

Médiation entre l'utilisateur et l'image :
Une étude exploratoire de l'utilisation des métadonnées par les professionnels

Mariane Léonard
École des sciences de l'information
Université d'Ottawa

Sous la supervision
de la Dre Mary Cavanagh
et du Dr Tom Delsey

Thèse présentée à la
Faculté des études supérieures et postdoctorales
dans le cadre des exigences
de la maîtrise en sciences de l'information (MSI)

© Mariane Léonard, Ottawa, Canada, 2015

Résumé

Avec la popularité grandissante de la mise en ligne d'images, il importe de se questionner sur le processus de repérage dans le but d'améliorer les résultats. En ce qui concerne notre recherche, nous nous intéressons au repérage d'images par un professionnel dans le contexte d'un service référence. Notre étude tente de comprendre comment les professionnels effectuant le service de référence dans un musée ou une université qui font de la recherche d'images ou d'informations sur une image utilisent les métadonnées associées à ces dernières.

La collecte de données a été effectuée grâce à trois méthodes : un journal de bord, une entrevue semi-dirigée et une observation. Même si notre étude est de type exploratoire, il a été possible d'affirmer que les professionnels ont un rôle de médiateur important et que le type d'établissement dans lequel ils travaillent a une incidence sur la nature du rôle de médiateur.

Les résultats de la collecte de données prouvent que la sélection de l'outil pour le repérage dépend du type d'établissement. Par ailleurs, la formulation de la recherche est étroitement liée avec les termes présents dans la requête de l'utilisateur et cette étape détermine le niveau de médiation des professionnels. L'examen de l'usage des métadonnées par les professionnels a démontré que les éléments identifiés comme utiles varient selon le type de requêtes.

Mots clés : repérage d'images, métadonnées, médiation, professionnels

Abstract

With the growing popularity of the online publication of images, it is important to examine the image retrieval process in order to improve results. In our research, we focus on image retrieval by professionals in the context of reference services. Our study seeks to understand how professionals providing reference services in museums and universities search for images or information on an image using the metadata associated with the image.

Data collection was done using three methods: a log book, a semi-structured interview and observation. Although our study is exploratory in nature, it has been possible to assert that professionals have a significant role as mediators and that the type of institution in which they work has an impact on the nature of that role.

The results of the data collection have demonstrated that the selection of the search tool is directly linked with the type of institution. Furthermore, the formulation of the search is closely linked with the terms found in the user's request, and this step determines the level of mediation by the professional. The analysis of the use of metadata by professionals has shown that the elements identified as useful differ according to the type of request.

Keywords : metadata, image retrieval, mediation, professional

Remerciements

L'accomplissement de cette thèse n'aurait pas été possible sans l'appui et le support de plusieurs personnes. Tout d'abord, je voudrais remercier mes directeurs, la Dre Mary Cavanagh et le Dr Tom Delsey, pour leur soutien, leur aide, leur compréhension, leur patience, leurs conseils et j'en saute. Je crois que je n'aurais pas pu être mieux dirigé, ils sont pour moi des modèles et ils ont su me transmettre leur passion.

De plus, il va de soi que sans les participants, la recherche n'aurait pas eu lieu. Je les remercie infiniment, c'est grâce à vous que tout a été possible. Merci pour votre temps et votre aide.

Je voudrais remercier ma famille qui a toujours été là pour moi, plus précisément, mes parents, André et Andrée, qui sont, selon moi, les meilleurs parents qu'on peut avoir; ma sœur, MJ, qui a toujours su m'encourager et mon copain, Fred, pour sa patience. Ils ont su me supporter dans ce projet et sans eux je n'y serais pas arrivé.

Je voudrais également remercier mes amis, plus précisément mon amie Isa qui a su m'écouter quand j'avais besoin de parler de mon projet, je pouvais lui transmettre mes peurs et mes inquiétudes et elle a été une écoute bien chaleureuse. Merci également d'avoir été là quand j'avais besoin de décompresser.

Aussi, un petit mot pour remercier Jeanne, qui a été pour moi une source d'inspiration en tant que femme. Sa force et sa détermination m'ont guidée lors de mes longues journées en solitude.

Table des matières

Chapitre 1	Introduction.....	1
1.1	Formulation du problème.....	2
1.2	Objectif de la recherche.....	3
1.3	Questions de recherche.....	5
1.4	Méthodologie.....	5
1.5	Étendue et limite de la recherche.....	6
1.6	Plan.....	7
Chapitre 2	Revue de la littérature et cadre conceptuel.....	8
2.1	La recherche sur le repérage d'images.....	9
2.1.1	La sémantique et la syntactique de l'image.....	9
2.1.2	Le comportement de l'utilisateur.....	12
2.1.3	La médiation par les professionnels.....	14
2.1.4	Schémas de métadonnées.....	15
2.1.5	L'analyse de la requête.....	17
2.2	Cadre conceptuel.....	20
2.2.1	Fonctionnalités requises des notices bibliographiques (FRNB).....	22
2.2.2	Pyramide.....	24
2.3	Conclusion.....	26
Chapitre 3	Méthodologie.....	27
3.1	Revue d'éthique.....	28
3.2	Échantillonnage et recrutement.....	28
3.3	Rencontre d'information.....	30
3.4	Participants.....	30
3.5	Mode de collecte de données.....	31
3.5.1	Journal de bord.....	31
3.5.2	Entrevue.....	33
3.5.3	Observation et prise de note.....	34
3.6	Conservation des données et confidentialité.....	35
3.7	Analyse de données.....	35
3.7.1	Validation.....	36
3.7.2	Limitations.....	37
3.8	Conclusion.....	38
Chapitre 4	Résultats.....	40
4.1	Journaux de bord.....	40
4.1.1	Trouver.....	42
4.1.2	Identifier.....	48
4.1.3	Sélectionner.....	49
4.1.4	Obtenir.....	52
4.2	Entrevues.....	53
4.3	Observation.....	54
4.4	Conclusion.....	56
Chapitre 5	Discussion.....	58

5.1	Le contexte.....	58
5.2	L'analyse de la requête	59
5.3	Médiation	61
5.3.1	Sélection de l'outil.....	61
5.3.2	Formulation de la requête.....	62
5.3.3	L'utilisation des métadonnées	64
5.4	Conclusion	66
Chapitre 6	Conclusion	67
6.1	Résumé.....	67
6.2	Recherches futures	68
6.3	Critique de la recherche	69
6.4	Réflexion finale.....	70
	Liste de références.....	71
	Appendices.....	77

Liste des figures

Figure 2.1 Cadre de recherche	21
Figure 2.2 La pyramide	25
Figure 4.1 Métadonnées utiles et inutiles.....	50
Figure 4.2 Métadonnées utiles et inutiles.....	55

Liste des tableaux

Tableau 3.1 Description des participants	31
Tableau 4.1 Ressources utilisées pour la recherche	41
Tableau 4.2 Analyse des mots clés utilisés dans la requête	44
Tableau 4.3 Analyse des mots clés utilisés dans les recherches	45
Tableau 4.4 Analyse des termes utilisés pour la recherche.....	46
Tableau 4.5 Nombre de résultats selon la recherche lancée.....	47
Tableau 4.6 Utilité des métadonnées	51
Tableau 4.7 Métadonnées utiles pour la recherche d'image	53
Tableau 4.8 Utilité des métadonnées pendant l'observation.....	55
Tableau 5.1 Comparaison des niveaux de la pyramide de la requête et la recherche	63

Liste des acronymes

FRNB : *Fonctionnalités requises des notices bibliographiques*

Liste des appendices

Appendice A : Journal de bord.....	78
Appendice B : Questions d’entrevue.....	81
Appendice C : Certificat d’approbation déontologique	82
Appendice D : Courriel d’invitation à participer	84
Appendice E : Description du projet.....	85
Appendice F : Formulaire de consentement.....	88
Appendice G : Liste des codes.....	91
Code 4.2.1.1	91
Code 4.2.1.2	92
Code 4.2.1.2 – 4.2.2.2 - 4.2.3.1	92
Code 4.2.1.3	94
Code 4.2.1.4	94
Code 4.2.4.1	94
Appendice H : Exemple de code.....	95
Appendice I : Classification pyramide des mots clés.....	96

Chapitre 1 Introduction

À l'époque actuelle, les images, qu'elles soient nées numériquement ou qu'elles proviennent d'une représentation numérisée, deviennent de plus en plus nombreuses, importantes, mais également populaires. Non seulement elles se retrouvent sur les bases de données scientifiques, artistiques, culturelles, universitaires, mais aussi, dans plusieurs domaines, elles sont utilisées, partagées, mais par-dessus tout, elles sont créées.

Pour surenchérir sur l'ampleur d'images disponibles sur le Web ou dans les bases de données, des statistiques croissantes de certains moteurs de recherches ou de bases de données d'images illustrent très bien cette affirmation. En effet, Google Images a indexé entre 2001 et 2010 10 billions d'images (Smith, 2010). Quant aux bases de données à abonnement, Artstor¹ contient plus de 280 collections et qui une fois que les divers contributeurs auront mis toutes les images en ligne en comprendra plus de 2,624,291. Précisons également que ces collections sont disponibles grâce à 267 contributeurs qui numérisent et déposent leurs images sur cette base de données.

Enfin, en tenant compte de ces statistiques, il est indéniable que les images numériques sont plus que présentes sur le Web. Par conséquent, le partage de ces dernières

¹ Description de Artstor « The Artstor Digital Library is a nonprofit resource that provides over 1.8 million digital images in the arts, architecture, humanities, and sciences with an accessible suite of software tools for teaching and research. Our community-built collections comprise contributions from outstanding international museums, photographers, libraries, scholars, photo archives, and artists and artists' estates. »(Artstor, n.d.)

via divers sites et bases de données a alors créé un nouveau besoin dans l'indexation de celles-ci (Jørgensen, Jaimes, Benitez, & Chang, 2001, p. 938). Conséquemment à la présence et la production des images numériques, ces dernières doivent être retrouvables, et ceci devient possible grâce à plusieurs méthodes d'indexation, thésaurus et schémas de métadonnées.

1.1 Formulation du problème

Avec la technologie qui s'améliore sans cesse, par exemple, grâce aux connexions Internet de plus en plus rapides permettant un téléchargement et un téléversement plus efficace, la numérisation des images et leur transmission vers des bases de données d'images se font beaucoup plus rapidement et facilement qu'avant (Kramer, 2005, p. 3). Avec de plus en plus d'images qui se retrouvent sur le Web avec les techniques de numérisation qui sont de plus en plus populaires et faciles à exécuter, les bibliothèques, les musées et les centres d'archives numérisent leurs collections d'images qu'ils possèdent sur papier ou sur film pour les rendre disponibles sur le Web pour que les usagers puissent les retracer eux-mêmes; ils créent des collections virtuelles. Les images numériques ajoutées à des bases se doivent d'être trouvables sinon leur présence sur le Web devient inutile.

Lorsque les usagers des bibliothèques, des musées et des centres d'archives ne trouvent pas l'image qu'ils recherchent, certains d'entre eux se tournent vers des intermédiaires, c'est-à-dire des professionnels, pour avoir de l'aide pour repérer leurs images. Le professionnel doit faire un travail d'interprétation pour obtenir les termes justes pour trouver l'objet désiré. Il interprète les informations reçues des participants en métadonnées. Ainsi, une meilleure compréhension du travail du professionnel pourrait aider

à comprendre les besoins de ces derniers lors de l'indexation des images. Ceci pourrait permettre de faciliter la recherche d'images.

Pour terminer, les changements relativement au Web2.0 apportent de nouveaux moyens de repenser l'indexation, par exemple la folksonomie.² Mais avant même de penser à instaurer de nouveaux schémas de métadonnées ou bien de changer les moteurs de recherche, ne serait-il pas important de se pencher sur l'utilisation des schémas déjà mis en place par les professionnels dans le but de les améliorer dans le futur? Qu'en est-il de l'utilisation des éléments des schémas de métadonnées, par exemple, VRA Core (Virtual Resources Association, 2014) et Dublin Core (Dublin Core Metadata Initiative, 2013) par les utilisateurs expérimentés en repérage d'images? Ces professionnels jouent un rôle de médiateur entre l'image et l'utilisateur et il serait important de comprendre comment les professionnels utilisent les métadonnées pour répondre aux demandes de référence pour une image.

1.2 Objectif de la recherche

Cette thèse a pour but de découvrir l'usage que font les intermédiaires des métadonnées associées aux images numériques lorsqu'ils en font la recherche dans le but de combler un besoin précis. Dans cette recherche, la population cible est les chercheurs expérimentés, comme les bibliothécaires, archivistes ou spécialistes de l'image qui font de la recherche d'images dans le cadre de leur travail, par exemple, un bibliothécaire recherchant une image pour un utilisateur. Nous les nommerons les professionnels. Ces professionnels

² Définition Folksonomie « Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information. » (Vander Wal, 2007)

travaillent dans le milieu de l'art ou de l'histoire, ils travaillent avec des images provenant d'objet d'art ou étant elles-mêmes une image d'art, des images d'artéfacts ou mêmes tout autres types d'images. Ils possèdent une connaissance de l'image, parfois celle-ci va de la connaissance de l'œuvre exacte à une connaissance assez vague.

Ces derniers effectuent dans le cadre de leurs tâches au travail un service de référence. La recherche d'images que les professionnels effectuent n'est pas aléatoire, elle remplit certains besoins d'information, car ils reçoivent une demande d'image ou de renseignements sur une image. La requête provient d'un usager du service de référence, qui peut être un étudiant ou un usager d'une bibliothèque ou d'un musée; nous le nommerons l'usager. Dans le dessein de combler ce besoin, les professionnels recherchent des images numériques, par conséquent, ils entreprennent quatre tâches : trouver, identifier, sélectionner, obtenir (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012). Cette recherche explorera comment les métadonnées associées aux images interviennent dans l'accomplissement de ces tâches faites par les professionnels.

Par l'observation de l'usage des métadonnées que font les professionnels qui effectuent des recherches pour les usagers des bibliothèques et musées, il sera possible de remarquer l'utilité des métadonnées dans les recherches d'images numériques relatives aux tâches effectuées. Comment interprètent-ils la demande reçue des usagers pour repérer l'image nécessaire? Quels éléments des métadonnées utilisent-ils (ou non) pour trouver, identifier, sélectionner, obtenir l'image? L'utilisation des métadonnées que font ces professionnels sera au centre de l'étude. La recherche apportera des contributions autant dans le milieu des bibliothèques que dans les musées. Elle fournira de nouvelles données sur l'utilisation des métadonnées par un type d'usagers précis, ceux qui ont une connaissance dans le domaine des images, ou du moins qui cherchent les images dans le but de combler un

besoin précis. Plusieurs études ont été faites au sujet des métadonnées entourant les images, mais pas en interrogeant ces types d'utilisateurs. Ces données seront utiles dans le futur, car l'indexation est en constante évolution; les métadonnées changent et de nouveaux schémas et de nouvelles techniques apparaissent (par exemple l'indexation automatique).

De plus, cette recherche apportera un regard nouveau sur l'utilisation que font les utilisateurs des informations qu'ils trouvent dans les métadonnées. Par exemple, les données révéleront quels éléments accompagnant l'image les professionnels estiment les plus utiles et quels éléments, selon eux, pourraient améliorer le repérage. Ceci permettra peut-être de repenser l'indexation des images et l'utilisation de certains schémas plutôt que d'autres dans le but de rendre les images le plus accessibles possible.

1.3 Questions de recherche

Cette recherche a pour but de combler un manque dans la littérature en ce qui concerne les métadonnées des images numériques et les professionnels qui font de la médiation entre l'utilisateur et l'image. Dans cette perspective, cette recherche tentera de répondre à ces questions de recherche :

1. Comment les professionnels interprètent-ils l'information des demandes des utilisateurs en métadonnées lors de la recherche d'images numériques?
2. Comment les métadonnées influencent-elles leur sélection d'images numériques?

1.4 Méthodologie

Des paramètres doivent être fixés pour répondre aux questions de recherche. La recherche a examiné comment les métadonnées influencent les tâches faites par les professionnels lors de la recherche d'images. Premièrement, les participants sont des

bibliothécaires ou spécialistes de l'image qui travaillent dans des musées ou des universités. Deuxièmement, les tâches qui ont été observées sont celles établies dans les *Fonctionnalités requises des notices bibliographiques* (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012), donc *trouver, identifier, sélectionner, obtenir*. Finalement, les métadonnées analysées sont celles qui sont liées aux images numériques.

Dans le but de bien cibler les données qui peuvent répondre correctement aux questions de recherche, trois modes de collecte de données ont été utilisés. Dans le dessein de récolter des données, les participants ont rempli un journal de bord (Appendice A) à chaque recherche d'image. Les participants devaient remplir un maximum de dix journaux. Ces formulaires ont été utiles lors des entrevues. Une entrevue semi-structurée (Appendice B) a permis de faire le point sur les rapports de recherche. Pour terminer, les participants devaient faire une recherche d'image numérique pendant laquelle une observation a été faite. Les participants ont reçu la même demande de recherche d'une image provenant d'un usager; ils ont accompli le repérage dans leur milieu de travail. Ils devaient également remplir le même journal de bord qu'ils ont complété pour les autres recherches.

1.5 Étendue et limites de la recherche

Cette recherche de type exploratoire à raisonnement inductif s'inscrit dans un processus de recherche pour l'obtention d'un diplôme de maîtrise en Sciences de l'information. Puisque seulement quatre participants ont répondu au questionnaire, il est certain que cette recherche ne sera pas représentative de tous les bibliothécaires et les chercheurs qui font de la recherche d'images numériques, mais elle met en lumière une piste de recherche qui peut porter à un développement sur la question de l'utilisation des métadonnées par les chercheurs expérimentés.

1.6 Plan

Le deuxième chapitre de la thèse comportera la recension des écrits provenant d'études sur le repérage d'images, les aspects sémantiques et syntactiques de l'image, le comportement de l'utilisateur, la médiation, les schémas de métadonnées et l'analyse des requêtes. Par la suite, le cadre conceptuel sera élaboré dans le but d'instaurer des balises qui délimiteront l'étendue de notre analyse. Cette section clôturera le deuxième chapitre.

Ensuite, la méthodologie sera expliquée dans le troisième chapitre qui se divisera comme suit : la revue d'éthique, l'échantillonnage et le recrutement, la rencontre d'information, les trois modes de collecte de données, leur conservation et leur analyse, la validation et les limitations de la recherche. Puis, le chapitre quatre comportera les résultats de la collecte de données qui seront divisés selon le mode de collecte : les journaux de bord, l'entrevue et l'observation. Le chapitre cinq permettra une discussion sur les résultats de la recherche selon ces thèmes : le contexte, l'analyse de la requête et l'utilisation des métadonnées. Enfin, cette thèse se terminera par une conclusion qui colligera les informations et les données. C'est dans cette section que nous proposerons également des pistes pour des recherches futures.

Chapitre 2 Revue de la littérature et cadre conceptuel

Metadata is much more than data about data; it expresses the « aboutness » of the data and all the forms in which « aboutness » can be explained and accessed (Zentner-Raasch, 2010, p. 3).

