27)].

#### 2. JEITA (Japan Electronics and Information Technology industries Association), pour les métadonnées Exif

« JEITA's Activities and Organization. What is JEITA ? » *JEITA* [site web] [En ligne : <http://www.jeita.or.jp/english/about/what/index.htm> (consultée le 10 juin 2013 à 22h45)].

JEITA. *Exchangeable image file format for digital still cameras : Exif Unified Version 2.3*. Préparé par : AV&IT Standardization Committee. Japan Electronics and Information Technology Industries Association. Révisé en avril 2010 [En ligne : [http://www.jeita.or.jp/japanese/standard/book/CP-3451B\\_E/#page=3](http://www.jeita.or.jp/japanese/standard/book/CP-3451B_E/#page=3) (lien vérifié le 11 juin 2013 à 9h50)].

(Cette spécification est consultable sur le site du JEITA à partir de la page « JEITA Standards. AV&IT Technology Standardization. Digital Cameras » [En ligne : [http://www.jeita.or.jp/cgi-bin/standard\\_e/list.cgi?cateid=1&subcateid=4](http://www.jeita.or.jp/cgi-bin/standard_e/list.cgi?cateid=1&subcateid=4) (consultée le 11 juin 2013 à 0h07)].)

#### **IV/ Informations communiquées par des spécialistes des métadonnées internes aux images**

- Entretien d'environ 1h30 avec M. Patrick PECCATTE, chercheur associé au Lhivic (Laboratoire d'histoire visuelle contemporaine) de l'EHESS, et créateur de la société Soft Experience, le 23 janvier 2013.
- Informations communiquées par mail par M. David RIECKS le 10 mai 2013. M. RIECKS a notamment participé à l'élaboration de l'IPTC Core et de l'IPTC Extension.
- Informations communiquées par mail par M. Phil HARVEY, créateur de l'outil ExifTool, les 6 et 10 juin 2013.



## ***Bibliographie et webographie***

(Nous n'indiquons pas ici les références des commentaires consultés sur certains sites web, comme par exemple sur *Déjà vu. Carnet de recherche visuel*. Ces informations sont notées dans le développement du mémoire).

### **I/ Ressources générales sur les métadonnées, des standards et des formats**

*Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008. 282 p.

(Nous avons également consulté un chapitre de cet ouvrage disponible en ligne :

DALBIN, Sylvie. « Représentation et accès à l'information : transformation à l'œuvre », dans *Métadonnées : mutations et perspectives*. Séminaire INRIA, 29 septembre-3 octobre 2008. Ouvrage coordonné par CALDERAN, Lisette, HIDOINE, Bernard, MILLET, Jacques. Paris : ADBS, 2008 [En ligne : [http://www.atd-doc.com/xmedia/publications/INRIA-2008\\_DALBIN\\_Chap01.pdf](http://www.atd-doc.com/xmedia/publications/INRIA-2008_DALBIN_Chap01.pdf) (lien vérifié le 14 juin 2013 à 1h23)].)

CLAERR, Thierry (dir.), WESTEEL, Isabelle (dir.). *Manuel de la numérisation*. Paris : Cercle de la librairie, 2012 (version e-book).

FOULONNEAU, Muriel, RILEY, Jenn. *Metadata for digital resources : implementation, systems design and interoperability*. Oxford : Chandos Publishing, 2008. 203 p.

### **II/ Indexation**

*L'indexation des ressources pédagogiques numériques : questions transversales*. Compte-rendu rédigé par l'ENSSIB à partir d'une transcription de la communication orale de Ghislaine CHARTRON du 16 novembre 2004. 6 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-1241> (lien vérifié le 14 juin 2013 à 1h48)].

Association des archivistes français. *Abrégé d'archivistique. Principes et pratiques du métier d'archiviste*. Paris : Association des archivistes français, 2012. 346 p.

COLLARD, Claude (dir.), MELOT, Michel (dir.). *Images et bibliothèques*. Paris : Éditions du Cercle de la librairie, 2011. 240 p.

LÉVESQUE, Michel. « L'indexation : luxe ou nécessité ? ». *Archives*, volume 33, n°1, 2001-2002, pp. 17-45  
[En ligne : [http://www.archivistes.qc.ca/revuearchives/vol33\\_1/33-1-levesque.pdf](http://www.archivistes.qc.ca/revuearchives/vol33_1/33-1-levesque.pdf) (lien vérifié le 14 juin 2013 à 8h43)].

SEGONDS, Amélie. *Indexation visuelle et recherche d'images sur le Web. Enjeux et problèmes*. Mémoire de master 2 dirigé par André GUNTHERT. Lhivic/EHESS, 2009. 164 p. [En ligne : <http://www.enssib.fr/bibliotheque-numerique/document-48722> (lien vérifié le 14 juin 2013 à 1h53)].

