

Université de Montréal

**Analyse documentaire en milieu universitaire : deux approches  
générales comparées**

par

Francis Hébert

École de bibliothéconomie et des sciences de l'information  
Faculté des arts et sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales  
en vue de l'obtention du grade de maître  
en Sciences de l'information

Octobre 2014

© Francis Hébert, 2014

Université de Montréal  
Faculté des études supérieures et postdoctorales

Ce mémoire intitulé :

Analyse documentaire en milieu universitaire : deux approches générales comparées

Présenté par :  
Francis Hébert

Évalué par un jury composé des personnes suivantes :

Michèle Hudon (Présidente-rapporteuse)  
Dominic Forest (Co-directeur)  
Lyne Da Sylva (Co-directrice)  
Bruno Bachimont (Membre du jury)

## Résumé

Ce mémoire porte sur l'analyse documentaire en milieu universitaire. Deux approches générales sont d'abord étudiées : l'approche centrée sur le document (premier chapitre), prédominante dans la tradition bibliothéconomique, et l'approche centrée sur l'utilisateur (deuxième chapitre), influencée par le développement d'outils le plus souvent associés au Web 2.0. L'opposition entre ces deux démarches reflète une dichotomie qui se trouve au cœur de la notion de sujet, c'est-à-dire les dimensions objective et subjective du sujet. Ce mémoire prend par conséquent la forme d'une dissertation dont l'avantage principal est de considérer à la fois d'importants acquis qui appartiennent à la tradition bibliothéconomique, à la fois des développements plus récents ayant un impact important sur l'évolution de l'analyse documentaire en milieu universitaire. Notre hypothèse est que ces deux tendances générales doivent être mises en relief afin d'approfondir la problématique de l'appariement, laquelle définit la difficulté d'accorder le vocabulaire qu'utilise l'utilisateur dans ses recherches documentaires avec celui issu de l'analyse documentaire (métadonnées sujet). Dans le troisième chapitre, nous examinons certaines particularités liées à l'utilisation de la documentation en milieu universitaire dans le but de repérer certaines possibilités et certaines exigences de l'analyse documentaire dans un tel milieu. À partir d'éléments basés sur l'analyse des domaines d'études et sur la démarche analytico-synthétique, il s'agit d'accentuer l'interaction potentielle entre usagers et analystes documentaires sur le plan du vocabulaire utilisé de part et d'autre.

**Mots-clés** : Analyse documentaire ; Indexation ; Classification ; Langage documentaire ; Folksonomie; Université.

## **Abstract**

The topic of this dissertation is subject analysis in a university environment. Two major approaches are studied at first: subject analysis centered on the document (first chapter), historically predominant in librarianship, and subject analysis centered on the user (second chapter), mostly influenced by the development of Web 2.0 technologies. The opposition between those two approaches reflects a dichotomy which is at the very heart of the notion of subject, meaning the objective and subjective aspects of the subject. The outline of the dissertation has the distinct advantage of presenting well established practices in the field of librarianship as well as recent developments that do have an impact on subject analysis in a university environment. Our hypothesis is that both major tendencies must be highlighted to study the question of mapping the terminology (subject metadata) that comes from subject analysis with the terminology that users tend to favor while searching for documents. In the third chapter, we examine more closely particularities of the university environment in an effort to look at distinct possibilities and requirements for subject analysis in such an environment. Reinforced by elements taken from domain and facet analysis, the goal is to accentuate the potential interaction between users and indexers on a terminological level.

**Keywords:** Subject analysis; Indexing; Classification; Controlled vocabulary; Folksonomy; University.

# Table des matières

<b>Résumé</b> .....	<b>i</b>
<b>Abstract</b> .....	<b>ii</b>
<b>Liste des sigles</b> .....	<b>v</b>
<b>Remerciements</b> .....	<b>vii</b>
<b>Introduction</b> .....	<b>1</b>
<b>Premier chapitre</b> .....	<b>8</b>
Approche centrée sur le document.....	8
1.1) La dimension objective du sujet et l' <i>aboutness</i> .....	9
1.2) L'analyse documentaire en tant qu'exercice pragmatiste.....	20
1.3) La réduction sémantique.....	30
<b>Deuxième chapitre</b> .....	<b>41</b>
Approche centrée sur l'utilisateur .....	41
2.1) Le relativisme dans le contexte de l'analyse documentaire en milieu universitaire .	44
2.1.1) La connaissance dans une optique relativiste.....	45
2.1.2) <i>Knowledge claims</i> .....	51
2.1.3) Les énoncés de recherche .....	55
2.2) La dimension subjective du sujet.....	59
2.2.1) La lecture en milieu universitaire .....	61
2.2.2) L'environnement de recherche documentaire .....	65
2.2.3) La caution usagère .....	69
2.3) La <i>folksonomie</i> .....	72
2.3.1) La problématique de l'appariement et la dimension subjective du sujet.....	73
2.3.2) La <i>folksonomie intégrée</i> .....	75
<b>Troisième chapitre</b> .....	<b>83</b>
Particularités de l'analyse des contenus en milieu universitaire.....	83
3.1) Analyser les domaines d'études .....	84
3.2) Favoriser une démarche analytico-synthétique .....	94
<b>Conclusion</b> .....	<b>104</b>
<b>Bibliographie</b> .....	<b>107</b>