Dans ce chapitre, ce sont les bases sur lesquelles cette recherche s'appuie qui seront illustrées. Celui-ci est divisé en deux grandes parties : la revue de la littérature et le cadre théorique. La revue de la littérature abordera cinq sujets : la sémantique et la syntactique de l'image, le comportement de l'utilisateur, la médiation des professionnels, les schémas de métadonnées et l'analyse des requêtes. Pour terminer le chapitre, le cadre théorique sera exposé et il repose sur deux écrits : « A Conceptual Framework and Empirical Research for Classifying Visual Descriptors » selon Jörgensen, Jaimes, Benitez et Chang (2001) et les « Fonctionnalités requises des notices bibliographiques » (FRNB) (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012). Ce cadre conceptuel est fondamental puisqu'il est également un pilier dans l'élaboration de la méthodologie, que ce soit dans la structure du journal de bord ou dans l'analyse des données obtenues pendant la collecte.

2.1 La recherche sur le repérage d'images

Dans cette section, nous relaterons les écrits qui ont marqué la littérature et qui discutent de notre sujet d'étude. Depuis plus de vingt ans, plusieurs recherches ont porté sur le repérage d'images numériques, et ce, dans plusieurs domaines, tout en abordant différents points de vue que nous développerons ici. Il est possible d'expliquer cet engouement pour le thème par le fait que les projets de numérisation d'images dans les centres de documentation se réalisent rapidement par l'influence de la nouvelle technologie qui ne cesse d'évoluer. D'ailleurs, les avancements dans les pratiques du traitement des images font que le sujet est en constante évolution.

2.1.1 La sémantique et la syntactique de l'image

Plusieurs études mettent l'accent sur la théorie de l'image en elle-même. Les recherches se développent autour de l'image et son contenu, et ce, selon deux aspects : la sémantique et la syntactique de l'image.

Tout d'abord, Shatford (1986) analyse comment indexer les images selon leurs utilisations et elle illustre l'ampleur de leurs usages, mais aussi de leurs perceptions. Elle fait la démonstration de tous les usagers qui peuvent utiliser les images et c'est à partir de cela qu'elle met en évidence l'importance de la signification de ces dernières. L'iconographie fait partie intégrante de sa recherche, car elle incorpore le « about » et le « of » d'une image. Selon elle, il est possible de décrire une image en affirmant : c'est l'image *de quelque chose* et c'est une image *à propos de quelque chose*. L'auteure ajoute que pour indexer les images, il faudrait satisfaire quatre aspects : *qui, quoi, quand* et *ou*. Ces aspects doivent être remplis au spécifique ou au générique du *de* et du *à propos*. C'est sa classification à facette qui serait la base de l'indexation des images.

Dans une autre étude, Shatford Layne (1994b) élabore un guide de référence pour l'indexation des images. Il est question d'utiliser des attributs de l'image pour l'indexation dans le but de les rendre accessibles. Ces caractéristiques sont biographiques, illustratives, de sujet et de relation. Dans la catégorie biographique, ce sont les informations des créateurs du lieu et du temps que l'on y trouve. En ce qui concerne les attributs du sujet d'une image, Shatford Layne utilise encore le *à propos* et le *de* d'une image. Tout comme dans son autre article (Shatford, 1986), elle parle également du générique et du spécifique. Enfin, ces sujets peuvent être aussi classifiés selon le temps, l'espace, l'activité ou l'évènement. Dans la catégorie illustrative, les informations qui y sont contenues sont de type technique, matériel, par exemple, une photographie. Et pour terminer, les informations de relation, ce sont les relations entre l'image et d'autres œuvres. Cette catégorie sert à regrouper des images dans le but de mieux les retrouver.

Toujours en ce qui concerne les images et le repérage, Fidel (1997) instaure la théorie des pôles où elle décrit l'image qui peut être employée selon deux qualifications. Tout d'abord, il y a le pôle de l'image comme *objet*, ce qui veut dire que l'image est recherchée pour ce l'on y perçoit sur elle. Par exemple, une agence de presse qui a besoin d'une image d'une personne précise. Par la suite, il y a le pôle de l'image comme *donnée*, ici ce sont les données qui sont sur les images qui sont nécessaires, par exemple, une carte d'une ville pour démontrer les anciens quartiers. Fidel explique également que dans le milieu de ces deux pôles se retrouvent aussi d'autres types d'images. Enfin, ajoutons aussi que l'auteur fait une nuance importante, à savoir que la même image peut servir pour divers buts pour deux personnes différentes. Enfin, selon le besoin de l'image, selon le pôle où elle se situe, la recherche de l'image se fera autrement.

Dans le même ordre d'idées sur le repérage d'image, Jørgensen, Jaimes, Benitez et Chang (2001) tentent de catégoriser les types d'attributs qui peuvent être attribués à des images numériques lorsqu'elles sont regardées. Ces attributs sont séparés selon plusieurs catégories qui se trouvent dans deux sphères : sémantique et syntactique. Les auteurs voulaient pouvoir élaborer des classifications qui pourraient englober toutes les descriptions d'images possibles. Cette étude est discutée plus en profondeur dans le cadre conceptuel puisqu'il est un pilier de notre recherche.

Pour terminer, Enser (2008) illustre l'étendue des études qui ont porté sur le sujet en abordant quelques problématiques qui touchent directement notre recherche. Par exemple, il met en lumière le travail important de médiation fait par le professionnel en faisant référence à une télé-série, « Shooting the past » (Poliakoff, 1999), dans laquelle les protagonistes narrent des histoires grâce aux images contenues dans leur collection. De plus, Enser discute également de la problématique de l'indexation des images, des besoins de l'utilisateur, de recherche d'images par le contenu (content-based image retrieval, CBIR), le repérage d'images sémantiques.

Dans un autre article, les auteurs (P. G. B. Enser, Sandom, Hare, & Lewis, 2007) font un recensement de la littérature sur le trou sémantique lors de l'indexation dans le but de démontrer le manque toujours présent dans les techniques. Pour terminer, toutes ces recherches illustrées plus haut démontrent que la perception de l'image est toujours et autant étudiée.

2.1.2 Le comportement de l'utilisateur

Avec la popularité grandissante des recherches réalisées à propos des images et de leurs repérages, conséquemment, des études se consacrent sur les usagers et leurs démarches de recherche d'images.

Jørgensen (1998) présente une recherche dans laquelle elle recrute 107 participants, des étudiants en première année à la maîtrise en science de l'information. Elle a étudié les mots utilisés pour décrire une image par les participants. Ces mots étaient ensuite classés selon une liste d'attributs qui eux étaient aussi classés selon 12 catégories. Elle affirme « This research demonstrates the need to address a variety of image attributes, including 'interpretive' as well as perceptual attributes, in order to facilitate retrieval within a potentially broad range of image descriptions as seem to be typically generated by humans when asked to describe images. »(Jørgensen, 1998, p. 172). Jørgensen note ici l'importance de son étude puisqu'elle apporte une connaissance sur la description faite par des « non-spécialistes ».

Dans la même voie, Matusiak (2006) utilise la perspective des usagers pour évaluer une collection d'images numériques constituées d'images et de plans typographiques créés par l'Université de Wisconsin-Milwaukee. L'auteure veut connaître la satisfaction des usagers qui utilisent cette collection. Dans cette étude, la population étudiée est plutôt large, il s'agit d'étudiants, mais aussi de membres du public. L'auteure arrive à la conclusion qu'il y a une différence dans la façon de chercher entre les deux groupes. Par exemple, les étudiants font plus une recherche par mots clés, tandis que les personnes du public utilisent plus la navigation comme méthode de recherche. De plus, elle constate aussi que les façons dont les usagers recherchent ont évolué pendant la période que l'étude a eu lieu.

L'étude du comportement chez les usagers a été également apportée jusqu'à un autre type d'utilisateur. Neal (2008) tente de déterminer la préférence des métadonnées chez un groupe de jeunes photographes et bibliothécaires. C'est grâce à un questionnaire que les participants ont rempli et à une entrevue que cette recherche a pu démontrer les préférences des informations sélectionnées lors du repérage de photographies. Il a été conclu que les informations favorites pour faire la recherche d'images étaient le nom des objets et des évènements, les termes assignés par les usagers et la date/grandeur/photographe. Les participants avaient également comme stratégie de recherche lors du repérage le furetage. Pour terminer, cette recherche a aussi démontré que les photographes étaient également conscients de l'importance de la qualité des métadonnées.

Une autre étude s'est penchée sur d'autres genres d'usagers : les archéologues, les architectes, les historiens de l'art et les artistes (Beaudoin & Brady, 2011). Cette fois-ci, ce sont les comportements des usagers qui ont été analysés dans le but de remarquer quelles ressources ils utilisent pour faire leurs recherches d'images. Les auteurs ont démontré que contrairement aux autres groupes étudiés, les artistes préfèrent l'imprimé au numérique pour l'obtention de leurs images. De plus, lorsque les participants devaient répondre quelle était leur ressource préférée, celle qui a reçu le plus de réponses est *Google Images*.

Enfin, toutes ses recherches offrent un portrait de divers types d'usagers et de plusieurs genres de comportements, ce qui situe notre recherche dans la littérature présente. Ces articles démontrent que selon le type d'usagers, les préférences lors du repérage d'images peuvent varier et qu'il importe de se pencher sur ces prédilections dans le but d'offrir des outils de recherche plus performants centrés sur le type d'utilisateurs ciblés.

2.1.3 La médiation par les professionnels

Plusieurs études se sont également penchées sur l'aspect d'un médiateur entre l'objet recherché et le repérage lui-même. Ces recherches étudient comment les professionnels qui jouent un rôle de médiateur font le repérage d'images et quelles sont leurs responsabilités précisément.

Shatford Layne (1994a) tente de comprendre le rôle du bibliothécaire lors de la recherche d'images par certains groupes, c'est-à-dire artistes et historiens de l'art. Elle affirme que les bibliothécaires doivent comprendre les besoins des usagers pour pouvoir ensuite juger et éventuellement améliorer les bases de données dans lesquelles les usagers cherchent. Selon elle, plusieurs caractéristiques de l'image peuvent rendre la recherche ardue : le nom de l'artiste, le titre de l'œuvre, la date et le lieu. Les ambiguïtés dans ces points d'accès peuvent rendre la tâche de la recherche laborieuse et les bibliothécaires se doivent de comprendre ces problèmes selon le type précis d'usagers.

De se questionner sur la fonction des bibliothécaires de référence ne faisait pas seulement couler l'encre dans les années 90. Le sujet devient de plus en plus d'actualité, car Goodrum (2003) parle également d'un projet permettant une collaboration entre un musée et une bibliothèque pour offrir un service de référence en ligne axé sur la recherche d'images. Le projet s'est fait à partir d'un modèle déjà créé, le Virtual Reference Desk. Goodrum (2005) s'est questionnée en plus sur la médiation plus dans un autre article qui tente de répondre à savoir comment l'expertise d'un humain médie entre les requêtes des usagers et les descriptions d'images sur le Web (p.47). La recherche s'est déroulée grâce à deux services de référence virtuelle : Le Virtual Reference Desk et AskERIC archives. Goodrum affirme :

This study confirmed earlier studies that demonstrated a gap between image index terms derived for retrieval purposes and the terms that users employ when describing images needed. When users request an image, they tend to describe the image itself rather than the meaning or emotive content of the image; their indexing vocabulary largely reflects what they expect to see in the image. (2005, p. 53)

Ces études sont le début de la recherche sur l'importance du rôle du professionnel de l'information et cette recherche vient en s'interrogeant sur certains aspects, comme les schémas de métadonnées, les services offerts et le rôle du professionnel.

2.1.4 Schémas de métadonnées

Il est indéniable que l'analyse du repérage d'images se doit de s'attarder sur les schémas de métadonnées puisqu'ils sont le centre de la classification des images, par conséquent un niveau important dans la recherche.

Greenberg (2001) a analysé quatre schémas de métadonnées qui peuvent être utilisés pour la recherche d'images : VRA Core, Dublin Core, REACH et EAD. Grâce à une méthodologie quantitative, elle analyse les schémas selon la distribution de quatre différentes classes de données (découverte, utilisation, authentification et administration) et tente de trouver la différence de granularité des schémas. En ce qui concerne les métadonnées qui supportent la découverte, Dublin Core et REACH sont les deux qui ont le plus d'éléments de ce type. Pour ce qui est de l'utilisation, c'est EAD qui a le plus d'éléments qui soutiennent cette fonction. En ce qui concerne l'authentification, Dublin Core et REACH sont les schémas qui ont le plus d'éléments qui supportent ces fonctions. Cependant, encore une fois, les quatre schémas n'ont pas une grande différence dans les totaux. En ce qui concerne l'administration, EAD et VRA Core sont les deux schémas qui supportent l'administration.

Pour terminer, cette étude démontre que chaque schéma de métadonnées est créé dans des desseins différents.

Une autre étude se penche sur les schémas de métadonnées tout en analysant le comportement des usagers (Attig, Copeland, & Pelikan, 2004). Dans une bibliothèque d'images universitaires, un schéma de métadonnées a été créé dans le but de combler les besoins du personnel et des étudiants lors du repérage. Ce dernier est né de trois normes : VRA Core, Dublin Core et IMS Learning Resource Meta-data. Trois étapes à la recherche ont permis d'arriver à un produit. Tout d'abord, les besoins des usagers ont été explorés, ensuite, deux prototypes de système de gestion d'images ont été créés et, pour terminer, ces systèmes ont été évalués par les usagers. Les auteurs sont venus à se poser la question à savoir quel était le minimum de métadonnées nécessaire pour le repérage d'image. Comme dans ce contexte de recherche les images provenaient de plusieurs sources, ils affirment que le système de métadonnées doit être assez complexe pour tous les incorporer.

D'ailleurs, en ce qui concerne les études sur les schémas de métadonnées, Fear (2010) mène une étude chez les étudiants de premiers cycles, c'est-à-dire, comme elle les nomme, les usagers non expérimentés. Elle veut savoir quelles métadonnées du schéma de Dublin Core les étudiants trouvent utiles et si certains éléments qui pourraient l'être ne sont pas disponibles. Grâce à des questionnaires, des groupes de discussion et une recherche assistée, la chercheuse a réussi à analyser l'utilité des métadonnées. Tout d'abord, en ce qui concerne les réponses provenant de la section du questionnaire où les participants pouvaient répondre librement, les informations qui ont été retenues comme les plus utiles sont celles que l'on peut récupérer en regardant l'image elle-même, par exemple, le contenu et l'impression que l'image donne. Enfin, l'autre catégorie qui a retenu le plus d'attention est celle qui est relative à l'utilisation et dans cette catégorie, c'est la grandeur de l'image qui a

été le plus souvent sélectionnée. Par la suite, en ce qui concerne les éléments de Dublin Core utilisés par les étudiants, dans ce cas-ci, le sujet, la description, le type et le format se sont démarqués.

Enfin, ces recherches démontrent que l'indexation des images et la sélection du schéma de métadonnées sont importantes, mais varient aussi selon le type d'utilisateur et le type d'images contenues dans la base de données.

2.1.5 L'analyse de la requête

Les analyses de repérage d'images ont également porté sur un autre sujet, l'analyse des requêtes. Remarquer comment les usagers lancent les requêtes peut également en dire long sur la façon dont ils font le repérage d'images.

Enser et McGregor (1992) étudient les requêtes des usagers reçues par une bibliothèque d'images dans le but de les catégoriser, de regarder l'efficacité du schéma de métadonnées utilisé et, pour finir, de s'attarder à la corrélation entre les images catégorisées, le nombre d'images retrouvées et l'usage fait par les usagers. Les auteurs affirment à propos de la spécificité de la collection :

The findings of the present project lend to support to the view that the subject indexing for commercial photographic material is of low utility; the judicious assignment of indexing terms to the images held within a subject-specific collection which services the more predictable requests of a specialist client base may be of considerably greater utility (P. G. B. Enser & McGregor, 1992, p. 26).

Toujours dans la même décennie, Hasting (1994) fait une dissertation sur l'analyse de la recherche d'images par les historiens de l'art. Sa recherche veut répondre à la question suivante : « Do the art historian's queries have a relationship to the access point used by

them for retrieval of the needed image(s)? » (p.8). À partir d'une collection de University of Central Florida, les requêtes pour des photographies et des images numérisées ont été analysées. Elle s'interroge pour savoir qu'est-ce qui change lors de la recherche d'images lorsque le format n'est plus le même, c'est-à-dire numérique ou papier. Enfin, elle conclut que le niveau de complexité lors d'une requête d'image numérique augmente comparativement à une recherche d'image papier.

D'autres études se sont intéressées à l'image, pas seulement la fixe, mais aussi la mouvante. Armitage et Enser (1997) ont conduit une analyse sur les requêtes des usagers dans sept bibliothèques. Les requêtes ont été séparées entre quatre catégories : contenu d'images, identification, accessibilité et divers. Les auteurs annoncent qu'ils ont trouvé une ressemblance entre les différentes bibliothèques dans lesquelles cette étude a eu lieu, même si les usagers eux se différencient les uns des autres puisqu'ils consultent des bibliothèques différentes. Les établissements ont été regroupés selon le type de ressources que l'on y retrouvait et, selon ces regroupements, il est possible d'affirmer que les requêtes se ressemblent selon le type d'établissement.

Toujours dans le domaine de l'histoire de l'art, Chen (2001) étudie les requêtes d'étudiants de cette discipline. Il met côte à côte ces résultats lors de sa recherche et ceux exposés dans des études antérieures : celles de Enser et McGregor (1992), Jörgensen (1995) et Fidel (1997). La conclusion de cette recherche a soulevé une question importante à propos de la différence des termes utilisés par les usagers contrairement à ceux utilisés pour l'indexation. Les participants ont utilisé des termes uniques pour lancer la requête et ils ont ajouté quelques facettes, par exemple une date. Cette recherche a démontré que les catégories de Enser et McGregor étaient les plus efficaces dans le cadre de cette recherche.

Chez un autre type de participants, Choi et Rasmussen (2003) étendent leur recherche chez les étudiants et les membres d'une faculté d'études supérieures qui font de la recherche pour combler leurs besoins d'images d'histoire de l'Amérique. Tout d'abord, les requêtes des usagers ont été analysées en terme du nombre de mots utilisés et ensuite ces requêtes ont été comparées entre types d'utilisateur : les membres de la faculté et les étudiants. Une moyenne de 4,87 termes a été utilisée pour les recherches et, à partir de ces termes, les auteurs ont démontré que près de 65 % étaient pour un nom spécifique d'un évènement, une action ou une place. Par ailleurs, cette recherche démontre qu'aucune recherche pour une émotion ou un concept abstrait n'a été faite.

Avec la disponibilité des services Internet fort répandue, certaines études mettent en lumière les requêtes des usagers avec l'utilisation des moteurs de recherche. Dans une recherche analysant les demandes sur Excite (Goodrum & Spink, 2001), un moteur de recherche qui s'est prévalu d'un thésaurus, les auteurs ont remarqué certaines particularités dans les requêtes d'images. Tout d'abord, 59,6 % des requêtes lancées étaient modifiées et les autres étaient uniques, donc une seule recherche. De plus, comme la recherche portait sur le Web, sur les 100 mots les plus populaires, 25 % des mots étaient reliés à la sexualité, voire à la pornographie. Enfin, dans les termes qui revenaient le plus souvent, puisque c'est un moteur de recherche, il est certain que « pictures » et « Pics », étaient des termes très populaires. En somme, la moyenne de mots utilisés par requête est de 3,74.