### III/ Métadonnées internes aux images fixes

« Commentary on Meta Information Formats (or "Why this or that Format Sucks") ». *ExifTool by Phil Harvey* [site web]. Page créée le 15 novembre 2005, mise à jour le 29 juillet 2010

[En ligne : <http://owl.phy.queensu.ca/~phil/exiftool/commentary.html> (consulté le 14 avril 2013 à 17h07)].

« The Controlled Vocabulary Survey regarding the Preservation of Photo Metadata by Social Media Websites ». *Controlled Vocabulary* [site web] [En ligne : <http://www.controlledvocabulary.com/socialmedia/index.html> (consulté le 9 mai 2013 à 15h05)].

CONTASSOT, Barbara. *Usages, avantages et inconvénients du standard IPTC*. Rapport de recherche bibliographique dirigé par Dominique CHICHEREAU. ENSSIB, 2005. 83 f.

FIGER, Jean-Paul. « Comment classer ses photos ? EXIF, IPTC, XMP et autres métadonnées : de la théorie à la pratique ». *Le PC simple* [site web]. Mis à jour le 3 décembre 2007 [En ligne : <http://www.figer.com/publications/collection.htm> (consulté le 12 mai 2013 à 16h17)].

GEORJON, Emmanuel. « Photos numériques : présentation des métadonnées ». *Emmanuel Georjon. Photo numérique et technologies Web* [site web]. 10 octobre 2008 [En ligne : <http://www.emmanuelgeorjon.com/photos-numeriques-presentation-des-metadonnees-270/> (consulté pour la dernière fois le 7 décembre 2012 à 18h04)].

PECCATTE, Patrick. « La France forte et ses métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 17 février 2012 [En ligne : <http://culturevisuelle.org/dejavu/1118> (consulté pour la dernière fois le 2 décembre 2012 à 17h59)].

PECCATTE, Patrick. « Métadonnées : une initiation. Dublin Core, IPTC, Exif, RDF, XMP, etc. ». *Soft Experience* [site web]. Août 2002, mis à jour le 13 décembre 2007 [En ligne : <http://peccatte.karefil.com/software/metadata.htm> (consulté pour la dernière fois le 11 juin 2013 à 14h27)].

PECCATTE, Patrick. « Photo Metadata toolkit IPTC-PLUS, la gestion intégrée des métadonnées ». *Déjà vu. Carnet de recherche visuel* [site web]. 24 juillet 2010 [En ligne : <http://culturevisuelle.org/dejavu/273> (consulté pour la dernière fois le 7 juin 2013 à 18h30)].

PECCATTE, Patrick. « Les photos d'agence de presse sur les sites d'actualité, entre illustration et confusion ». *Déjà vu. Carnet de recherche visuel* [site web]. 25 janvier 2011 [En ligne : <http://culturevisuelle.org/dejavu/509> (consulté le 7 juin 2013 à 11h08)].

#### **IV/ Compléments sur des formats, des langages ou des utilisateurs des métadonnées internes aux images**

« Introduction to XML ». *W3schools.com* [site web] [En ligne : [http://www.w3schools.com/xml/xml\\_what\\_is.asp](http://www.w3schools.com/xml/xml_what_is.asp) (consulté le 1<sup>er</sup> décembre 2012 à 16h18)].

« METS : Metadata Encoding and Transmission Standard ». *BnF* [site web]. 2 août 2012 [En ligne : [http://www.bnf.fr/fr/professionnels/formats\\_catalogage/a.f\\_mets.html](http://www.bnf.fr/fr/professionnels/formats_catalogage/a.f_mets.html) (consulté le 2 juin 2013 à 16h30)].

« MODS : Metadata Object Description Schema ». *BnF* [site web]. 2 août 2012 [En ligne : [http://www.bnf.fr/fr/professionnels/f\\_mods/s\\_mods\\_presentation.html](http://www.bnf.fr/fr/professionnels/f_mods/s_mods_presentation.html) (consulté le 29 décembre 2012 à 18h26)].

LAVILLE, Camille. « L'AFP : une agence d'avenir ? ». *InaGlobal* [site web]. 8 septembre 2010, mis à jour le 11 octobre 2010 [En ligne : <http://www.inaglobal.fr/presse/article/lafp-une-agence-davenir> (consulté le 12 juin 2013 à 22h18)].

PECCATTE, Patrick. *Le projet The Commons sur la plate-forme Flickr* [diaporama]. Paris : Séminaire Muséologie 2.0 : Institut de Recherche et d'Innovation. 14 février 2012 [Disponible en ligne sur la page : <http://culturevisuelle.org/dejavu/1116> (consulté le 8 juin 2013 à 17h38)].



# *Glossaire*

## **Profil ICC**

Le profil ICC est un profil colorimétrique intégré dans une image RVB.

Une image RVB est une image dont les couleurs sont restituées à partir de trois composantes : rouge (R), verte (V) et bleue (B). Pour que ce codage puisse représenter des couleurs précises, il faut adjoindre des informations complémentaires dans un profil ICC.