<b>Annexe A : Glossaire .....</b>	<b>i</b>
<b>Annexe B : Exemple d'énoncés de recherche .....</b>	<b>v</b>

## Liste des sigles

**CC** : *Colon Classification* (classification à facettes développée par Ranganathan)

**DDC** : *Dewey Decimal Classification* (classification décimale de Dewey).

**LCC** : *Library of Congress Classification* (classification de la Bibliothèque du Congrès).

**LCSH** : *Library of Congress Subject Headings* (vedettes-matières de la Bibliothèque du Congrès).

**OPAC** : *Online Public Access Catalog* (catalogue de bibliothèque accessible en ligne)

**RVM** : Répertoire des vedettes-matières de l'Université Laval.

**UDC** : *Universal Decimal Classification* (classification décimale universelle).

*Ce mémoire est dédié à mes proches.*

## **Remerciements**

Je voudrais d'abord remercier Dominic Forest et Lyne Da Sylva pour leur soutien, leur patience et leurs bons conseils. Ce mémoire est le résultat d'un travail de longue haleine (c'est le moins qu'on puisse dire) et d'un questionnement constant qui n'aurait pas été possible sans le regard objectif (et parfois critique) de mes codirecteurs. Je voudrais aussi remercier tous les membres de l'EBSI, aussi bien le corps professoral que le personnel administratif. J'ai vécu une expérience universitaire très enrichissante en votre compagnie. Finalement, je voudrais saluer mes collègues de la BIC, avec qui j'ai aussi beaucoup appris et pour qui j'ai beaucoup d'estime.

## Introduction

Dans ce mémoire, nous nous proposons d'examiner l'analyse documentaire en tant qu'activité intellectuelle (effectuée par un analyste documentaire). Avant d'introduire l'objectif du mémoire et les différentes parties qui le composent, regardons la définition qu'Hudon donne de l'analyse documentaire (*subject analysis*) : « Opération dont le double objectif est de déterminer les éléments d'information pertinents dans un document ou une collection de documents et de les représenter sous une forme concise et non ambiguë pour en faciliter le repérage ultérieur » (Hudon 2013, 264). Notre objectif est de présenter deux grandes approches (centrée sur le document versus centrée sur l'utilisateur) de manière à souligner les défis que rencontre l'analyse *humaine* de la documentation pour répondre aux attentes des usagers, mais aussi les possibilités qu'elle offre pour améliorer les options de recherche documentaire.

Plus spécifiquement, nous proposons une réflexion sur le défi incontournable qui consiste à établir une correspondance claire et pertinente entre le vocabulaire auquel ont recours les analystes documentaires et celui privilégié par les usagers pour représenter un même sujet. Il s'agit d'une question vaste et difficile qui accompagne d'un bout à l'autre le développement moderne de l'analyse documentaire (des règles de catalogage de Cutter au catalogue 2.0). C'est pour cette raison que nous avons choisi de rédiger une dissertation *classique* (ce qui n'est pas sans difficultés de lecture, nous en convenons) plutôt que de proposer la mise en application d'une approche plus précise, mais nécessairement incomplète. Notre but avoué est de fixer deux tendances incontournables (approche « document » / approche « usager ») en regard desquelles diverses approches particulières pourront éventuellement être évaluées. Si le lecteur veut bien nous suivre dans cet exercice, il trouvera matière à réflexion et multiples occasions de se questionner sur les défis et les exigences propres à l'analyse documentaire.