Avec l'omniprésence des ressources sur le Web, Chung et Yoon (2011) conduisent une étude sur le service de *Yahoo Answer* qui est un service de référence en ligne. Cette recherche constate que les questions reçues pour la recherche d'images contenaient en moyenne 55,7 mots qui décrivaient l'image recherchée. De plus, de ces termes, les requêtes ont été classées tout d'abord selon le pôle dans lequel elles se trouvaient. À la suite de cette

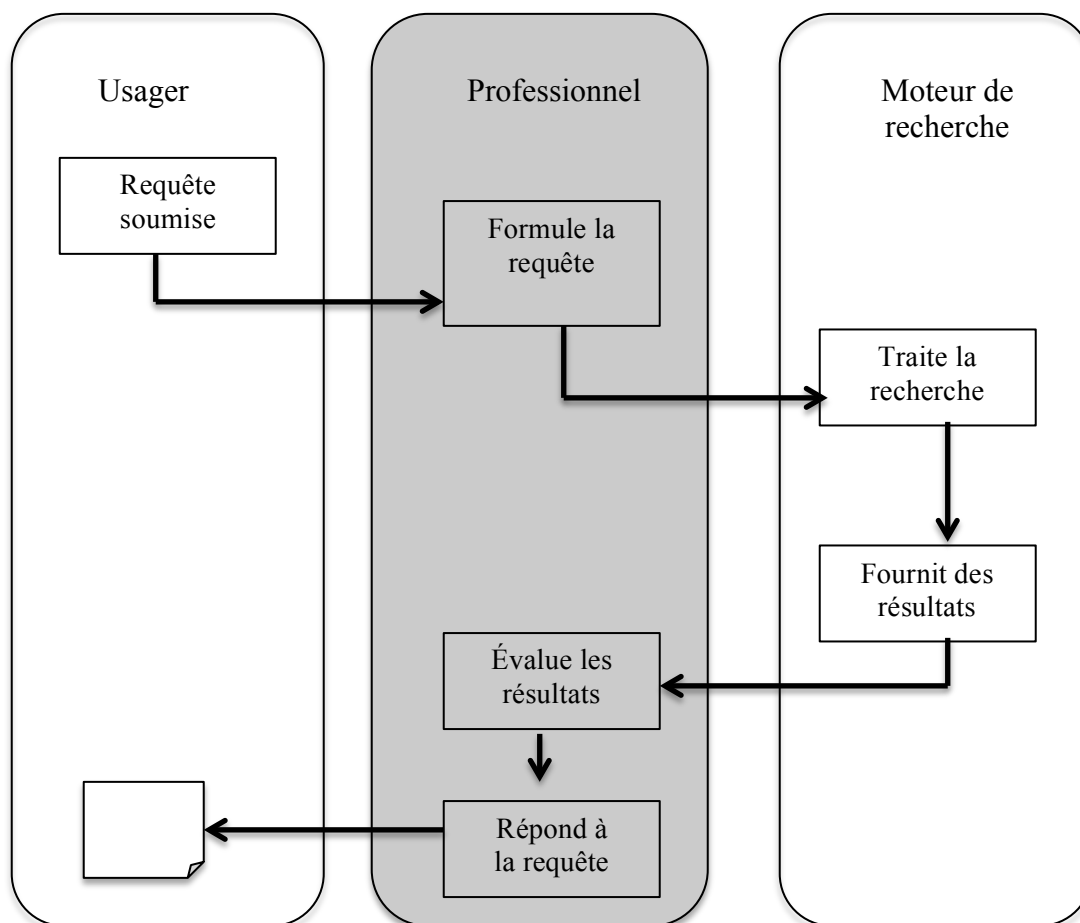
classification, les demandes ont été reclassées selon les attributs. Il est possible d'affirmer que cette recherche indique que la plupart des termes utilisés par les usagers se retrouvent dans la sphère sémantique et qu'ils sont pour la plupart de types génériques, plutôt que spécifiques. Enfin, les auteurs ont aussi démontré que les demandes étaient très variables, sur 24 catégories disponibles sur le site, les requêtes en couvraient 22.

2.2 Cadre conceptuel

Dans le but de bien établir les bases de cette analyse et de pouvoir exploiter les données selon des directives, la recherche doit s'appuyer sur un cadre conceptuel. « Un cadre conceptuel décrit, sous forme graphique ou narrative, les principales dimensions à étudier, facteurs clés ou variables clés, et les relations présumées entre elles » (Miles & Huberman, 2003, p. 41). Dans ce cas-ci, deux ressources serviront d'assise pour la recherche : « Fonctionnalités requises des notices bibliographiques » (FRNB) (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012) et « A conceptual framework and empirical research for classifying visual descriptors » (Jørgensen et al., 2001).

La figure 2.1 illustre les étapes du repérage d'images selon ce que nous avons observé pendant notre recherche. Nous nous attarderons sur la zone en gris au centre de la figure, car c'est dans cet espace que notre analyse a pris forme et c'est également là où la collecte de données a eu lieu. Notre cadre conceptuel permettra d'analyser cette zone.

Figure 2.1 Cadre de recherche



Tout d'abord, les usagers ont un besoin d'images ou d'informations sur une image. Pour débiter, les usagers vont se créer une idée personnelle de l'image pour ensuite la traduire en mots (Greisdorf & O'Connor, 2002, p. 8) qu'ils vont transmettre en requête aux professionnels. Par la suite, le professionnel fait un travail d'interprétation, c'est-à-dire que selon ses connaissances, il prend l'information qu'il juge utile et il lance la requête. Lors du retour de résultats, le professionnel a encore un travail à effectuer. Toutes ces étapes ont été analysées par notre cadre théorique. Premièrement, ce processus de repérage a été construit et divisé en étapes selon les quatre tâches identifiées dans les FRNB (2012). Les FRNB ont

été choisis parce qu'elles permettent une bonne exploration des tâches accomplies dans le repérage d'information et que ces quatre tâches peuvent être par la suite reliées avec les métadonnées supportant chacune d'elles. C'est de plus grâce à ces tâches que nous avons divisé les résultats de la recherche. Par la suite, les métadonnées utilisées ont été analysées selon les niveaux des attributs du contenu visuel définis dans le cadre conceptuel élaboré par Jørgensen, Jaimes, Benitez et Chang (2001).

2.2.1 Fonctionnalités requises des notices bibliographiques (FRNB)

Dans l'élaboration des journaux de bord et l'analyse des données, nous avons utilisé les « Fonctionnalités requises des notices bibliographiques » (FRNB) (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012). Les FRNB sont un modèle conceptuel utilisé dans le catalogage en bibliothèque (Carlyle, 2006); ce modèle ne se résume pas à être élaboré que pour les bibliothèques, mais il convient partout où les notices bibliographiques peuvent être utilisées. En plus de vouloir couvrir le plus de types d'utilisateurs, il veut aussi pouvoir être utilisé pour plusieurs types de documents et de supports. Ce modèle s'adapte très bien à notre type de recherche. « Le modèle élaboré pour la présente étude constitue une amorce de cadre conceptuel visant à aider à la compréhension et à l'élaboration de futures conventions de description bibliographique » (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012, p. 9).

Ce modèle complexe comprend plusieurs entités divisées selon trois groupes. « Les entités du premier groupe représentent les différents aspects de ce qu'un utilisateur peut trouver dans les produits d'une activité intellectuelle ou artistique » (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012, p. 14), c'est-à-dire œuvre, expression, manifestation, item. Dans le deuxième groupe, on retrouve deux entités :

personne et collectivité. Pour terminer, dans le troisième groupe se trouvent le concept, objet, évènement, lieu. Ces entités ont également des attributs, « c'est au moyen des attributs de l'entité que les utilisateurs formulent leurs requêtes et interprètent les réponses qu'ils reçoivent lorsqu'ils effectuent une recherche sur une entité déterminée. » (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012, p. 26). S'ajoutent à ces fonctions les opérations qui peuvent être faites par les utilisateurs. Dans le cas de cette recherche, ce sont ces opérations, que nous appellerons tâches, qui nous intéressent.

Ces quatre opérations ont permis de construire les journaux de bord. Il est important de bien les définir puisque ces tâches sont un pilier du journal de bord et par le fait même de la collecte de données. Voici comment ils y sont définis :

- trouver les entités qui correspondent aux critères de recherche formulés par l'utilisateur (c'est-à-dire, localiser soit une entité isolée soit un ensemble d'entités dans un fichier ou dans une base de données au terme d'une requête portant sur un attribut ou une relation de l'entité);
- identifier une entité (c'est-à-dire, obtenir confirmation que l'entité décrite correspond bien à l'entité recherchée, ou bien établir une discrimination entre deux ou plusieurs entités présentant des caractéristiques analogues);
- sélectionner une entité qui soit appropriée aux besoins de l'utilisateur (c'est-à-dire, choisir une entité qui corresponde aux spécifications émises par l'utilisateur en ce qui concerne le contenu, la présentation matérielle, etc., ou bien écarter une entité qui ne serait pas appropriée aux besoins de l'utilisateur);
- obtenir (gratuitement ou moyennant paiement) l'accès à l'entité décrite (c'est-à-dire, obtenir une entité par achat, prêt, etc., ou bien accéder numériquement à une

entité par le biais d'une connexion en ligne à un ordinateur distant) (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012, p. 59).

Enfin, ce modèle complexe nous a permis de structurer la recherche selon les tâches, ce qui permet une analyse en étape de chaque période du repérage, c'est-à-dire une vision complète des sphères de la recherche d'image.

2.2.2 Pyramide

Un autre cadre conceptuel qui a été utilisé pendant cette étude est celui de Jørgensen, Jaimes, Benitez et Chang (2001) qui s'applique au matériel visuel que l'on a nommé la pyramide à cause de sa forme.

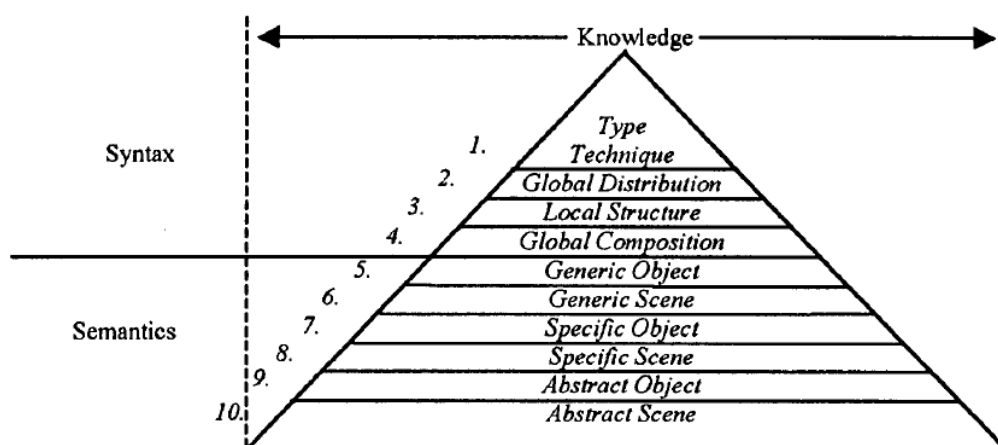
Cette pyramide contient dix niveaux qui permettent de classer les attributs du contenu visuel. Selon les auteurs, « visual content carries information at many different levels (e.g., syntactic: the colors in the image; semantic: the objects in the image) » (2001, p.939). De plus, ils affirment que les

descriptions of visual content can be highly subjective, varying both across indexers and users, and for a single user over time. Such descriptions depend on other factors that include, for example, the indexer's knowledge (e.g., art historian), purpose of the database (e.g., education), database content (e.g., fine art images; commercial images), and the task of the user (find a specific image or a "meaningful" image) (2001, p. 939).

C'est selon la division des niveaux que la pyramide s'achemine. Cette pyramide couvrirait tout le savoir et se divise selon dix niveaux (figure 2.2). Les quatre premiers niveaux font

partie des niveaux de la syntaxe, tandis que les six autres eux combrent les niveaux sémantiques.

Figure 2.2 La pyramide



Pour les niveaux sémantiques, un savoir est requis pour faire l'indexation, mais au niveau syntaxique, l'indexation ne nécessite pas une connaissance pointue (p.940). La longueur des niveaux de la pyramide reflète la quantité de connaissance à avoir pour pouvoir indexer à ce niveau. D'ailleurs, les niveaux génériques, spécifiques et abstraits ne sont pas définis, les auteurs préfèrent fournir des exemples qui reflètent chaque niveau, en ce sens ils affirment : « Although it is possible to mathematically define specific, generic, and abstract, such definitions would be difficult to apply. Instead, we present intuitive definitions with examples » (Jørgensen et al., 2001, p. 940). De plus, certaines métadonnées ne sont pas classées dans la pyramide, elles sont des informations extérieures, car elles ne peuvent pas être trouvées en regardant directement l'image, elles proviennent de l'externe, par exemple, la date. Cette pyramide a permis d'analyser les demandes des usagers lors de la demande d'images numériques et il a permis également de comparer quel niveau qui a été le plus

souvent utilisé par les professionnels dans l'accomplissement des quatre tâches définies plus haut. Le degré de connaissance mis en corrélation avec les métadonnées utilisées par les usagers et les professionnels permettra d'avoir un tableau clair de la façon que les professionnels utilisent les métadonnées en fonction des demandes qu'ils reçoivent des usagers.

2.3 Conclusion

Enfin, la recherche faite sur le repérage d'images a démontré l'ampleur des écrits sur le sujet. La sémantique et la syntactique de l'image ont permis d'analyser beaucoup de sujets en lien avec l'image. De plus, plusieurs études ont également été faites en lien avec le comportement des utilisateurs, mais aussi des médiateurs entre les usagers et l'image. Nous avons également démontré que certaines études ont été interpellées par les problèmes liés avec les schémas de métadonnées et l'analyse de la requête. De plus, la mise en place de notre cadre théorique permettra une analyse complète des données récoltées pendant la collecte.

Chapitre 3 Méthodologie

Cette étude est de type exploratoire et elle se concentre sur la recherche d'informations. Elle analyse les techniques de repérage d'images par les professionnels lors de la réception d'une demande en référence par un usager. Cette recherche n'évalue pas les systèmes de recherche d'informations, mais plutôt la façon dont les utilisateurs effectuent leur repérage d'images.

Dans le but de remplir le dessein de cette étude, les méthodes utilisées pour la collecte de données sont soit qualitatives ou quantitatives, elles se chevauchent. Plusieurs raisons peuvent pousser un chercheur à choisir cette approche, mais en ce qui concerne notre étude, cette approche a été sélectionnée, car une seule source de donnée était insuffisante (Creswell & Plano Clark, 2011, p. 8). Autant les données qualitatives que les données quantitatives enrichissent un portrait plus détaillé et juste de la réalité que nous observons. Il est important aussi de préciser que les données qualitatives et quantitatives n'étaient pas meilleures les unes que les autres. Même si pendant la collecte de données nous avons obtenu plus de données qualitatives (à cause du choix de type de collecte), une valeur équivalente a été apportée à tous les types de données.

De plus, le design de cette recherche est convergent, ce qui consiste à faire une collecte de données de types différents (textuelles et numériques) dans la même période de temps (Creswell & Plano Clark, 2011, p. 77) pour ensuite en faire l'analyse séparément. De

plus, il y a une démarche inductive qui « consiste à induire des énoncés généraux (des vérités) à partir d'expériences particulières, rigoureuses et systématiques » (Depelteau, 2000, p. 56). Donc, aucune hypothèse n'a été posée avant d'avoir étudié la population choisie et d'en avoir analysé les résultats obtenus lors de la collecte de données. De ce fait, pendant toute la démarche de recherche, la chercheuse a tenté d'être la plus objective possible.

3.1 Revue d'éthique

Cette recherche s'est fait conformément au code d'éthique en recherche à l'Université d'Ottawa à la suite de la réception de l'approbation du projet (Appendice C). Le projet d'éthique a été renouvelé jusqu'au 29 avril 2014. À cette date, la collecte de données était terminée. De plus, tous les actes de cette recherche ont été régis selon quatre principes moraux définis : consentement éclairé, sans tromperie, précision, respect de la vie privée et confidentialité (Christians, 2011, p. 65). Pendant la collecte de données, tous les actes entrepris par la chercheuse ont respecté ces règles.

3.2 Échantillonnage et recrutement

La population étudiée est un bassin de professionnels qui effectuent de la recherche d'images pour des usagers dans un cadre de référence soit dans des bibliothèques universitaires ou muséales. Certains sont également impliqués dans des projets de numérisation d'images dans le but de les rendre disponibles en ligne. Ces professionnels sont d'ailleurs des personnes ayant une connaissance technique des images, mais également du sujet entourant ces objets, donc soit une connaissance quelconque en art ou en histoire. Ils travaillent à rechercher les images dans le cadre de leurs fonctions, ils sont en quelque sorte des experts dans leur domaine, puisqu'ils sont désignés à cette tâche : le repérage d'images.

D'ailleurs, selon le domaine dans lequel ils œuvrent, certains sont appelés à faire de la formation à la recherche pour leur clientèle.

En ce qui concerne l'échantillonnage, il a été non probabiliste de volontaires, donc c'est sur une base de volontariat que les participants se sont proposés pour participer à l'étude. Une affiche a été créée au Congrès des milieux documentaires à Montréal en septembre 2012 dans le but de faire du recrutement. La chercheuse a eu l'opportunité de discuter du projet avec des professionnels et les personnes ont pu également s'informer sur le sujet grâce à l'affiche et à un dépliant. De plus, un courriel (Appendice D) a été envoyé aux professionnels susceptibles de faire du repérage d'images numériques dans la région de Montréal et d'Ottawa. Leurs adresses courriel ont été retrouvées sur les sites Web des établissements. Lorsque le nom du professionnel et ses informations n'apparaissaient pas, un courriel était envoyé tout simplement à l'adresse courriel de la référence. Dans certains cas, la personne ayant reçu le courriel l'a transféré à la personne responsable du repérage d'images numériques. Enfin, l'invitation à participer a aussi été publiée sur le site Web de Art Libraries Society of North America (ARLIS). Chaque invitation à participer au projet était accompagnée de la description de la recherche (Appendice E).

Les participants ont été choisis lorsqu'ils répondaient à certains critères. Ils devaient, dans le cadre de leurs tâches régulières au travail, faire du repérage d'images numériques. De plus, ils étaient tenus d'effectuer ce repérage dans un cadre de référence, où un usager demande une image en particulier et qu'ils agissent comme intermédiaires pour effectuer la tâche. Ils devaient travailler dans les deux milieux sélectionnés, c'est-à-dire université ou musée. Enfin, ils devaient également consentir à remplir les journaux de bord et à faire les rencontres.

3.3 Rencontre d'information

Lorsque les professionnels ont répondu au courriel d'invitation, une rencontre d'information a été planifiée dans le but d'expliquer plus en détail en quoi consistaient la recherche et également les prochaines rencontres à venir. Ces rencontres se sont faites soit par téléphone, soit en personne dans l'endroit désiré par les professionnels. C'est pendant cette rencontre que le journal de bord à compléter a été expliqué plus en profondeur dans le but d'uniformiser la façon de le remplir. Cette explication voulait aussi par le fait même permettre d'obtenir des résultats plus uniformes. Les rencontres avaient une durée indéterminée, le temps variait selon le nombre de questions posées par le professionnel. Dans un cas, la rencontre a duré plus de deux heures. Somme toute, certains professionnels rencontrés ont décidé de ne pas participer à la recherche, soit par manque de temps ou par manque de ressources. C'est à la fin de cette rencontre que les participants ont signé le formulaire de consentement (Appendice F) en deux exemplaires. Ils ont pu en conserver une et l'autre est conservée avec les données par la chercheuse.