La structure de ce profil a été standardisée par l'ICC (International Color Consortium).

(Source : CLAERR, Thierry (dir.), WESTEEL, Isabelle (dir.). *Manuel de la numérisation*. Paris : Cercle de la librairie, 2012 (version e-book).)



## *Table des illustrations*

Illustration 1: Photographie testée avec l'outil « Jeffrey's Exif Viewer » (photographie de l'auteur).....	19
Illustration 2: Métadonnées de la photographie présentée en Illustration 1 (page précédente), affichées par « Jeffrey's Exif Viewer ». Capture d'écran.....	20
Illustration 3: Affichage de métadonnées internes d'une photographie sur Flickr. Capture d'écran (2 juin 2013).....	40
Illustration 4: Interface cartographique de Flickr. Capture d'écran (4 juin 2013).....	42
Illustration 5: Exemple de notice sur MédiHAL, avec le positionnement géographique de la photographie sur une carte. Capture d'écran (26 mai 2013).....	47
Illustration 6: Exemple de notice sur MédiHAL (2) : notice détaillée. Photographie identique à celle de l'Illustration 5. Capture d'écran (26 mai 2013).....	47



# Table des matières

SIGLES ET ABRÉVIATIONS.....	9
INTRODUCTION.....	11
<b>PREMIÈRE PARTIE : LES MÉTADONNÉES INTERNES AUX IMAGES FIXES ET LEUR CONTEXTE TECHNIQUE .....</b>	<b>19</b>
<b>I. Présentation des métadonnées internes aux images fixes.....</b>	<b>19</b>
1.1. <i>Présentation des standards IPTC/IIM, Exif, XMP, IPTC Core, IPTC Extension et Exif 2.3 metadata for XMP.....</i>	21
A. Les métadonnées IPTC/IIM.....	21
B. Les métadonnées Exif.....	23
C. Les métadonnées XMP.....	24
D. Les métadonnées IPTC Core et IPTC Extension .....	26
E. Les métadonnées Exif 2.3 metadata for XMP.....	28
<b>II. Comparaisons, champs d'application et conjonction des métadonnées IPTC, Exif et XMP.....</b>	<b>29</b>
2.1. <i>Exif et IPTC.....</i>	29
2.2. <i>IPTC/IIM et XMP.....</i>	30
<b>III. Exploitation des métadonnées .....</b>	<b>31</b>
3.1. <i>Les outils de gestion des métadonnées internes aux images.....</i>	31
3.2. <i>Archivage des métadonnées et indexation.....</i>	33
<b>DEUXIÈME PARTIE : EXEMPLES D'UTILISATION DES MÉTADONNÉES INTERNES AUX IMAGES.....</b>	<b>35</b>
<b>I. Réseaux sociaux et services de partage de photographies en ligne.....</b>	<b>35</b>
1.1. <i>Les services de partage de photographies en ligne.....</i>	37
1.2. <i>Les réseaux sociaux : Facebook et Twitter.....</i>	43
<b>II. Des acteurs de la recherche, de la culture et du patrimoine.....</b>	<b>44</b>
2.1. <i>L'archive ouverte MédiHAL.....</i>	45
2.3. <i>L'ECPAD (Établissement de Communication et de Production Audiovisuelle de la Défense).....</i>	48
2.4. <i>La Bibliothèque municipale de Lyon.....</i>	52
<b>III. La presse, l'édition et les photographes professionnels.....</b>	<b>54</b>
<b>TROISIÈME PARTIE : ÉLÉMENTS SUR LES CHAMPS D'APPLICATION ET SUR LES POSSIBILITÉS D'ÉVOLUTIONS DES MÉTADONNÉES INTERNES AUX IMAGES.....</b>	<b>57</b>
<b>I. Les champs d'application possibles des métadonnées internes aux images .....</b>	<b>57</b>
1.1. <i>L'association étroite entre l'image et ses métadonnées : transport et préservation des informations.....</i>	57
1.2. <i>Indexation et renseignement des métadonnées.....</i>	61
<b>II. Quelques évolutions possibles dans le domaine des métadonnées internes aux images.....</b>	<b>64</b>
2.1. <i>La place des métadonnées internes dans les pratiques relatives aux images : entre désintérêt et pédagogie.....</i>	64
2.2. <i>Regards sur les possibilités d'évolution technique des métadonnées internes aux images .....</i>	67
<b>CONCLUSION.....</b>	<b>71</b>

<b>SOURCES.....</b>	<b>73</b>
<b>BIBLIOGRAPHIE ET WEBOGRAPHIE.....</b>	<b>79</b>
<b>GLOSSAIRE.....</b>	<b>83</b>
<b>TABLE DES ILLUSTRATIONS.....</b>	<b>85</b>
<b>TABLE DES MATIÈRES.....</b>	<b>87</b>