L'analyse documentaire, comme nous le savons, vise entre autres à régler des problèmes de synonymie et de polysémie nuisibles dans un « système documentaire »<sup>1</sup>. L'outil principal pour y parvenir, le langage documentaire, résulte d'un effort de normalisation dont l'une des conséquences est un appauvrissement langagier :

Quelle que soit la solution adoptée, le danger de la normalisation est qu'elle réduit inévitablement les nuances exprimées par les langues naturelles (c'est son but) et qu'elle risque, en traitant les termes les plus spécifiques comme des quasi-synonymes, d'appauvrir la capacité sémantique d'un langage et d'aboutir à une nouvelle forme de silence, dû à l'absence de concepts spécifiques. (Maniez 2002, 177-178)

Cette stratégie de normalisation peut avoir un impact sur la recherche documentaire. La problématique étudiée dans ce mémoire, inspirée de Maniez, émerge d'un désaccord potentiel entre le langage privilégié lors de l'analyse documentaire et celui utilisé lors de la recherche documentaire. En d'autres mots, les moyens pour formaliser l'expression des contenus de la documentation présentent des difficultés qui, selon Maniez, engendrent un problème fondamental :

De ces difficultés techniques de normalisation formelle découle un problème pratique de communication tout aussi fondamental : comment harmoniser l'expression du même sujet par l'indexeur – professionnel du système documentaire – et par l'utilisateur occasionnel du système documentaire? (Maniez 2002, 178)

C'est ce que nous appellerons tout au long du mémoire la « problématique de l'appariement » (Maniez 2002, 178-179), laquelle souligne, dans une perspective générale, la difficulté d'harmoniser le vocabulaire utilisé par l'utilisateur et celui que privilégie habituellement l'analyste documentaire dans son travail<sup>2</sup>.

---

<sup>1</sup> Nous reprenons ici une expression utilisée par Maniez, bien que d'autres termes tels que « système d'information » soient parfois privilégiés. Nous reviendrons sur le choix terminologique que nous avons fait à la fin de cette introduction.

<sup>2</sup> Maniez place l'appariement au cœur non seulement de toute recherche documentaire, mais de tout système de recherche (dans un sens très large) : « L'appariement est au cœur de tous les systèmes de recherche, même si le processus qui y conduit prend des formes et des délais variés. Sans cette reconnaissance de l'identité, qu'elle soit le fait d'une perception humaine ou d'une machine, toute sélection est impossible, puisque le tri suppose qu'on distingue ce qui est identique et ce qui est différent » (Maniez 2002, 25).

Dans une perspective plus technique, Hjørland nous fournit l'occasion d'examiner plusieurs activités, procédés et outils qui donnent son orientation à l'organisation des connaissances, laquelle englobe l'analyse documentaire dans son acception la plus courante :

In the narrow meaning Knowledge Organization (KO) is about activities such as document description, indexing and classification performed in libraries, bibliographical databases, archives and other kinds of "memory institutions" by librarians, archivists, information specialists, subject specialists, as well as by computer algorithms and laymen. KO [*Knowledge Organization*] as a field of study is concerned with the nature and quality of such knowledge organization processes (KOP) as well as the knowledge organization systems (KOS) used to organize documents, document representations, works and concepts. Library and Information Science (LIS) is the central discipline of KO in this narrow sense (although seriously challenged by, among other fields, computer science). » (Hjørland 2008a, 86)

Lorsqu'il rattache et même subordonne les sciences de l'information (*Library and Information Science*) à l'organisation des connaissances (*Knowledge Organization*), Hjørland se réfère avant tout à des tâches telles que le traitement et l'analyse documentaires (*document description, indexing and classification*). La bibliothéconomie et, de manière plus générale, les sciences de l'information comptent bon nombre d'activités, de savoirs et d'objectifs qui dépassent les limites de l'organisation des connaissances. Parallèlement, l'organisation des connaissances, entendue de manière plus large que ce premier sens que lui donne Hjørland, rassemble une panoplie de disciplines, outils et tendances, au sein desquels l'automatisation joue un rôle de plus en plus important. Néanmoins, le croisement est significatif puisque ces deux domaines (organisation des connaissances et sciences de l'information) ont contribué et continue de contribuer à leurs développements respectifs.

Dans sa thèse de 2002, intitulée *Subject determination during the cataloging process*, Šauperl résume dans une proposition synthétique la contribution du catalogage et, plus précisément, de l'analyse documentaire (*subject cataloging*) dans le secteur de la bibliothéconomie :

The main purpose of cataloging is to help people find information about library materials through library catalogs. Subject cataloging, a part of the cataloging process, focuses on providing information on the topic of the

document, that is, what the document is about. This subject representation is a critical element in providing access to materials. (Šauperl 2002, 1)

L'analyse documentaire a donc un rôle à jouer dans la transmission de l'information en facilitant l'accès aux documents. C'est un objectif pratique que l'analyste documentaire doit toujours garder en tête, surtout dans un contexte où les sources d'information se multiplient. Plusieurs outils sont maintenant proposés pour permettre à chacun de trouver l'information qu'il cherche et d'ainsi parfaire ses connaissances dans une discipline ou sur un sujet donné. Si bien qu'on est en droit de se demander si le catalogue de type OPAC occupe toujours une place importante dans les stratégies de recherche des usagers. Par extension, il faut aussi se questionner sur les méthodes utilisées dans le domaine de l'analyse documentaire, sur leur validité actuelle et leur développement futur.