3.4 Participants

Après avoir rencontré les professionnels pour l'explication du projet, quatre professionnels (n=4) ont décidé de participer à la recherche. Les participants proviennent de milieux similaires, c'est-à-dire qu'ils travaillent dans des bibliothèques universitaires (n=2) ou muséales (n=2). Bien que ces participants n'aient pas tous le même titre d'emploi, ils font tous les mêmes activités au travail et ils sont tous considérés comme les experts en repérage d'images numériques dans le milieu dans lequel ils travaillent. Ces participants ont fait le même processus, c'est-à-dire qu'ils ont complété les trois étapes de la collecte de données. Tout d'abord, ils ont rempli un nombre de journaux de bord relatif aux demandes d'utilisateurs

qu'ils ont reçues. La nature des demandes a varié en fonction du milieu dans lequel les professionnels travaillent et du type d'utilisateurs. Un maximum de dix journaux de bord a été demandé et un total de 24 journaux a été amassé (n=24) (tableau 1). Ces journaux de bord étaient remplis dans le moment qu'ils désiraient dans le cadre de leur travail. En deuxième lieu, les participants ont participé à une entrevue semi-dirigée dans un endroit de leur choix. Pour terminer, tous les participants se sont soumis à une observation par la chercheuse pendant laquelle ils recevaient tous la même requête provenant d'un utilisateur réel. Ils devaient faire la recherche d'image selon l'information qu'ils avaient reçue.

Tableau 3.1 Description des participants

Participants	Type d'institution	Nombre de questionnaires remplis
P1	Musée	6
P2	Université	3
P3	Musée	9
P4	Université	6
		Total : 24

3.5 Mode de collecte de données

Trois modes de collectes ont été employés dans le but de trianguler les données. En tout d'abord, les participants ont dû remplir un journal de bord lors de la recherche d'images. C'est par la suite qu'une entrevue semi-dirigée a eu lieu et pour terminer, les participants ont dû se soumettre à une observation d'une recherche d'image.

3.5.1 Journal de bord

La première étape de la collecte de données a été le journal de bord qui contenait des données qualitatives et quantitatives. Un journal de bord (Appendice A) a été rempli par les

participants à chaque recherche d'images qu'ils ont faite. Il leur a été demandé de compléter un maximum de dix journaux. Ces derniers ont été remplis dans un fichier Word qui a été remis à la chercheuse par courriel avant l'entrevue. Les participants avaient un mois pour remplir les journaux de bord. Chaque fichier a été renommé comme suit : le numéro du participant, puis *jdb*. et le numéro du journal, par exemple, P3-jdb2, qui correspond au deuxième journal de bord rempli par le troisième participant. En ce qui concerne la création du journal de bord, il a été divisé en quatre sections, selon les tâches de FRNB : *trouver*, *identifier*, *sélectionner* et *obtenir* (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012).

Dans la section *trouver*, les participants devaient retranscrire la requête qu'ils ont reçue de l'utilisateur. Ils ont dû également inscrire les mots clés qu'ils ont utilisés pour lancer la requête.

Ensuite, dans la section *identifier*, les participants devaient mentionner s'ils avaient obtenu des résultats qui étaient pertinents à la requête qu'ils avaient lancée. De plus, ils devaient copier l'URL de l'image sur laquelle ils cliquaient et indiquer si cette image était en lien avec la recherche, si non, ils devaient indiquer pourquoi.

C'est dans la section *sélectionner* que le participant devait inscrire l'utilité des métadonnées pour eux. Une liste de champs de métadonnées leur était fournie et ils devaient cocher la pertinence de chaque métadonnée pour cette recherche. Cette liste a été créée selon les éléments définis dans les schémas VRA CORE (Virtual Resources Association, 2014) et Dublin Core (Dublin Core Metadata Initiative, 2013). Dublin Core a été choisi, car c'est un schéma qui est utilisé autant dans le domaine des arts, de l'archivistique et des ressources visuelles (Greenberg, 2001, p. 920). VRA Core a été sélectionné, car il est un standard employé pour la description des œuvres d'art et les images des œuvres (Virtual Resources

Association, 2014). Cette section devait être remplie même si le participant n'a pas encore sélectionné d'image à ce stade, puisque nous voulions également savoir quel élément était utile pour discréditer l'image et recommencer la sélection.

Enfin, l'obtention de l'image est la dernière section du journal. C'est à ce moment que le participant indique comment il a obtenu l'image, par exemple s'il a dû numériser une image avec une qualité supérieure ou s'il a pu tout simplement envoyer l'image à l'utilisateur.

La complexité du journal de bord était nécessaire pour pouvoir analyser chaque tâche que le participant devait remplir. Cette division permet une analyse séparée des étapes et une comparaison des résultats dans les différentes tâches. Cette complexité rend le journal de bord un peu laborieux à remplir, ce qui a découragé certains potentiels participants à s'inscrire à notre étude.

3.5.2 Entrevue

Des entrevues semi-dirigées ont été conduites avec tous les participants. Une série de questions étaient déjà prédéterminées (Appendice B) dirigeant ainsi la conversation. Pendant cette entrevue, ce sont également les journaux de bord que les participants ont remplis qui ont alimenté la discussion. En fait, utiliser ce type d'entrevue permet une plus grande liberté dans les questions posées, ce qui permet de personnaliser la rencontre selon les réponses obtenues avec les journaux de bord. De plus, l'entrevue semi-dirigée a pour but de « permet [tre] de rendre explicite l'univers de l'autre [...], vise[r] la compréhension du monde de l'autre [...] permet d'apprendre, à propos du monde de l'autre et aux interlocuteurs, d'organiser, de structurer leur pensée [...], a une fonction émancipatrice » (Savoie-Zajc, 2000, p. 342). Par ailleurs, puisqu'aucune hypothèse n'a été émise sur l'utilisation des métadonnées, il était important de laisser place aux questions imprévues, de cette façon, la

discussion pouvait se diriger dans le sens qui sera important pour la recherche. Ces rencontres ont été enregistrées sur une piste audio à l'aide d'un iPad et elles ont été par la suite retranscrites dans le but d'analyser les rencontres. La retranscription des entrevues a été faite grâce au logiciel Word et chaque fichier a été nommé comme suit : le numéro du participant et *ent.*, par exemple P2-ent. Ces entrevues ont été conduites dans le milieu de travail des participants, de cette façon, ils avaient tous les outils qu'ils ont utilisés pour la recherche à la portée de la main. De plus, les entrevues ont duré de 20 à 40 minutes selon le nombre de journaux remplis, puisque dans certains cas, les participants avaient plus de précision sur leur recherche, mais aussi certains participants font plus de recherche d'images que d'autres dans le cadre de leur travail.

3.5.3 Observation et prise de note

Le troisième mode de collecte de données est une observation d'une recherche d'image par le professionnel. Une demande de recherche de type référence a été distribuée à chaque participant. Ils ont tous reçu la même requête juste avant l'observation :

Je recherche un tableau représentant, je crois, le Cap-à-l'original au Bic, Bas-Saint-Laurent, QC. J'ai vu ce tableau au musée il y a longtemps; on y peignait la silhouette du cap et il y avait un feu et des amérindiens. Je peux vous envoyer une photo de la silhouette du C.àO., silhouette très particulière avec un rocher qui se détache du cap et forme un pic.

Les participants devaient remplir un journal de bord pendant cette recherche. Pendant qu'ils effectuaient la recherche, ils devaient également penser à voix haute. La recherche d'image s'est déroulée dans l'espace de travail des participants, de cette façon, ils baignaient dans leur environnement de travail et il n'était pas distrait par le changement

d'outil ou de local. Les participants se sentaient tout de même observés puisque je regardais leur démarche. De plus, les recherches ont été enregistrées avec un iPad. Chaque recherche a été analysée avec celles des autres participants dans le but de corrélérer les données obtenues.

3.6 Conservation des données et confidentialité

Dans le but de ne pas révéler l'identité des volontaires, nous avons attribué un pseudonyme à chaque participant qui consiste à un numéro, par exemple, P1 pour participant numéro 1. Le nom des lieux de travail des participants ne sera pas révélé. Les seules personnes qui ont accès au code arbitraire sont la chercheuse, la directrice et le codirecteur. Les participants sont cités selon leur pseudonyme. Dans aucun cas il n'est possible de reconnaître le participant dans l'étude.

Les données brutes sont sauvegardées sur un ordinateur possédant un mot de passe empêchant quiconque autre que la chercheuse de pouvoir accéder au contenu de ce dernier. Seulement la chercheuse principale, la directrice et le codirecteur ont accès aux données. Les données seront conservées pour un minimum de cinq ans à la suite du dépôt de la thèse. Pendant la période de rétention, les données seront transférées sur un disque compact qui sera conservé dans un bureau verrouillé. Les données qui se trouvaient sur l'ordinateur et le disque dur seront effacées. Une copie électronique sera également conservée à l'École des sciences de l'information de l'Université d'Ottawa. À la fin de la période de rétention, le disque compact ainsi que les copies papier seront déchiquetés à l'Université d'Ottawa.

3.7 Analyse de données

L'analyse des données s'est fait grâce à certains logiciels. Les données qualitatives proviennent des entrevues, mais également des journaux de bord. Elles ont été analysées et

compilées grâce au logiciel Nvivo10 pour Mac qui est un logiciel d'analyse de données qualitatives. Toutes les sections des journaux de bord des quatre participants ont été rassemblées dans le but de remarquer les similitudes. Les journaux remplis pendant l'observation ont également été compilés et ont été analysés tout comme les autres journaux de bord. Une liste de code a été créée (Appendice F). Cette liste est séparée selon chaque section des journaux de bord; chacun des codes a été décrit et précisé. Enfin, chaque journal de bord a été analysé grâce à cette même liste. Certains de ces codes qui ont été assignés proviennent de la pyramide (Jørgensen et al., 2001) et de la liste de champs de métadonnées offerts dans le journal de bord.

En ce qui concerne les données quantitatives provenant du journal de bord, elles ont été compilées grâce à Excel dans le but de remarquer si des incidences avaient lieu. Chaque journal de bord a été compilé dans le même fichier pour pouvoir analyser réciproquement chacun des journaux. Des tableaux ont été créés à la suite pour pouvoir imager les similitudes et les récurrences.

Les fichiers des entrevues ont été également analysés à l'aide du logiciel Nvivo10. Chaque entrevue a été regardée en fonction des autres entrevues, c'est-à-dire que les quatre transcriptions ont été mises côte à côte et des similitudes ont été soulignées. Enfin, pour terminer, les observations ont été analysées selon les mêmes codes que les journaux de bord. Ils ont été compilés séparément dans le but de les comparer avec les premiers journaux de bord.

3.7.1 Validation

Premièrement, dans le but d'obtenir des données valides, une triangulation a été choisie, c'est-à-dire que trois modes de collecte de données ont été choisis dans le but de

confirmer les résultats. « Essentiellement, la triangulation est censée confirmer un résultat en montrant que les mesures indépendantes qu'on en a faites vont dans le même sens, ou du moins ne se contredisent pas » (Miles & Huberman, 2003, p. 480). C'est pour cette raison qu'en plus de faire une entrevue et de remplir les questionnaires, il a été demandé aux participants de faire une recherche avec la même requête provenant d'un usager. De cette façon, les résultats de cette recherche pouvaient démontrer que même si les participants proviennent de milieux différents, les tâches effectuées pendant la recherche peuvent démontrer des similitudes. Ceci a permis de donner une autre perspective à la recherche d'image puisque les experts devaient entreprendre une recherche avec les mêmes informations. Enfin, la triangulation permet d'obtenir des données le plus précises possible.

Deuxièmement, en ce qui concerne la validité des données, une certaine menace peut être « le passage du temps », « cette menace est reliée à la maturation des groupes stratégiques, à l'évolution de leurs expériences et de leurs connaissances par rapport à des sujets reliés à l'objet d'observation » (Roy, 2009, p. 193). Par exemple, dans ce cas-ci, c'est pour cette raison que les participants n'ont seulement eu qu'un mois pour remplir les journaux de recherche, puisque le passage du temps aurait pu fausser les résultats, si dans le cas échéant, pendant la recherche, où les participants se seraient intéressés au sujet et auraient lu plus en profondeur sur les métadonnées, ceci aurait pu fausser les données de la recherche.

3.7.2 Limitations

Étant donné que cette recherche est de type exploratoire, il importe de préciser qu'elle comporte quelques limites. Premièrement, le nombre de participants peu élevés est une première limitation. D'abord, ceci peut s'expliquer par le fait que la structure de la

collecte de données rigide a découragé certains participants. Par exemple, dans certains cas, les professionnels rencontrés ont affirmé faire de la recherche d'images dans le cadre de leur fonction, mais il n'en faisait pas suffisamment pour remplir les journaux de bord dans une période de temps raisonnable. De plus, étant donné que seulement quatre milieux ont été étudiés, dont deux universités et deux musées, il est possible d'affirmer que certaines pratiques peuvent différer dans d'autres genres d'établissements. En outre, il est possible d'affirmer que deux universités ne sont peut-être pas représentatives de tout le réseau des universités du Canada. D'ailleurs, puisque les données proviennent de milieux différents, ceci a pour résultats que les données obtenues ne seront pas totalement homogènes. L'échantillonnage non homogène peut cependant s'expliquer par le fait que si nous avons ciblé seulement un type de participants, nous n'aurions pas eu assez de réponses à nos invitations à participer à la recherche. Somme toute, il importe de préciser que le nombre de participants n'a pas empêché une récolte de données considérable et efficace.

Enfin, il faut préciser que la recherche commanditée se prête moins bien à l'induction pure, il faut souvent produire des résultats sous pression, trouver réponse à une question précise dans un temps limité, en respectant des délais fixés à d'avance et un budget restreint. (Deslauriers, 1991, p. 86).

Pour terminer, précisons que l'absence de complétude est reprochée à l'induction, car il est pratiquement impossible d'avoir un portrait complètement absolu (Depelteau, 2000, p. 69).

3.8 Conclusion

Cette recherche n'a pas l'audace de se prévaloir représentative de tous les spécialistes et les milieux, étant donné qu'elle est de type exploratoire, il est certain qu'une recherche

plus approfondie d'un milieu en particulier pourrait suivre celle-ci. Cependant, il est clair que les données récoltées pendant cette dernière se feront de façon à ce qu'elles soient le plus juste possibles. Choisir trois types de récoltes de données permet de trianguler ces dernières et de les confirmer les unes et les autres. Les journaux de bord ont permis d'obtenir des données précises sur la façon dont les professionnels recherchent. Les entrevues confirment ces données, mais aussi les élaborent. Enfin, l'observation a permis de corroborer toutes les données obtenues dans la recherche. C'est dans le prochain chapitre que toutes ces données seront exposées.

Chapitre 4 Résultats

Dans ce chapitre, les résultats de la collecte de données seront divulgués. Les vingt-quatre journaux de bord, les quatre entrevues ainsi que les quatre observations (qui sont également accompagnées de quatre autres journaux de bord supplémentaires) ont fourni un nombre considérable de données dont nous en afficherons les résultats. Ces données seront exposées selon la chronologie de leur collecte, tout d'abord les résultats des journaux de bord, ceux de l'entrevue, puis les données de l'observation. Dans les journaux de bord, les requêtes des usagers ont été analysées selon la pyramide (Jørgensen et al., 2001), les tâches établies dans le FRNB (Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, 2012) et les éléments de métadonnées. Subséquemment, les données obtenues pendant l'entrevue seront exposées selon les questions inscrites dans les journaux de bord. Pour terminer, les résultats de l'observation seront exposés. Les journaux de bord obtenus pendant l'observation ont été analysés selon les mêmes codes que ceux utilisés avec les journaux collectés en début de recherche. Pendant ce chapitre, les résultats permettront de répondre aux deux questions de recherche identifiées dans le premier chapitre.

4.1 Journaux de bord

Les participants ont rempli un total de vingt-quatre journaux de bord provenant de questions réelles d'usagers de référence en relation avec des images. Lors de ces recherches,

les participants devaient indiquer quelles métadonnées ils ont trouvé utiles pour effectuer la recherche d'images, ainsi que tout le cheminement de la recherche. Les journaux de bord se divisaient en quatre sections correspondants aux tâches établies dans le SFNB : *trouver*, *identifier*, *sélectionner* et *obtenir*. C'est selon la logique de ces subdivisions que les données seront exposées dans cette partie du chapitre.

Tout d'abord, avant même de remplir les quatre sections des journaux de bord, les participants devaient mentionner, dans la première partie, en plus de leur nom et de la date de recherche, le moteur de recherche qu'ils ont utilisé pour compléter la demande de référence. Cette question était utile pour pouvoir reproduire la recherche et par le fait même retrouver les schémas de métadonnées employés dans la base de données si nécessaire. Pendant la recherche d'images, il est arrivé parfois que les participants refassent la même requête dans deux moteurs de recherche. Le tableau 4.1 illustre les moteurs de recherches employés selon les participants et le nombre de journaux de bord qu'ils ont remplis, c'est-à-dire le nombre de recherches d'images.

Tableau 4.1 Ressources utilisées pour la recherche

Participant	Journaux remplis	Google	Catalogue maison	Base de données à abonnement	Autre site Web
P1	6	6	4	0	0
P2	3	1	0	3	0
P3	9	3	7	0	0
P4	6	1	0	4	1
Total	24	11	11	7	1

Notes : Il est arrivé qu'en cours de route le participant recherche sur plus d'une base de données et l'inscrive sur un seul journal de bord. C'est pour cette raison que le nombre de bases de données utilisées est plus élevé que le nombre de journaux de bord. Les bases de données identifiées comme *Catalogue maison* diffèrent entre les deux participants (P1 et P3) selon l'endroit où ces derniers travaillent. Conséquemment, ces données ne sont pas révélées puisque celles-ci permettraient de reconnaître les milieux de travail des participants et par conséquent de cibler leur identité.

De ce tableau, il est possible de noter que *Google* et les bases de données maison sont les endroits les plus utilisés pour les recherches. D'autre part, remarquons que le Participant 1 et le Participant 3 n'ont pas utilisé de bases de données à abonnement; ceci peut s'expliquer tout simplement par le milieu de travail dans lequel ils opèrent, contrairement aux Participants 2 et 4, les Participants 1 et 3 travaillent dans des musées où ce type de ressources n'est pas disponible, tandis que les deux autres travaillent dans des universités où les bases de données payantes sont des ressources qui sont souvent mises à la disposition des usagers et par conséquent, des professionnels. De plus, la section *Autre site Web* représente les sites Web qui ne nécessitent pas d'abonnement et qui ne sont pas liés à un musée ou une bibliothèque. Une seule des recherches a été faite sur un site comme celui-ci.

4.1.1 Trouver

La première partie du journal de bord porte sur l'élaboration de la recherche du participant. C'est dans cette étape qu'il indique la requête reçue et les préparatifs de la recherche. Cette section est divisée en quatre parties distinctes : la requête, la conception des mots clés, la description de la recherche et le nombre de résultats obtenus lors du lancement de la recherche.

4.1.1.1 Quelle est la requête que vous avez reçue?

La recherche de l'image numérique par les participants a débuté par la réception d'une requête d'un usager pour une demande particulière. Dans cette section, c'est la demande de l'utilisateur qui a été recopié.

Deux types de requêtes ont été reçus par les usagers. Tout d'abord, une partie des demandes des usagers étaient pour une recherche d'une image précise ou qui devait s'orienter vers un thème dont l'utilisateur avait précisé. Par exemple, dans P4-jdb5, un usager

demande : « Looking for posters for Godard's film pierrot le fou ». Dans un deuxième temps, les requêtes des usagers étaient d'un autre genre, les usagers connaissaient l'image, mais ils voulaient obtenir de l'information sur cette dernière, par exemple dans le P1-jdb1

I am in possession of a circa 1800's portrait that I would like some info on or if you can direct me to whom I should be consulting. I have researched on the internet and found that the National Gallery of London UK has a little information, but they have directed me to obtain from within Canada. The portrait is of the Duchess of Beaufort, Emily Frances Somerset (Smith), drawn by Chelan, engraved by WH Mote, printed by McQueen. Any help would be greatly appreciated.

Cette division du journal a été analysée en lien avec les niveaux de la pyramide de Jørgensen, Jaimes, Benitz et Chang (2001) et la liste de métadonnées créée à partir de Dublin Core et VRA Core. Chaque requête a été analysée en fonction des niveaux de la pyramide. Premièrement, chaque fois que dans la requête il y avait une référence pour un niveau de la pyramide, cette référence était notée. Par la suite, chacune de ces références a été analysée en fonction du nombre de mots clés qui s'y trouvaient. Il est possible de voir un exemple du codage dans les journaux de bord dans l'Appendice H. Dans cet exemple, l'utilisateur fait référence à trois niveaux de la pyramide : type/ technique, scène générique et objet spécifique. Dans chacun de ces niveaux, nous avons comptabilisé le nombre de mots clés présents. D'ailleurs, la classification des mots clés en lien avec la pyramide s'est faite selon l'article de Jørgensen, Jaimes, Benitz et Chang (2001) (Appendice I).

Tableau 4.2 Analyse des mots clés utilisés dans la requête

Niveau de la pyramide	Requêtes reçues ^a	Mots clés ^b
Type/Technique	5	8
Objet générique	6	9
Scène générique	4	4
Objet spécifique	10	16
Scène spécifique	4	10
Objet abstrait	1	2
Scène abstraite	1	1
Total	31	50

^a Nombre de requêtes reçues qui contenaient une ou plus références au niveau de la pyramide spécifié.

^b Nombre de mots clés utilisés dans les requêtes qui étaient associés au niveau de la pyramide spécifié

De ces mots clés analysés, certains ne pouvaient pas être reliés selon un niveau de la pyramide parce qu'ils ne sont pas des informations qui peuvent être perceptibles en regardant l'image, par exemple, une date ou le créateur de l'image. Une fois ces informations ont été compilées comme des informations externes, 22 mots clés sur 72 provenaient de ce type d'informations.

4.1.1.2 Quel(s) mot(s) clé(s) vous avez retenu pour faire la recherche?

Dans la deuxième section du journal de bord, les participants devaient indiquer quels mots clés ils ont retenus pour lancer la requête. La première étape pour le participant dans cette recherche d'image est d'extraire les informations dans la demande pour les adapter en métadonnées qui pourront être utilisées pour remplir la requête : trouver l'image. Cette section a été analysée en deux parties.

Tout d'abord, les mots retenus pour la recherche ont été analysés comme pour la première section, avec la pyramide et les mots clés. Le tableau 4.3 indique les résultats obtenus.

Tableau 4.3 Analyse des mots clés utilisés dans les recherches

Niveau de la pyramide	Requêtes reçues ^a	Mots clés ^b
Type/Technique	5	7
Objet générique	6	9
Scène générique	4	1
Objet spécifique	10	13
Scène spécifique	4	8
Objet abstrait	1	12
Scène abstraite	1	1
Total	31	51

^a Nombre de requêtes reçues qui contenaient une ou plus références au niveau de la pyramide spécifié.

^b Nombre de mots clés utilisés dans les recherches qui étaient associés au niveau de la pyramide spécifié

Dans les mots clés retenus par les professionnels pour la recherche d'image, certains provenaient d'informations externes. Dans cette section-ci, c'est 22 mots clés sur 73 qui ont été retenus.

Par la suite, ces mots utilisés ont été analysés par rapport à ceux qui se trouvaient dans la requête des usagers. Les mots ont été comparés dans le but de remarquer si les participants lançaient la requête avec de nouveaux termes ou s'ils utilisaient les mêmes que les usagers. Dans le tableau 4.4, le pourcentage des relations entre les mots utilisés par les participants par rapport aux mots utilisés dans les requêtes des usagers démontre une tendance dans le choix des mots.

Tableau 4.4 Analyse des termes utilisés pour la recherche

Participant	Pareil		Variante		Synonyme		Terme associé		Autre	
P1	20	67 %	9	30 %	0	0 %	0	0 %	1	3 %
P2	2	10 %	1	5 %	0	0 %	3	14 %	15	71 %
P3	15	83 %	1	6 %	0	0 %	0	0 %	2	11 %
P4	10	56 %	4	22 %	0	0 %	0	0 %	4	22 %
Total	47	54 %	15	17 %	0	0 %	3	3 %	22	25 %

Dans la colonne pareil, ce sont les mots qui ont été utilisés exactement de la même typographie qui ont été classés dans cette section. Dans la colonne variante, c'est le même mot qui a été utilisé, mais avec une typographie variante. Le synonyme est un mot différent utilisé qui a le même sens. Un terme associé est un mot ayant une sémantique rapprochée. Et la catégorie autre englobe les mots totalement différents et les traductions d'une langue à l'autre³.

Le tableau 4.4 permet de répondre à la première question de recherche : comment les professionnels interprètent l'information des demandes des usagers lors de la recherche d'images numériques? C'est dans cette section du journal de bord que nous pouvons remarquer le processus d'interprétation que les professionnels devaient faire dans le but de combler le besoin d'image. Il est possible d'affirmer que dans la plupart des cas, les professionnels vont extraire des mots de la demande reçue de l'utilisateur et les réutiliser pour lancer la requête.

³ Pour avoir une description complète des codes pour cette section, voir l'appendice G dans la section Code 4.2.1.2.

4.1.1.3 *Comment avez-vous fait la recherche?*

Dans la troisième section, les participants pouvaient inscrire de l'information sur la façon dont ils ont fait la recherche d'images. En ce qui concerne les informations inscrites dans cette section, un peu plus de la moitié (n=17) les réponses données (n=30), les participants ont fait une recherche simple en utilisant le moteur de recherche de la base de données. Dans les autres cas, les participants ont soit modifié leur recherche pour trouver les images à l'extérieur du Web sur un autre support (n=5), par exemple sur des négatifs ou dans des livres; soit fait une recherche à facette (n=3) ou bien rechercher sur le Web pour trouver plus d'informations sur l'image recherchée (n=4). Dans seulement une (n=1) recherche, un participant a indiqué qu'il a contacté le demandeur pour avoir plus d'informations sur l'image.

4.1.1.4 *Combien de résultats avez-vous obtenus?*

Enfin, pour terminer la section *trouver*, les participants devaient indiquer combien de résultats d'images ils ont obtenus avec la recherche qu'ils ont lancée. Le tableau 4.5 indique le nombre de résultats obtenus par les participants. Il est possible de remarquer que le P1 et le P2 ont des résultats qui se trouvent aux extrémités des échelons du tableau, tandis que les deux autres (P3 et P4) ont des résultats plutôt partagés. Enfin, plus de la moitié (n=15) des recherches ont obtenu des résultats de moins de 10 images.

Tableau 4.5 Nombre de résultats selon la recherche lancée

Participant	Aucune	1 image	2-5 images	6-9 images	10 ou plus	Plus de 100
P1	3	3	0	0	0	0
P2	0	0	0	1	1	3
P3	0	3	0	2	1	4
P4	1	0	2	0	3	0

Participant	Aucune	1 image	2-5 images	6-9 images	10 ou plus	Plus de 100
Total	4	6	2	3	5	7

Note : Il est arrivé qu'en cours de route le participant recherche sur plus d'une base de données et l'inscrive sur un seul journal de bord. C'est pour cette raison que le nombre de bases de données utilisées est plus élevé que le nombre de journaux de bord.

4.1.2 Identifier

Dans la section *identifier*, les participants inscrivait le nombre de résultats obtenus dans la recherche, les URL des images qu'ils ont consultées, puis ils mentionnaient si certaines images ont été désélectionnées et pourquoi. Ils devaient indiquer s'ils avaient eu des résultats qui ne sont pas pertinents à la recherche, c'est dans cette section qu'ils discréditaient les résultats.

4.1.2.1 *Avez-vous obtenu des résultats qui ne sont pas pertinents avec la recherche que vous avez faite?*

Tout d'abord, il était demandé aux participants d'indiquer si des résultats non pertinents à la recherche avaient été obtenus. Nous n'avons pas indiqué une définition précise de cette partie de journal de bord puisque la pertinence est très subjective selon la situation et le participant. Dans la plupart (15/24) des journaux de bord, les participants affirment avoir obtenu des résultats qu'ils ne considèrent pas comme pertinents à la recherche qu'ils ont effectuée. Dans certaines situations la date n'était pas la bonne (P1-jdb4), l'image ne se téléchargeait pas (P2-jdb5), ou l'image d'un joueur de hockey apparaissait lorsque le participant cherchait l'image d'un métis (P3-jdb4).

4.1.2.2 *Pour chaque image, pouvez-vous indiquer l'URL et l'information qui vous a indiqué pourquoi l'image n'était pas pertinente?*

Dans cette partie du journal, il était demandé aux participants d'indiquer l'URL de l'image et l'information utilisée pour indiquer que l'image n'était pas pertinente. Les

résultats obtenus ont été compilés tous simplement par le type de réponses inscrit. Dans la majorité des cas (21/33), les participants ont indiqué que c'est l'image elle-même qui leur avait permis de discréditer l'image ou non. Il n'y avait pas plus d'information à savoir si c'était une miniature ou l'image en pleine résolution qui leur a permis de faire cette étape de la recherche. De plus, quelques autres résultats étaient plus parlants, par exemple les mots clés ont également été utilisés dans cette étape de la recherche (10/33) et la technique (7/33).

4.1.3 Sélectionner

Dans la partie *sélectionner*, c'est l'utilité des types de métadonnées pour la sélection d'images que les participants devaient évaluer. La section *sélectionner* permet de répondre à la deuxième question de recherche : Comment les métadonnées influencent-elles leur sélection d'images numériques?

4.1.3.1 Pourquoi avez-vous sélectionné ou désélectionné cette image?

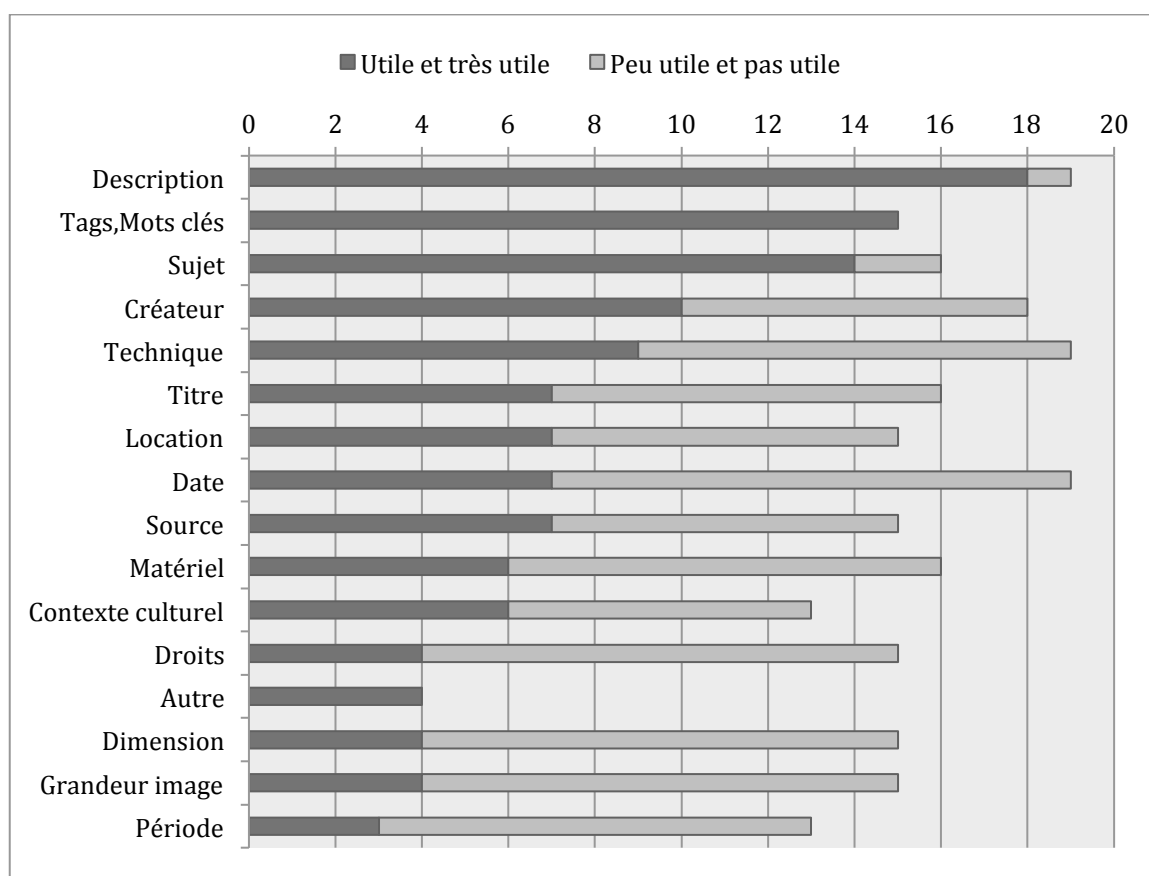
Tout d'abord, les participants devaient indiquer pourquoi ils avaient décidé que l'image était la bonne ou non. Dans les journaux qui ont obtenu des réponses pour cette section (n=20) pour la plupart du temps, c'est encore l'image elle-même qui a permis aux participants de sélectionner l'image (n=12). Sinon, dans deux autres cas, l'image a été choisie à cause de l'iconographie et la date de la prise de photo.

4.1.3.2 Classez chaque information selon leur utilité lors de la sélection de l'image.

À partir d'une liste prédéterminée, les participants devaient indiquer l'utilité de chaque métadonnée. Pour toutes les recherches d'images effectuées, les participants devaient réfléchir aux métadonnées et leur donner une utilité sur une échelle allant de *très utile* à *ne s'applique pas*.

Dans un premier temps, pour illustrer les résultats des journaux de bord, les valeurs des métadonnées indiquées par les participants ont été rassemblées en deux catégories : *utile* et *inutile*. Dans la catégorie *utile* se retrouvent les données indiquées *très utile* et *utile* et dans la catégorie *inutile* se trouvent les données *peu utile* et *pas utile*. La figure 4.1 démontre la classification de l'utilité des métadonnées dans les 24 recherches par les participants. Les données qui se trouvent en haut sont celles qui sont perçues comme les plus utiles par les participants. Les trois premiers sont alors la *description*, les *tags*, *mots clés* et *sujet*. Enfin, les *tags*, *mots clés* n'ont reçu aucune mention non utile.

Figure 4.1 Métadonnées utiles et inutiles



Dans le but de corroborer les données, les mêmes résultats ont été classés dans un tableau dans lequel chaque niveau d'utilité a été multiplié avec une valeur. Ceci permet de faire ressortir les résultats plus explicitement. Le tableau 4.6 démontre le classement provenant des usagers des métadonnées selon leur utilité dans la recherche d'images. Ces chiffres illustrent les 24 journaux de bord remplis par les usagers pendant qu'ils cherchaient les images. L'utilité des métadonnées a été classée sur une échelle de 0 à 3, zéro étant inutile et trois très utile. Chaque résultat a été multiplié par le nombre correspondant à la valeur. Les métadonnées ont été classées dans le tableau en ordre descendant selon l'utilité.

Tableau 4.6 Utilité des métadonnées

Éléments	Très utile (x3)	Utile (x2)	Peu utile (x1)	Pas utile (x0)	Total
Créateur	24	4	4	0	32
Titre	21	0	5	0	26
Date	18	2	7	0	27
Description	27	18	1	0	46
Sujet	18	16	1	0	35
Tags, Mot clés	24	14	0	0	38
Source	9	8	1	0	18
Droits	12	0	1	0	13
Location	18	2	1	0	21
Technique	18	6	1	0	25
Matériel	12	4	1	0	17
Dimension	12	0	0	0	12
Grandeur image	12	0	0	0	12
Contexte culturel	6	8	1	0	15
Période	6	2	2	0	10

À la lumière de ces données, il est possible de répondre à la deuxième question de recherche : comment les métadonnées influencent-elles la sélection d'images numériques?

Certaines métadonnées ont permis aux participants de faire leur sélection d'images, les trois plus utiles sont la *description*, les *mots clés* et les *tags*, puis le *sujet*.

4.1.4 Obtenir

C'est dans la dernière section que les participants devaient indiquer la manière qu'ils avaient employée pour obtenir l'image. Dans ce champ, les participants inscrivait une courte description de cette dernière étape de la recherche.

4.1.4.1 Comment avez-vous fait pour obtenir l'image?

Pour le total des images obtenues (n=25⁴), la plupart des fois (n=13), l'image a été obtenue en passant par un site Web, par exemple via une base de données. De plus, dans certaines bases de données les usagers peuvent y accéder par eux-mêmes et peuvent télécharger l'image désirée, ce n'est plus aux professionnels d'obtenir l'image. Nonobstant la disponibilité du téléchargement de l'image par l'utilisateur, le P2 affirme qu'il prend le temps d'informer l'utilisateur sur le droit d'auteur des images. Sept des recherches ont mené vers un serveur interne, donc les images avaient été scannées et étaient en ligne, mais non disponibles pour le public. Enfin, dans certains cas, l'image était bien celle désirée, cependant certains aspects de l'image faisaient que les participants devaient recommencer la recherche pour en trouver une autre. Parfois, l'image que l'utilisateur cherchait n'était pas encore accessible numériquement, donc, cinq recherches les ont menés à des images qui n'étaient pas numérisées, ainsi soit dans des livres, sur des négatifs ou des CD.

⁴ Ici, le nombre d'images obtenues est de 25 même si 24 journaux de bord ont été remplis. Dans un cas, l'image a été obtenue grâce aux informations d'un site Web mais aussi d'un serveur interne.

4.2 Entrevues

Les entrevues devaient être un moment où les participants pouvaient ajouter de l'information qu'il n'aurait pas eu la chance d'inscrire dans les journaux de bord. Par ailleurs, lors de l'entrevue, ils ont tous répondu à la question sur l'utilisation des métadonnées pendant la recherche d'images. Dans le tableau 4.7, les réponses ont été compilées dans le but de remarquer les incidences. Le *créateur* et les *sujets* sont les métadonnées qui ont été nommées utiles par trois des participants pendant l'entrevue. Les différences de réponse entre les participants peuvent s'expliquer par le milieu où ils travaillent et le type de demande qu'ils reçoivent. Par exemple, le P4 fournit des réponses très différentes des autres dans ce cas. Le P4 explique sa réponse comme suit : « Les étudiants en beaux-arts eux ils aimeraient avoir d'autres façons de chercher. Parce qu'eux vont s'intéresser à un style, une technique. Couleur par exemple. Il ne cherche pas quelques choses de précis. Il cherche plus une approche une façon de travailler, ça serait technique, j'imagine. » (P4-ent).

Tableau 4.7 Métadonnées utiles pour la recherche d'image

Éléments	P1-ent	P2-ent	P3-ent	P4-ent	Total
Créateur	1	1	1	0	3
Sujet	1	1	1	0	3
Titre	1	0	1	0	2
Date	1	0	1	0	2
Description	1	0	1	0	2
Tags, mots clés	0	1	1	0	2
Location	0	0	0	1	1
Technique	0	1	0	0	1
Période	0	0	0	1	1
Source	0	0	0	0	0
Droits	0	0	0	0	0

Éléments	P1-ent	P2-ent	P3-ent	P4-ent	Total
Matériel	0	0	0	0	0
Dimension	0	0	0	0	0
Grandeur image	0	0	0	0	0
Contexte culturel	0	0	0	0	0

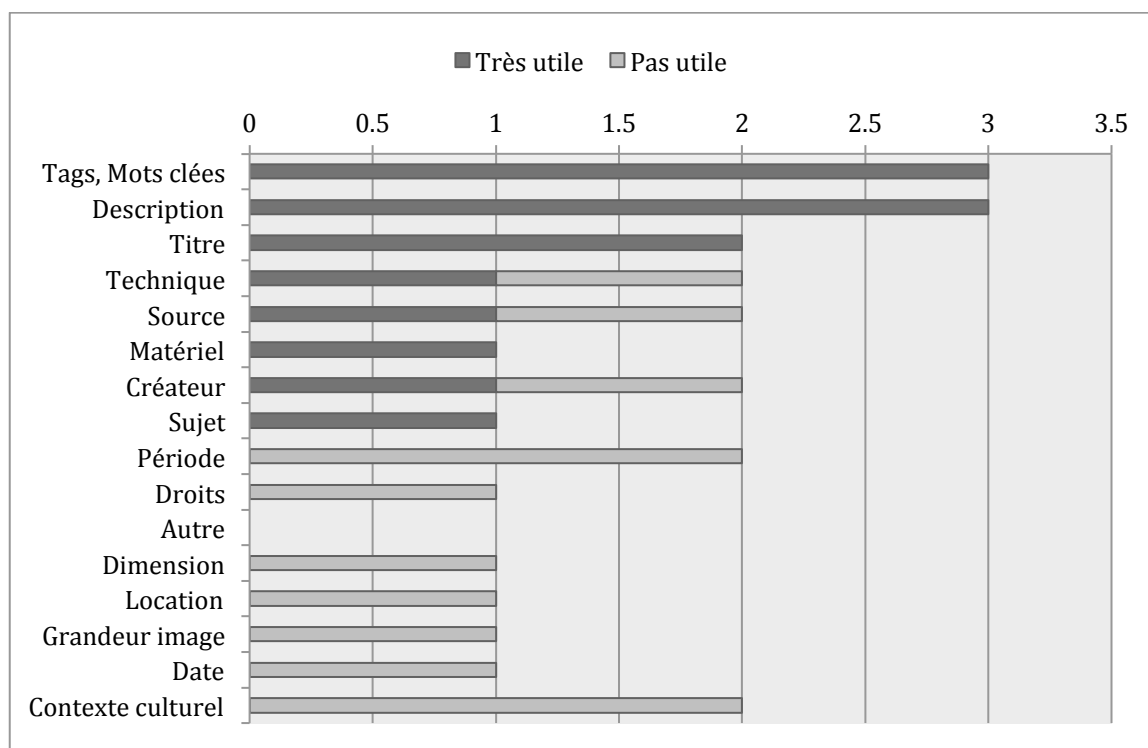
4.3 Observation

Lors de l'observation, les participants devaient faire une recherche d'image numérique selon la même demande d'un usager. Pour cette partie de la recherche, nous avons pris une demande qui venait réellement d'un usager qui est celle-ci :

Je recherche un tableau représentant, je crois, le Cap-à-l'original au Bic, Bas-Saint-Laurent, QC. J'ai vu ce tableau au musée il y a longtemps; on y peignait la silhouette du cap et il y avait un feu et des amérindiens. Je peux vous envoyer une photo de la silhouette du C. à O., silhouette très particulière avec un rocher qui se détache du cap et forme un pic.

Les résultats ont également été compilés selon leur positivité et négativité, donc utiles et non utiles tout comme pour la section des journaux de bord. Les résultats sont disponibles dans la figure 4.2. Les éléments complètement en haut de la figure sont ceux qui ont été retenus comme plus utiles. *Description* et *tags, mots clés* sont arrivés à égalité en première position. Puis *sujet* est également un élément qui est choisi comme utile dans la recherche d'images.

Figure 4.2 Métadonnées utiles et inutiles de l'observation



Tout comme pour les résultats des autres journaux de bord, ceux obtenus pendant l'observation ont été multipliés selon leur utilité allant de trois pour très utile et zéro pour pas utile. Le tableau 4.9 démontre les résultats obtenus.

Tableau 4.8 Utilité des métadonnées pendant l'observation

Champs	Très utile x3	Utile x2	Peu utile x1	Pas utile x0	Total
Créateur	3	0	0	0	3
Titre	3	0	1	0	4
Date	0	0	1	0	1
Description	3	4	0	0	7
Sujet	0	4	0	0	4
Tags, Mots clés	3	4	0	0	7
Source	3	0	0	0	3
Droits	0	0	0	0	0
Location	0	0	0	0	0

Champs	Très utile x3	Utile x2	Peu utile x1	Pas utile x0	Total
Technique	3	0	0	0	3
Matériel	3	0	0	0	3
Dimension	0	0	0	0	0
Grandeur image	0	0	0	0	0
Contexte culturel	0	0	0	0	0
Période	0	0	0	0	0

Pendant l'observation, les participants ont eu de la difficulté à faire la recherche de l'image, car c'est ce genre de demande des usagers qui demande une communication entre le professionnel et l'utilisateur. Les participants avaient beaucoup d'informations qu'ils auraient aimé vérifier avec l'utilisateur. Le Participant 4 affirme que la présence de l'utilisateur pourrait aider, beaucoup de vérification à faire. P2 a fait une recherche pour voir à quoi ressemblait le Cap à l'Original, pour pouvoir le discerner dans une image.

4.4 Conclusion

Pour terminer, dans ce chapitre, les résultats de la collecte de données ont été présentés selon l'ordre de leur collecte. Les journaux de bord, l'entrevue et l'observation pour les quatre participants ont créé un nombre impressionnant de données. Seulement celles qui sont les plus pertinentes ont été présentées dans la section. Certaines données très intéressantes, mais ne se reliant pas au sujet de recherche ont été omises, par exemple les participants approfondissaient sur divers sujets dans les entrevues, mais puisque ces derniers n'étaient pas reliés directement avec les questions de recherches, nous n'avons pas jugé bon de les insérer dans les résultats. Enfin, les informations récoltées pendant la collecte de données ont servi à répondre aux questions de la recherche qui avaient été établies dans le

premier chapitre. C'est dans le chapitre suivant que les informations récoltées pourront ouvrir la discussion.

Chapitre 5 Discussion

Dans le chapitre précédent, les données obtenues pendant les collectes de données ont été exposées et, de ces dernières, nous en avons tiré quelques sujets qui seront approfondis et discutés ici. Comme nous l'avons déjà mentionné, nous avons eu une démarche inductive, qui

s'agit de procéder à des observations particulières de la réalité étudiée, de regarder, de chercher à tout voir si possible, à tout entendre, à tout sentir, etc., puis d'en induire des énoncés généraux (des concepts, des hypothèses, des théories, des lois) qui rendent compte de la réalité ». (Depelteau, 2000, p. 56)

Ces sujets, concepts, diviseront ce chapitre et ils sont : le contexte, l'analyse de la requête et la médiation (la sélection de la base de données, la formulation de la requête et l'utilisation des métadonnées). D'ailleurs, c'est également dans ce chapitre que la réponse à la question générale de recherche prendra forme : comment les chercheurs expérimentés qui font de la recherche d'images numériques utilisent-ils les métadonnées associées aux images numériques?

5.1 Le contexte

Comme nous l'avons démontré lors de la revue de la littérature, bien des études exposent des recherches qui s'attardent aux comportements des usagers relativement au

contexte dans lequel la recherche a été réalisée. Dans notre étude, les participants et les requêtes proviennent de deux milieux différents et ceci a un impact direct sur tous les résultats, car cet environnement sculpte le paysage dans lequel les participants évoluent.

Enfin, le genre d'institution attire un type d'utilisateur typique par le fait qu'il consulte cet endroit selon son besoin d'images. Conséquemment, les usagers qui ont adressé une requête d'images étaient également un public spécifique, parce que les demandes étaient axées sur ce genre d'établissements. Le contexte dans lequel la demande de recherche d'images a été présentée par l'utilisateur a été un point important, puisque ceci a des répercussions non seulement sur le type de demandes reçues, mais aussi sur le type de bases de données consultées et par conséquent sur l'utilisation des métadonnées. Nous avons établi que nous avons reçu différents types de requêtes pendant la recherche et nous discuterons dans ce chapitre à savoir comment les distinctions de ces requêtes teintent les résultats de la recherche.

5.2 L'analyse de la requête

Dans un premier temps, regardons les requêtes des usagers des services de référence. Il est important d'en discuter, puisque c'est à partir de ces dernières que les professionnels lancent leur propre requête et font la recherche de l'image. Nous pouvons classer les demandes des usagers selon leurs types et celles-ci varient selon le contexte dans lequel elles sont acheminées.

En ce qui concerne les musées, les requêtes ont pu être classées selon deux types de demandes. En premier lieu, certains usagers demandaient des informations sur une œuvre qu'ils possédaient. Par exemple:

I have an enquiry about the artwork that I own, and I kindly ask for your professional opinion, if it's possible. I have recently purchased it from a private owner in London. It's an abstract oil and mixed media on board. The owner told me his father bought it on [sic] moving sale in Bristol a few years ago. I was researching the signature on this painting, and I think it is similar to a signature of Harold Town. I am not familiar enough with this artist [sic] work to be able to tell if the style is also similar. The painting is dated '60'. (P1-jdb6)

Par la suite, un deuxième type de demande a été observé; les usagers réclamaient une image qui se trouvait dans le musée ou qui était étroitement liée avec la collection, par exemple, « Photo du livre d'autographe de Louis Riel » (P3-jdb7).

Puis, les requêtes provenant des universités ont été également classées en deux autres catégories. Tout d'abord, les usagers ont demandé des images pour un concept ou pour illustrer un travail. Une illustration de ce genre de demande serait « Images pour un essai photographique discutant l'harcèlement [sic] ou l'abus sexuel des femmes au domicile » (P2-jdb3). Le deuxième type de requêtes ressemble plus à celles reçues par les musées lors d'une demande pour une image reliée à une œuvre spécifique, par exemple (P4-jdb5) « Looking for posters for Godard's film pierrot le fou [sic] ».

Ces différences entre les requêtes des universités et des musées se sont fait également sentir lors de l'analyse avec la pyramide. Par exemple, les requêtes provenant des musées étaient plus vers les niveaux de la pyramide spécifiques et moins de référence aux niveaux génériques. D'ailleurs, nous avons remarqué qu'aucune référence n'était en lien avec les niveaux abstraits. Par contre, lorsque nous regardons les requêtes des universités, il y avait plus de références avec les niveaux génériques et abstraits que ceux spécifiques.

Malgré les différences entre les requêtes reçues par les deux types d'établissement, d'ensemble, plusieurs similitudes sont ressorties selon la catégorie de requêtes. Les informations de la pyramide qui ont reçu le plus de résultats sont dans la sphère spécifique, les objets et les scènes. Par la suite, c'est la sphère générique qui a reçu le plus de résultats.

Quoi que nos résultats diffèrent un peu entre les deux sources de requêtes, celles reçues par les musées et celles reçues dans les universités, nos résultats globaux peuvent être comparés à ceux d'autres études. Même si notre étude est de type exploratoire lorsque nos résultats globaux sont comparés avec d'autres, deux autres recherches soutiennent des résultats semblables aux nôtres en ce qui concerne les proportions relatives à la sphère générique, spécifique et abstraite, celles de Armitage et Enser (1997, p. 293) et de Choi et Rasmussen (2003, p. 506). Plus précisément, les proportions du nombre d'informations selon ses divisions dans les sphères sont de deux pour un pour les spécifiques comparativement au générique. Enfin, peu de résultats se retrouvent dans la zone abstraite.

5.3 Médiation

Le professionnel a joué tout au long du repérage d'image un rôle important, celui du médiateur entre l'image et l'utilisateur du service de référence. À partir de l'importance de ce rôle, nous pensons que trois sujets méritent d'être approfondis : la sélection de l'outil, la formulation de la requête et l'utilisation des métadonnées.

5.3.1 Sélection de l'outil

Comme nous avons déjà discuté précédemment, le type d'institution attire un type d'utilisateur distinct puisqu'il consulte cet endroit selon son besoin. Par conséquent, les usagers qui ont adressé une requête d'images étaient également un public spécialisé, parce que les demandes étaient axées sur ces types d'institutions. C'est pourquoi les ressources

disponibles varient d'un endroit à l'autre, les musées n'ont pas les mêmes besoins que les universités, et par conséquent ils n'ont pas les mêmes ressources.

Toutefois, en ce qui concerne le choix des bases de données dans les musées fait par le professionnel, nous pouvons affirmer que ce choix a été grandement influencé par le type de requêtes puisque la plupart avaient un lien direct avec leur propre collection. Par ailleurs, en ce qui concerne les requêtes pour les universités, certaines requêtes étaient directement relatives à l'histoire de l'art. C'est pour cette raison que des bases de données à abonnement ont été sélectionnées. Enfin, ceci démontre que le choix de la base de données est influencé par le type de requêtes qui diffèrent dans les deux types d'établissements. En contrepartie, lorsque nous avons récolté les résultats des choix d'outils de recherche, il est intéressant de remarquer que peu importent les différences des outils disponibles, *Google* a été utilisé dans 11 fois sur 24.

5.3.2 Formulation de la requête

Dans la deuxième partie de la médiation, les participants devaient prendre la requête des usagers et la traduire pour lancer la recherche d'image. « Image retrieval is essentially an act of translation. Users' cognitive image needs must be translated into external descriptions or depictions to communicate these needs to other humans and to information systems » (Goodrum, 2005, p. 46). Lorsque l'on observe les niveaux de la pyramide qui ont été représentés pour commencer le repérage, il y a les objets spécifiques, les objets abstraits et les objets génériques; les requêtes lancées par les participants pour cette section ressemblent beaucoup à celles provenant des usagers. À cet instant, nous pouvons affirmer que l'acte de traduction n'est pas encore pleinement accompli, car si nous mettons côte à côte les informations de la pyramide provenant des usagers et les recherches créées par les

participants, nous pouvons remarquer que la division selon la pyramide est semblable (tableau 5.1).

Tableau 5.1 Comparaison des niveaux de la pyramide de la requête et la recherche

Niveau de la pyramide	Requêtes reçues ^a	Mots clés dans les requêtes ^b	Mots clés dans les recherches ^c
Type/Technique	5	8	7
Objet générique	6	9	9
Scène générique	4	4	1
Objet spécifique	10	16	13
Scène spécifique	4	10	8
Objet abstrait	1	2	12
Scène abstraite	1	1	1
Total	31	50	51

^a Nombre de requêtes reçues qui contenaient une ou plus mots clés associés au niveau de la pyramide spécifié.

^b Nombre de mots clés utilisés dans les requêtes qui étaient associés au niveau de la pyramide spécifié.

^c Nombre de mots clés utilisés dans les recherches qui étaient associés au niveau de la pyramide spécifié.

Néanmoins, dans un niveau de la pyramide, les informations utilisées par les participants ne sont pas les mêmes que celles reçues par les usagers : ce sont les objets abstraits. Cette différence marquée peut s'expliquer par deux requêtes provenant du même participant, le P2. Les usagers cherchaient pour un thème en particulier, par exemple « Images pour un essai photographique discutant l'harcèlement ou l'abus sexuel des femmes au domicile » (P2-jdb3). Le professionnel a alors ajouté des synonymes et des traductions pour la recherche. Puis, le deuxième niveau de la pyramide qui n'a pas été utilisé par les participants mêmes si les usagers en avaient indiqué dans leur demande est la scène générique. Ce niveau de la pyramide est très général, par exemple, un portrait en fait partie.

Une recherche (Goodrum, 2005) démontre que les professionnels utilisent souvent des synonymes lors de la méditation. Dans notre recherche, en ce qui concerne l'utilisation

des termes par les participants lors de la première opération du repérage d'image, *trouver*, les mots utilisés pour lancer la recherche ont été analysés par rapport à ceux qui se trouvaient dans la demande des usagers. Le Participant 1 et le Participant 3 sont plus enclins à utiliser les mêmes termes pour la recherche. Encore une fois, cette ressemblance peut s'expliquer par l'environnement dans lequel ils travaillent.

5.3.3 L'utilisation des métadonnées

Lors de l'observation de l'usage des métadonnées par les participants, nous avons pu faire deux constats : l'image comme métadonnée et la tâche en lien avec l'usage des métadonnées.

5.3.3.1 L'image comme métadonnée

En ce qui concerne l'utilisation des métadonnées lors du repérage d'images, les données recueillies pendant la collecte nous ont permis de faire quelques remarques. Non seulement les métadonnées qui étaient associées avec les images recherchées ont été utilisées par les participants, mais aussi les images elles-mêmes. Dans la tâche *identifier* et dans la première étape de la sélection de l'image où les participants devaient répondre à une question ouverte, les participants ont grandement affirmé que l'image elle-même permettait d'accomplir ces tâches, qu'elle leur était utile. Ici, l'image peut être une miniature ou l'image en pleine résolution, mais elle devient utile par l'information qu'elle véhicule. Même si notre bassin de participants est petit et que notre recherche est de type exploratoire, d'autres études remarquent également l'importance de l'image dans les étapes du repérage. Notamment, Hasting (1999) démontre dans un article sur les moteurs de recherche d'images que 60 % des recherches faites pouvaient être répondues par la miniature (p.448). De surcroît, une autre recherche observe les mêmes résultats que nous. « Browsing a page of thumbnails in order to

locate the specific picture they want to retrieve follows keyword searching» (Neal, 2008, p. 28).

Enfin, notre recherche tend à démontrer que l'image fait partie du processus de repérage par elle-même. Selon la définition qui suit des métadonnées : « Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information » (National Information Standards Organization, 2004, p. 1). Pourrions-nous affirmer que l'image numérique en soit devient métadonnée lors du repérage?

5.3.3.2 La tâche et l'usage de métadonnée

De plus, lors de la sélection de l'image, les participants devaient remplir la section avec la liste des métadonnées, et ici il importe de regarder les résultats. Chaque métadonnée de la liste a été au minimum remplie comme très utile à deux reprises. Il est certain que certaines métadonnées ont reçu un plus grand nombre de résultats, mais il importe tout de même d'affirmer que chaque métadonnée est utilisée par les participants et que faire un choix dans le but d'en éliminer serait une erreur. Chaque métadonnée est alors utilisée par les participants et ceci dépend de la recherche qu'ils font et de la requête qu'ils reçoivent.

Pour terminer, nous avons également remarqué une différence dans l'utilisation des métadonnées par les professionnels selon l'étape du repérage. Nous avons observé que dans le début du processus, les professionnels ne font pas encore beaucoup de traduction et qu'ils lancent la requête avec beaucoup des informations provenant des usagers. Ces informations se trouvent bien souvent dans les informations extérieures selon la pyramide, donc créateur, date. Puis, dans les autres étapes de repérage, les professionnels débutent vraiment un travail

d'interprétation, car ils utilisent d'autres métadonnées pour remplir les tâches. Par exemple, nous avons observé que la *description* est bien souvent utilisée et l'image aussi. « Les dates aussi ça peut être pratique, mais pas... des fois tu tapes une date dans le catalogue et tu vas avoir des milliers de résultats, mais avec la description tu élimines déjà, sujet keyword. C'est vraiment celles qui pour moi les plus importantes. » (P3-ent)

5.4 Conclusion

Pour terminer, les discussions que nous avons apportées ici sont un tremplin pour les recherches futures. Grâce à nos résultats de la collecte de données, nous avons ainsi démontré que le contexte de la requête a établi des différences dans les résultats de l'étape de la médiation, par exemple, l'incidence directe sur la sélection de l'outil. De plus, nous avons également remarqué que la formulation de la recherche est étroitement liée avec la requête de l'utilisateur. Enfin pour terminer, nous avons aussi exposé l'utilisation des métadonnées selon les participants. Ces discussions permettront une réflexion sur des recherches futures.

Chapitre 6 Conclusion

Ce chapitre clôt la thèse. Nous débuterons par un résumé de notre recherche pour, par la suite, exposer des pistes de recherches futures. Une courte critique de la recherche suivra avant de terminer avec les dernières réflexions.

6.1 Résumé

Cette recherche de type exploratoire a analysé l'utilisation des métadonnées grâce à des données provenant de trois étapes de collecte. Elle voulait examiner l'emploi des métadonnées chez les professionnels qui font de la recherche d'images dans le contexte d'un service de référence en répondant à ces deux questions de recherche :

1. Comment les professionnels interprètent-ils l'information des demandes des usagers en métadonnées lors de la recherche d'images numériques?
2. Comment les métadonnées influencent-elles leur sélection d'images numériques?

Quatre participants provenant de deux musées et deux universités ont rempli en tout 24 journaux de bord, participé à une entrevue et à une observation; toutes les données obtenues ont été cumulées. Quatre types de requêtes ont été comparés. Nous avons remarqué plusieurs phénomènes dont nous avons discuté précédemment. Tout d'abord, le contexte dans lequel les participants travaillent est très important à la recherche; ceci ne nous avait pas

tout de suite frappés au début de la recherche. Enfin, nous avons démontré qu'il y a des différences dans les requêtes selon l'environnement; les requêtes qui provenaient des participants pouvaient être catégorisées selon quatre types d'après le contexte et le type de requête. De plus, la sélection de l'outil est également caractérisée par ce contexte, c'est-à-dire que le type de ressource disponible varie selon l'établissement et par conséquent varie selon les besoins, donc les requêtes.

Nous avons établi que l'interprétation des informations dans les requêtes faites par les usagers était la source première pour la recherche du professionnel, c'est-à-dire que la plupart des mots clés utilisés dans la recherche provenaient directement de la requête de l'usager. De plus, en ce qui concerne l'utilisation des métadonnées, nous avons observé que toutes les métadonnées ont été au moins à trois reprises comme utiles ou très utiles.

Nous avons également avancé des pistes de recherche sur le type de demande que les professionnels reçoivent, par exemple, une demande de référence n'est pas toujours une demande pour une image, mais parfois pour de l'information sur une image. Nous avons également apporté un regard différent sur l'utilisation des métadonnées lors du repérage d'image, et ce, en déconstruisant le repérage en quatre tâches bien définies, ce qui est innovateur dans les recherches faites sur le repérage. Enfin, tous ces résultats nous permettent de réfléchir sur des sujets de recherches postérieures qui permettraient d'éclairer le rôle du professionnel comme médiateur lors du repérage d'image.

6.2 Recherches futures

Il est certain que cette recherche n'est pas exhaustive, mais elle apporte des pistes pour des recherches ultérieures, car le repérage d'images est un sujet qui est d'actualité et qui est très étudié.

The main difficulty arises, I believe, because there is not a happy marriage of image and the supporting texts that would provide full identification, historical background, and comprehensive indexing. The images, of course, are vital, but the infrastructure of the archive that cataloguing and indexing systems provide us with is what dictates how it can be used and what results are possible. » (Roberts, 2001, p. 911).

En ce sens, il serait important de se questionner sur la granularité de l'indexation relativement au domaine dans lequel les recherches sont faites. Est-ce qu'une bibliothèque d'art visuel devrait avoir un niveau d'indexation différent d'une bibliothèque muséale? Est-ce que les schémas de métadonnées doivent également être différents dans des établissements différents? Que faire pour les bibliothèques avec des collections diversifiées? Jørgensen (2003) se questionnait déjà sur ce propos, elle apporte ces raisonnements : « What is the appropriate level of specificity of indexing? Are varying levels of specificity needed for different domains and tasks? » (p.262). Dans une recherche future, il pourrait être intéressant de classer les types de requêtes selon les établissements selon les pôles de Fidel (1997) et de voir si selon le type de requêtes, les façons de faire le repérage se ressemblent et si dans ce cas, l'établissement dans lequel le professionnel évolue joue vraiment un rôle dans l'utilisation des métadonnées.

Enfin, il serait intéressant aussi de s'attarder aux choix des bases de données par les professionnels en fonction du type de repérage, car peu de recherches couvrent le sujet de la sélection des bases de données par les professionnels.

6.3 Critique de la recherche

Pour terminer, puisque c'est une recherche de type exploratoire, cette étude est nécessaire pour démontrer que dans ce domaine encore peu d'études ont été faites et qu'un

manque y règne. D'avoir seulement quatre participants, deux dans chaque établissement, n'est pas suffisant pour dire que la recherche couvre toutes les réalités sur le sujet.

Cependant, puisque nous n'avions pas ciblé d'avance ces quatre participants, nous avons réussi à démontrer les distinctions dans les différents centres documentaires. Pour terminer, nous avons analysé 24 journaux de bord, même si ce nombre semble peu, nous avons récolté un nombre de données suffisant pour s'interroger et lancer des pistes de réflexion sur le sujet.

6.4 Réflexion finale

Pour terminer, nous avons démontré que le repérage d'images par les professionnels est un milieu où peu d'études ont été faites en lien avec les schémas de métadonnées. Avec les technologies en mouvement constant et les innovations dans ce domaine, il importe de se pencher sur la question des schémas de métadonnées dans les musées, les bibliothèques et les centres d'archives. De mettre ses images sur la toile est une façon de communiquer, de partager sa collection, mais si ces dernières tombent dans l'oubli et ne peuvent être retracées, le partage de l'image devient alors inutile.

Liste de références

- Armitage, L. H., & Enser, P. G. B. (1997). Analysis of user need in image archives. *Journal of Information Science*, 23(4), 287–299.
- Artstor. (n.d.). Index. Retrieved from www.artstor.com
- Attig, J., Copeland, A., & Pelikan, M. (2004). Context and meaning: The challenges of metadata for a digital image library within the university. *College & Research Libraries*, 65(3), 251–261.
- Beaudoin, J. E., & Brady, J. E. (2011). Finding visual information: A study of image resources used by archaeologists, architects, art historians, and artists. *Art Documentation*, 30(2), 24–36.
- Carlyle, A. (2006). Understanding FRBR as a Conceptual Model. *Library Resources & Technical Services*, 50(4), 264–274. Retrieved from <http://alcts.metapress.com/index/J83Q31W747N53531.pdf>
- Chen, H. (2001). An analysis of image queries in the field of art history. *Journal of the American Society for Information Science and Technology*, 52(3), 260–273. [http://doi.org/10.1002/1532-2890\(2000\)9999:9999<::AID-ASI1606>3.0.CO;2-M](http://doi.org/10.1002/1532-2890(2000)9999:9999<::AID-ASI1606>3.0.CO;2-M)
- Choi, Y., & Rasmussen, E. M. (2003). Searching for images: The analysis of users' queries for image retrieval in American history. *Journal of the American Society for Information Science and Technology*, 54(6), 498–511. <http://doi.org/10.1002/asi.10237>

- Christians, C. G. (2011). Ethics and politics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (4th ed., pp. 61–80). Sage Publications.
- Chung, E., & Yoon, J. (2011). Image needs in the context of image use: An exploratory study. *Journal of Information Science*, 37(2), 163–177.
<http://doi.org/10.1177/0165551511400951>
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles: SAGE Publications.
- Dégez, D., & Ménillet, D. (2001). *Thésauriglossaire des langages documentaires : un outil de contrôle sémantique*. Paris: Paris ADBS Éditions.
- Depelteau, F. (2000). *La démarche d'une recherche en sciences humaines : de la question de départ à la communication des résultats*. Sainte-Foy, Québec: Presses de l'Université Laval.
- Deslauriers, J.-P. (1991). *Recherche qualitative : guide pratique*. Montréal: McGraw-Hill.
- Dublin Core Metadata Initiative. (2013). DCMI - metadata terms. Retrieved from
<http://dublincore.org/documents/dcmi-terms/>
- Enser, P. (2008). The evolution of visual information retrieval. *Journal of Information Science*, 34(4), 531–546. <http://doi.org/10.1177/0165551508091013>
- Enser, P. G. B., & McGregor, C. G. (1992). *Analysis of visual information retrieval queries*.
- Enser, P. G. B., Sandom, C. J., Hare, J. S., & Lewis, P. H. (2007). Facing the reality of semantic image retrieval. *Journal of Documentation*, 63(4), 465–481.
<http://doi.org/10.1108/00220410710758977>
- Fear, K. (2010). User understanding of metadata in digital image collections: Or , what exactly do you mean by “coverage”? *The American Archivist*, 73(1), 26–60.

- Fidel, R. (1997). The image retrieval task: Implications for the design and evaluation of image databases. *New Review of Hypermedia and Multimedia*, 181–199. Retrieved from <http://www.tandfonline.com/doi/full/10.1080/13614569708914689>
- Goodrum, A. (2003). Visual resource reference collaboration between digital museums and digital libraries. *D-Lib*, 9(2). Retrieved from <http://www.dlib.org/dlib/february03/goodrum/02goodrum.html>
- Goodrum, A. (2005). I can't tell you what I want, but I'll know It when I see it. *Reference & User Services Quartely*, 45(2), 46–53.
- Goodrum, A., & Spink, A. (2001). Image searching on the Excite Web search engine. *Information Processing & Management*, 37(2), 295–311. [http://doi.org/10.1016/S0306-4573\(00\)00033-9](http://doi.org/10.1016/S0306-4573(00)00033-9)
- Greenberg, J. (2001). A quantitative categorical analysis of metadata elements in image-applicable metadata schemas. *Journal of the American Society for Information Science and Technology*, 52(11), 917–924.
- Greisdorf, H., & O'Connor, B. (2002). Modelling what users see when they look at images: A cognitive viewpoint. *Journal of Documentation*, 58(1), 6–29. <http://doi.org/10.1108/00220410210425386>
- Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques. (2012). *Fonctionnalités requises des notices bibliographiques*.
- Hastings, S. K. (1994). *An exploratory study of intellectual access to digitized art images*. The Florida State University.
- Hastings, S. K. (1999). Evaluation of Image Retrieval Systems: Role of User Feedback. *Library Trends*. Retrieved from <http://eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ603797>

- Jørgensen, C. (1995). *Image attributes: An investigation*. Syracuse University. Retrieved from http://surface.syr.edu/it_etd/39/
- Jørgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, 34(2), 161–174. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0306457397000770>
- Jørgensen, C. (2003). *Image retrieval: Theory and research*. Laham, Maryland: Scarecrow Press.
- Jørgensen, C., Jaimes, A., Benitez, A. B., & Chang, S. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, 52(11), 938–947.
- Kramer, E. F. (2005). IUPUI image collection: A usability survey. *OCLC Systems & Services*, 21(4), 1–20.
- Matusiak, K. K. (2006). Information seeking behavior in digital image collections: A cognitive approach. *The Journal of Academic Librarianship*, 32(5), 479–488. <http://doi.org/10.1016/j.acalib.2006.05.009>
- Miles, M. B., & Huberman, A. M. (2003). *Analyse des données qualitatives*. Paris: De Boeck.
- National Information Standards Organization. (2004). *Understanding metadata*. Bethesda: NISO Press.
- Neal, D. (2008). News photographers, librarians, tags, and controlled vocabularies: Balancing the forces. *Journal of Library Metadata*, 8(3), 199–219. <http://doi.org/10.1080/19386380802373936>
- Poliakoff, S. (1999). *Shooting the past*.

- Roberts, H. E. (2001). A picture is worth a thousand words: Art indexing in electronic databases. *Journal of the American Society for Information Science and Technology*, 52(11), 911–916.
- Roy, S. N. (2009). L'étude de cas. In B. Gauthier (Ed.), *Recherche sociale : de la problématique à la collecte des données* (pp. 199–226). Montréal: Presses de l'Université du Québec.
- Savoie-Zajc, L. (2000). L'entrevue semi-dirigée. In B. Gauthier (Ed.), *Recherche sociale : de la problématique à la collecte des données* (pp. 335–360). Montréal: Presses de l'Université du Québec.
- Shatford Layne, S. (1994a). Artists, Art Historians, and Visual Art Information. *The Reference Librarian*, 22(47), 23–36. http://doi.org/10.1300/J120v22n47_03
- Shatford Layne, S. (1994b). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45(8), 583–588.
- Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly*, 6(3), 39–62. <http://doi.org/10.1300/J104v06n03>
- Smith, N. (2010). Official blog: Ooh! Ahh! Google Images presents a nicer way to surf the visual web. Retrieved from <http://googleblog.blogspot.ca/2010/07/ooh-ahh-google-images-presents-nicer.html>
- Vander Wal, T. (2007). Folksonomy coinage and definition. *Vanderwal.net*. Retrieved from <http://vanderwal.net/folksonomy.html>
- Virtual Resources Association. (2014). VRA CORE - A data standard for the description of works of visual culture: Official web site. Retrieved September 5, 2014, from <http://www.loc.gov/standards/vracore/schemas.html>

Zentner-Raasch, C. (2010). Digital images, management and metadata: The long tail, or the order of order? *Library Philosophy and Practice*. Retrieved from <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1356&context=libphilprac>

Appendices

Appendice A : Journal de bord

Recherche d'images numériques et utilisation des métadonnées chez des chercheurs expérimentés.

Mariane Léonard
 Étudiante à la maîtrise
 École des sciences de l'information, Université d'Ottawa
 [REDACTED]@uottawa.ca

Questionnaire à remplir pour chaque image recherchée.

Nom :	
Date de la recherche :	
URL du moteur de recherche	

Trouver	
Quelle est la requête que vous avez reçue?	
Quel(s) mot(s) clé(s) vous avez retenu pour faire la recherche?	
Comment avez-vous fait la recherche? (Courte description)	
Combien de résultats avez-vous obtenus?	

Identifier		
<i>Répondez à cette section si vous avez obtenu plus d'un résultat de recherche.</i>		
Avez-vous obtenu des résultats qui ne sont pas pertinents avec la recherche que vous avez faite?		
Pour chaque image, pouvez-vous indiquer l'URL et l'information qui vous a indiqué pourquoi l'image n'était pas pertinente? <i>Vous pouvez ajouter des lignes si vous avez plus d'images à ajouter.</i>	URL de l'image	Information

Sélectionner							
<i>Vous répondez à ces questions lorsque vous êtes sur la page où se trouvent les métadonnées. Vous pouvez dupliquer cette section si vous avez cliqué sur plus d'une image.</i>							
URL de l'image							
Pourquoi avez-vous sélectionné ou désélectionné cette image?							
Classez chaque information selon leur utilité lors de la sélection de l'image.	Informations	Très utile	Utile	Peu utile	Pas utile	Ne sait pas ce que cela signifie	Ne s'applique pas
	Creator						
	Title						
	Date						
	Description						
	Subject						
	Tags, keywords						
	Source						
Rights							

	Location						
	Technique						
	Material						
	Measurements						
	Image size						
	Cultural Context						
	Style Period						
	Autre : <i>Veillez indiquer le type d'informations</i>						

Obtenir	
Comment avez-vous fait pour obtenir l'image?	
Notes	

Les données qui seront récoltées dans cette recherche seront totalement anonymes, en aucun cas le nom des participants ne sera cité et il sera impossible de les reconnaître. Les données seront également conservées sur un disque dur crypté conservé à l'École de sciences de l'information à l'Université d'Ottawa, donc aucune personne extérieure au projet ne pourra accéder aux résultats de la recherche et aux notes d'entrevues. Seulement la chercheuse principale Mariane Léonard, la directrice D^{re} Mary Cavanagh et le codirecteur le Dr Tom Delsey auront accès aux données. Le tout sera fait selon la réglementation du comité d'éthique de l'Université d'Ottawa. D'ailleurs, le projet sera supervisé par les D^{re} Mary Cavanagh et D^r Tom Delsey de l'École des sciences de l'information de l'Université d'Ottawa.

Appendice B : Questions d'entrevue

Recherche d'images numériques et utilisation des métadonnées chez des chercheurs expérimentés.

L'entrevue sera semi-structurée, alors quelques questions permettront de laisser la parole aux professionnels qui pourront développer sur ces sujets. Quelques questions ont été notées dans le but de diriger la discussion.

Questions pour les entrevues

1. Quel moteur de recherche utilisez-vous le plus souvent pour faire votre recherche d'images?
2. Quels sont les critères qui sont importants lorsque vous sélectionnez une image?
3. Y-a-t-il de l'information qui pourrait vous aider lors de la recherche d'image que vous ne trouvez pas sur le moteur de recherche?
4. Pouvez-vous me parler un peu plus de la façon dont vous avez cherché les images? (Par exemple, recherche par mots clés, par artiste, etc.)
5. Quelles sont les métadonnées qui vous ont aidé à sélectionner l'image?

Appendice C : Certificat d'approbation déontologique

Numéro de dossier: 03-12-22

Date (mm/jj/aaaa): 04/30/2012



Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Certificat d'approbation déontologique

CÉR Sciences sociales et humanités

Chercheur principal / Superviseur / Co-chercheur(s) / Étudiant(s)

<u>Prénom</u>	<u>Nom de famille</u>	<u>Affiliation</u>	<u>Rôle</u>
Mary F.	Cavanagh	Arts / Autres	Superviseur
Tom	Delsey	Arts / Autres	Co-Superviseur
Mariane	Léonard	Arts / Autres	Étudiant-chercheur

Numéro du dossier: 03-12-22

Type du projet: Thèse de maîtrise

Titre: L'Utilisation des métadonnées des usagers experts lors de la recherche d'images

Date d'approbation (mm/jj/aaaa) Date d'expiration (mm/jj/aaaa) Approbation

04/30/2012

04/29/2013

I

(Ia: Approbation complète, Ib: Autorisation préliminaire de libération de fonds de recherche)

Conditions Spéciales / Commentaires:

Approbation : Le recrutement et la collecte de données peuvent commencer à l'Université d'Ottawa.

Approbation conditionnelle:

- Veuillez soumettre la copie d'approbation éthique de l'Université de Montréal.
- Veuillez soumettre la correspondance de l'Université du Québec à Montréal soutenant que l'approbation éthique de l'Université d'Ottawa suffira afin de recruter les membres de leur institution.
- Veuillez noter que l'approbation déontologique complète sera accordée seulement lorsque ces autorisations seront reçues.



Université d'Ottawa **University of Ottawa**
 Bureau d'éthique et d'intégrité de la recherche Office of Research Ethics and Integrity

La présente confirme que le Comité d'éthique de la recherche (CER) de l'Université d'Ottawa identifié ci-dessus, opérant conformément à l'Énoncé de politique des Trois conseils et toutes autres lois et tous règlements applicables de l'Ontario, a examiné et approuvé la demande d'approbation déontologique du projet de recherche ci-nommé. L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales / Commentaires".

Lors de l'étude, le protocole ne peut être modifié sans approbation préalable écrite du CER sauf si le sujet doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques de l'étude comme par exemple un changement de numéro de téléphone. Les chercheurs doivent aviser le CER dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou affecter considérablement le déroulement du projet. Ils devront aussi rapporter tout événement imprévu et / ou dommageable et devront soumettre toutes les nouvelles informations pouvant nuire à la conduite du projet et/ou à la sécurité des participants. Toutes modifications apportées au projet, aux lettres d'information / formulaires de consentement ainsi qu'aux documents de recrutement doivent être soumises pour approbation à ce Service en utilisant le document intitulé "Modification au projet de recherche" au: http://www.ssr.d.uottawa.ca/deontologie/application_dwn_f.asp.

Veillez soumettre un rapport annuel au Responsable de la déontologie en recherche, quatre semaines avant la date d'échéance indiquée afin de fermer le dossier ou demander un renouvellement de l'approbation déontologique. Le document nécessaire est disponible en ligne au: http://www.rges.uottawa.ca/ethics/application_dwn.asp.

Pour toutes questions, vous pouvez communiquer avec le bureau de déontologie en composant le poste 5387 ou en nous contactant par courriel à: ethics@uOttawa.ca.

Signature:



Kim Thompson

Responsable de la déontologie en recherche
 Pour Barbara Graves, Présidente du CER en Sciences sociales et humanités

Appendice D : Courriel d'invitation à participer

OBJET : INVITATION À PARTICIPER

Madame,

Monsieur,

Je suis étudiante à la maîtrise en sciences de l'information de l'Université d'Ottawa. Je fais un mémoire sur l'utilisation des métadonnées lors de la recherche d'images numériques par les chercheurs professionnels. Je suis à la recherche de professionnels qui voudraient participer à mon étude et par conséquent je me demandais si certains bibliothécaires ou archivistes souhaiteraient collaborer à ce genre de projet. Je vous envoie la description de ma recherche.

Vous trouverez ci-joint une description détaillée du projet de recherche. N'hésitez pas à communiquer avec moi pour tout renseignement complémentaire.

En espérant qu'il vous sera possible de donner suite à ma demande, je vous prie d'agréer, Madame, Monsieur, mes salutations distinguées.

Appendice E : Description du projet

Recherche d'images numériques et utilisation des métadonnées chez des chercheurs expérimentés.

Mariane Léonard
Étudiante à la maîtrise
École des sciences de l'information
Université d'Ottawa

Le but de ce projet est de comprendre l'usage que font les chercheurs expérimentés des métadonnées associées aux images numériques lorsqu'ils en font la recherche dans le cadre d'un accomplissement d'un besoin ou d'une demande. Dans ce cas-ci, la population cible est les bibliothécaires ou les archivistes qui font de la recherche d'images numériques dans le but de combler un besoin spécifique, par exemple, trouver une image demandée par un usager. Dans le dessein de combler ce besoin, les bibliothécaires et archivistes utilisent des outils de recherche et par le fait même ils entreprennent quatre tâches : trouver, identifier, sélectionner, obtenir. Cette recherche tentera de comprendre comment les métadonnées associées aux images interviennent dans ces tâches. Quel est le lien entre l'accomplissement du besoin et l'utilisation des métadonnées? Par l'observation de l'usage des métadonnées que font ces chercheurs, il sera possible de comprendre mieux l'utilisation qu'ils font des métadonnées liées aux images.

Le projet de recherche s'attentionnera sur l'utilisation que font les professionnels des données qu'ils rencontrent dans leurs recherches d'images, que ce soit dans les collections

virtuelles de bibliothèques ou sur le Web. Les professionnels devront, alors qu'ils font de la recherche, remplir un questionnaire dans lequel ils inscriront des données, par exemple la base de données visitée, la décision de la sélection d'images, la raison de ce choix. Ils devront écrire un maximum de dix questionnaires et un minimum de six. Deux entrevues de 45-60 minutes se feront individuellement, une au milieu du projet et une autre à la fin. La recherche s'étalera sur une période d'un mois, pendant laquelle les données seront compilées.

La participation des professionnels est indispensable, car c'est leur perception qui est au centre de cette étude. Cette recherche permettra d'avoir une vision plus juste de l'utilisation des métadonnées par les chercheurs expérimentés. Ceci permettra de comprendre mieux ce sujet et par conséquent, de parfaire la présentation des données associées aux images numériques. De plus, il sera possible de repérer si certaines informations indisponibles dans les schémas actuels pourraient être ajoutées dans le but d'améliorer la recherche d'image.

Les données qui seront récoltées dans cette recherche seront totalement anonymes, en aucun cas le nom des participants ne sera cité et il sera impossible de les reconnaître. Les données seront également conservées sur un disque dur crypté conservé à l'École de sciences de l'information à l'Université d'Ottawa, donc aucune personne extérieure au projet ne pourra accéder aux résultats de la recherche et aux notes d'entrevues. Seulement la chercheuse principale Mariane Léonard, la directrice D^{re} Mary Cavanagh et le codirecteur le D^r Tom Delsey auront accès aux données. Le tout sera fait selon la réglementation du comité

d'éthique de l'Université d'Ottawa. D'ailleurs, le projet sera supervisé par les D^{re} Mary Cavanagh et D^r Tom Delsey de l'École des sciences de l'information de l'Université d'Ottawa.

Pour plus d'information, veuillez contacter :

Mariane Léonard



[@uottawa.ca](mailto:mariane.leonard@uottawa.ca)

École des sciences de l'information

Pavillon Desmarais

55, av. Laurier Est,

11^e étage

Ottawa ON Canada

K1N 6N5

Appendice F : Formulaire de consentement

Formulaire de consentement

Recherche d'images numériques et utilisation des métadonnées chez des chercheurs expérimentés.

Chercheuse principale :
Mariane Léonard
École des sciences de l'information
Université d'Ottawa
[redacted]
[redacted]@uottawa.ca

Directrice
D^{re} Mary Cavanagh
École des sciences de l'information
Université d'Ottawa
[redacted]
[redacted]@uOttawa.ca

Codirecteur
D^r Tom Delsey
École des sciences de l'information
Université d'Ottawa
[redacted]
[redacted]@uOttawa.ca

Invitation à participer : Je suis invité(e) à participer à la recherche nommée ci-haut qui est menée par Mariane Léonard, les D^{re} Mary Cavanagh et D^r Tom Delsey.

But de l'étude : Le but de l'étude est d'étudier l'utilisation des métadonnées que font les chercheurs expérimentés lors de la recherche d'images numériques.

Participation : Ma participation consistera essentiellement à tenir un journal de bord lors de la recherche d'images, et ce pour un maximum de 10 entrées et un minimum de 6. Je devrai remplir ce journal de bord dans une période d'un mois. De plus, je devrai participer à deux entrevues en français avec la chercheuse principale, Mariane Léonard. Ces rencontres seront

d'une durée maximum d'une heure chacune. La première entrevue sera au milieu du projet, tandis que la dernière sera à la fin. Les questions d'entrevues seront à propos de la recherche d'images que les participants auront faite ainsi que des journaux qui auront été remplis. Les entrevues seront enregistrées sur une piste audio. Les entrevues pourront se faire en personne.

Bienfaits : Ma participation à cette recherche aura pour effet de permettre une compréhension de l'utilisation qu'ont les chercheurs expérimentés des métadonnées liées les images numériques. Ceci permettra de comprendre mieux ce sujet et par conséquent, d'améliorer la présentation des données.

Confidentialité et anonymat : J'ai l'assurance de la chercheuse principale que l'information que je partagerai avec elle restera strictement confidentielle. Je m'attends à ce que le contenu ne soit utilisé que pour la rédaction de la thèse et selon le respect de la confidentialité. L'anonymat est garanti de la façon suivante : lorsque je serai cité(e) et/ou que je serai mentionné(e) dans la thèse, je serai nommé(e) par un pseudonyme et toutes informations pouvant permettre de me reconnaître seront omises.

Conservation des données : Les données seront recueillies sur papier en ce qui concerne le journal de bord et sur support audio pour les entrevues. Les transcriptions audios ainsi que les formulaires papier seront conservés de façon sécuritaire dans un ordinateur muni d'un mot de passe dont seul Mariane Léonard le connaît. De plus, une copie des données sera conservée dans le bureau de la D^{re} Mary Cavanagh à l'École des Sciences de l'information

sur un disque cryptée. À la fin de la période de rétention, 5 ans après le dépôt de la thèse, tous les documents seront détruits à l'Université d'Ottawa.

Participation volontaire : Ma participation à la recherche est volontaire et je suis libre de me retirer en tout temps, et/ou refuser de répondre à certaines questions, sans subir de conséquences négatives. Si je choisis de me retirer de l'étude, les données recueillies jusqu'à ce moment seront détruites.

Acceptation : Je, (_____), accepte de participer à cette recherche menée par Mariane Léonard de l'École des sciences de l'information de l'Université d'Ottawa, laquelle recherche est supervisée par D^{re} Mary Cavanagh et D^r Tom Delsey.

Pour tout renseignement additionnel concernant cette étude, je peux communiquer avec le chercheur ou ses superviseurs.

Pour tout renseignement sur les aspects éthiques de cette recherche, je peux m'adresser au Responsable de l'éthique en recherche, Université d'Ottawa, Pavillon Tabaret, 550, rue Cumberland, pièce 154, (613) 562-5387 ou ethics@uottawa.ca.

Il y a deux copies du formulaire de consentement, dont une copie que je peux garder.

Signature du participant :

Date :

Signature du chercheur :

Date :

Appendice G : Liste des codes

Code 4.2.1.1

Codes	Définitions	Exemples
Type/Technique	« describes the general type of image or video sequence or the technique used to produce it (e.g., black and white, color) »	Photographie, Peinture, Affiche
Global Distribution	« classifies images or video sequences based on their global content, and is measured in terms of low-level perceptual features such as spectral sensitivity (color), and frequency sensitivity (texture) »	
Local Structure	« is concerned with the extraction and characterization of the individual components of the image. At the most basic level, those components result from low-level processing and include elements such as the Dot, Line, Tone, Color, and Texture »	
Global Composition	« the focus is on the specific arrangement or spatial layout of elements in the image. In other words, we analyze the image as a whole, but only use the basic elements described in the previous level (e.g., line and circle) for the analysis. Traditional analysis in art describes composition concepts such as balance, symmetry, center of interest (center of attention or focus), leading line, and viewing angle »	
Generic Objects	« the basic level categories: the most general level of object description, which can be recognized with everyday knowledge (e.g., apple, man, chair) »	Canot,
Generic Scene	« refers to scenes at their most general level of description. Examples of scene classes include city, landscape, indoor, outdoor, still life, and portrait. It is not necessary to know a specific street or building name to determine that it is a city scene, nor is it necessary to know the name of an individual to know that the individual is a woman. »	
Specific Object	« refers to identified and named objects. Specific knowledge of the objects in the image or the video sequence is required, and such knowledge is usually objective because it relies on known facts. Examples include individual persons or objects (e.g., Bill Clinton, Eiffel Tower) »	
Specific Scene	« refers to attributes that carry specific knowledge about the scene (e.g., Paris, Times Square, Central Park, etc.). »	

Codes	Définitions	Exemples
Abstract Object	"refer to specialized or interpretative knowledge about what the objects represent. This indexing level is the most difficult one in the sense that it is completely subjective, and assessments between different users may vary greatly. For example, a woman in a picture may represent anger to one observer and pensiveness to another »	
Abstract Scene	"refers to what the image as a whole represents, and may be very subjective Jörgensen (1995b) showed that users sometimes describe images in affective (e.g., emotion) or abstract (e.g., atmosphere, theme) terms. Examples at the abstract scene level include sadness, happiness, power, heaven, and paradise. »	
Jörgensen, C., Jaimes, A., Benitez, A. B., & Chang, S. (2001). A Conceptual Framework and Empirical Research for Classifying Visual Descriptors. <i>Journal of the American Society for Information Science and Technology</i> , 52(11), 938–947.		

Code 4.2.1.2

Codes	Définitions
Other Autres	Autres mots totalement différents, par exemple le terme en anglais
Related Terme associé	« Terme possédant une proximité sémantique avec d'autres termes, en excluant les relations hiérarchiques et d'équivalence » (Dégez & Ménillet, 2001, p. 42)
Same Même	Le même mot
Synonym Synonyme	« Caractère de deux ou plusieurs termes de formes différentes et de même sens ou de sens très voisins » (Dégez & Ménillet, 2001, p. 41)
Variant Variance	Le même mot, mais écrit de façon différente

Code 4.2.1.2 – 4.2.2.2 - 4.2.3.1

Metadata	Définitions
Creator	« An entity primarily responsible for making the resource. » (Dublin Core)
Title	« A name given to the resource. » (Dublin Core)

Metadata	Définitions
Date	« A point or period of time associated with an event in the lifecycle of the resource. »(Dublin Core)
Description	« An account of the resource. »(Dublin Core)
Subject	« Terms or phrases that describe, identify, or interpret the Work or Image and what it depicts or expresses. These may include generic terms that describe the work and the elements that it comprises, terms that identify particular people, geographic places, narrative and iconographic themes, or terms that refer to broader concepts or interpretations. » (VRA Core)
Tags, keywords	« Terme en langage naturel, choisi généralment dans le titre ou le texte d'un document pour en caractriser le contenu et en permettre la recherche. Il est à disintiguer d'un descripteur, qui est un terme normalisé dans un thésaurus » (Dégez & Ménillet, 2001, p. 31).
Source	« A related resource from which the described resource is derived. »(Dublin Core)
Rights	« Information about rights held in and over the resource. »(Dublin Core)
Location	« The geographic location and/or name of the repository, building, site, or other entity whose boundaries include the Work or Image. Inclusion of a type attribute distinguishes between different kinds of locations, e.g. repository locations, creation locations, discovery locations, etc" (VRA Core)
Technique	« The production or manufacturing processes, techniques, and methods incorporated in the fabrication or alteration of the work or image. » (VRA Core)
Material	"The substance of which a work or an image is composed" VRA core)
Measurements	« The physical size, shape, scale, dimensions, or format of the Work or Image. Dimensions may include such measurements as volume, weight, area or running time. » (VRA Core)
Image size	Fait référence à la dimension de l'image numérique.
Cultural Context	« The name of the culture, people (ethnonym), or adjectival form of a country name from which a Work, Collection, or Image originates, or the cultural context with which the Work, Collection, or Image has been associated. » (VRA Core)
Style Period	« A defined style, historical period, group, school, dynasty, movement, etc. whose characteristics are represented in the Work or Image. Cultural and regional terms may be combined with style and period terms for display purposes. » (VRA Core)

Code 4.2.1.3

Codes	Définitions
Négatif	La recherche apporte à des négatifs non numérisés
Google	Le participant fait une recherche sur Google pour obtenir plus d'information
Recherche avec facette	Le participant utilise les facettes pour faire la recherche
Livres	Le participant fait une recherche dans un livre pour trouver plus d'information
Recherche information	
Problème technique	Le participant rencontre un problème technique pendant la recherche
Contacteur participant	Le participant contacte l'utilisateur pour lui demander plus d'informations
Recherche simple	Le participant fait une recherche simple dans la base de données

Code 4.2.1.4

Codes	Définitions
1 image	
Aucune	
2-5 images	
10 ou plus	

Code 4.2.4.1

Codes	Définitions
Négatif	Image non numérisée
Site Web	Obtenu en ligne
Book	Image non numérisée
Serveur interne	Image en ligne, mais sur un serveur interne

Appendice H : Exemple de code

Trouver	
Quelle est la requête que vous avez reçue?	I am in possession of a circa 1800's portrait that I would like some info on or if you can direct me to whom I should be consulting. I have researched on the internet and found that the National Gallery of London UK has a little information, but they have directed me to obtain from within Canada. The portrait is of the Duchess of Beaufort, Emily Frances Somerset (Smith), drawn by Chelan, engraved by WH Mote, printed by McQueen. Any help would be greatly appreciated
Quel(s) mot(s) clé(s) vous avez retenu pour faire la recherche ?	Chelan artist Mote engraving McQueen printing A.E. Chalon, WH Mote, McQueen Duchess of Beaufort, Emily Frances Somerset.
Comment avez-vous fait la recherche ? (Courte description)	In order to identify the individuals I did a quick google search and found information on Mote but nothing on Chelan. I then referred to the Dictionary of British Artists, still without any luck. So then I contacted the individual again. It turns out that they had spelled the artist's name incorrectly- it was actually A.E Chalon. Furthermore this individual was trying to ascertain whether or not they had a valuable print. So then I searched the names in combination and found similar prints in an assortment of collections in the UK including the collections of the National Portrait Gallery, the British Museum, and the Royal Academy of Arts.

Legend for the table content:

- date
- circa
- Artist type/technique
- Object/Genre/Scene
- Creator
- Zee-specific object
- Technique

Appendice I : Classification pyramide des mots clés

Indexing template (group & attribute type)	Indexing template (attribute)	Explanation/example	Pyramid level
"EXTERNAL" INFO			
	> Image ID		NA
	> Creator/Author		NA
	> Title		NA
	> Publisher		NA
	> Date		NA
	> Image Type	color X-ray graphics	TYPE/TECHNIQUE
	> Access Conditions		NA
	> Technical	resolution, file size, etc.	NA
"INFERRED" INFO			
> "Environment"	⊗ When—time <i>in image</i>	Middle Ages, summer	GENERIC SCENE
	⊗ Where—General	city, rural, indoor, office	GENERIC SCENE
	⊗ Where—Specific	Paris, Chrysler Building	SPECIFIC SCENE
> Subject/Topic		overall subject/theme	ABSTRACT OBJECT/SCENE
> Medium		oil, watercolor, digital	TYPE/TECHNIQUE
> Symbolism		Garden of Eden, afterlife	ABSTRACT OBJECT/SCENE
> "Why"	⊗ Emotions/Mental	sadness, laughter	ABSTRACT OBJECT/SCENE
	⊗ Relationships	brothers, romance	ABSTRACT OBJECT/SCENE
> "Miscellaneous"	⊗ Point of	bird's-eye, close-up	GLOBAL COMPOSITION
	⊗ Style	abstract, realism, etc	ABSTRACT SCENE
	⊗ Genre	landscape, portrait	GENERIC SCENE
	⊗ Atmosphere/overall	gloomy, mysterious	ABSTRACT OBJECT/SCENE
VISUAL ELEMENTS			
> Color	⊗ Color	red, blue	GLOBAL DIST/LOCAL STRC
	⊗ Color Quality	dark, bright	GLOBAL DIST/LOCAL STRC
	⊗ Placement	center, overall	LOCAL STRUCTURE
> Shape	⊗ Placement	square, elongated, curved	GLOBAL DIST/LOCAL STRC
> Texture	⊗ Placement	smooth, shiny, fuzzy	GLOBAL DIST/LOCAL STRC
LITERAL OBJECTS			
> Category—General		What group: tool	GENERIC/SPECIFIC
> Type—Specific	⊗ Placement	What it is—hammer	GENERIC/SPECIFIC
	⊗ Shape		
	⊗ Texture		
	⊗ Size		
	⊗ Number		
	⊗ Color		
LIVING THINGS			
> Type		human or what animal	GENERIC OBJECT
	⊗ Placement		
	⊗ Size	large, very small	
	⊗ Gender	male, female	SPECIFIC OBJECTS
	⊗ Age		SPECIFIC OBJECTS
	⊗ Number		
	⊗ Pose	seated, standing, lying	GENERIC/SPECIFIC SCENE
	⊗ Name	Ghandi	SPECIFIC OBJECT/SCENE
	⊗ Physical	running, talking, football	GENERIC/SPECIFIC SCENE
	⊗ Status	occupation, social, status	ABSTRACT OBJECT/SCENE
COLLATERAL INFO			
> Caption			
> Related Text			
> Voice Annotations			

(Jørgensen et al., 2001, p. 947)