Comme nous l'avons souligné d'entrée de jeu, nous allons nous intéresser à l'analyse documentaire *intellectuelle* (effectuée par un analyste documentaire<sup>3</sup>) et, plus précisément, aux deux activités les plus répandues dans cette discipline : l'indexation et la classification<sup>4</sup>. Ces dernières constituent le cœur de l'analyse documentaire, aussi bien si l'on considère la tradition que si l'on regarde les pratiques actuelles. Tradition et actualité seront d'ailleurs mis en parallèle fréquemment dans le but d'établir une forme de continuité, mais, comme le remarque à juste titre Breeding, il s'agit d'un défi important dans la conjoncture actuelle : « It's quite a challenge to provide a new approach for the growing ranks of Websavvy users while preserving something familiar to other library users who rely on the traditional online catalog » (Breeding 2010, 10). Par ailleurs, il faut mentionner que le point de vue que nous voulons adopter n'est ni celui d'un manuel, ni celui d'un article scientifique : nous voulons examiner l'analyse documentaire, son fondement intellectuel et ses possibilités, de manière à engendrer une réflexion suffisamment large pour en décrire le fonctionnement général et assez précise pour donner une image juste de ce qui en constitue le moteur, la notion de sujet. Une forme d'indétermination caractérise certains concepts évoqués dans ce mémoire, ce qui a été

---

<sup>3</sup> Il ne sera pas question, dans ce mémoire, de techniques automatisées d'analyse des contenus.

<sup>4</sup> La condensation n'est pas à négliger, son utilité étant évidente par exemple dans le cas de l'analyse documentaire des articles scientifiques. Néanmoins, il faut opérer certains choix pour circonscrire la réflexion dans un travail dont les limites sont préétablies. Nous considérons, par ailleurs, qu'elle est trop complexe sur le plan formel et, par la même occasion, trop peu répandue dans les services techniques des bibliothèques pour mériter sa place dans ce mémoire.

vu comme l'occasion de questionner les bases de l'analyse documentaire en tant qu'exercice intellectuel. Plusieurs questions de vocabulaire seront abordées tout au long du mémoire, ainsi que certains aspects sémantiques et épistémologiques de l'analyse documentaire. Un glossaire auquel pourra se rapporter le lecteur est fourni dans l'annexe A.

Il est important de préciser aussi que le milieu universitaire occupera une place non négligeable dans cette dissertation. Cela ne veut pas dire que l'analyse documentaire sera soumise à une conception *définie par* ce milieu en particulier, puisque certains principes incontournables sont d'un ordre plus général. Une réflexion sur la bibliothèque universitaire sera mise à contribution surtout dans le deuxième chapitre du mémoire, puisque celui-ci concerne une approche de l'analyse documentaire centrée sur l'utilisateur. Notre hypothèse, concernant cette limitation relative (certains développements pourraient sûrement être repris pour le bénéfice de toute institution documentaire, avec quelques modifications), repose sur le potentiel intéressant que représente le milieu universitaire quant au caractère convergent des activités des usagers. Les objectifs, les habitudes et les critères de recherche peuvent différer bien évidemment, mais le recours à la documentation pour élucider un énoncé de recherche et la lecture critique semblent deux éléments sur lesquels les bibliothécaires peuvent tabler pour cerner, au moins en partie, les besoins de leurs usagers.

Nombre d'articles et de monographies ont façonné le propos développé dans ce mémoire, mais deux textes ont permis d'en orienter plus précisément le contenu et la forme. En premier lieu, le livre *Actualité des langages documentaires : les fondements théoriques de la recherche d'information*, écrit par Jacques Maniez (cité plus haut), nous a permis de définir une problématique dont les ramifications sollicitent l'approfondissement de la notion de sujet. En second lieu, nous nous sommes inspiré d'un article de Jens-Erik Mai (2005), « Analysis in indexing : document and domain centered approaches », pour développer les différents chapitres du mémoire et accorder son contenu avec la forme classique d'une dissertation.

La forme du mémoire s'inspire donc d'un article de Mai (2005) portant sur l'approche centrée sur le document et celle centrée sur la discipline (ou le domaine d'études). Comme nous l'avons déjà dit, nous avons opté en faveur d'une forme classique de la dissertation pour aborder la problématique de l'appariement (inspirée de Maniez). Pour cette raison,

l'opposition entre deux démarches nous est apparue comme une occasion intéressante d'approfondir la notion de sujet (ses aspects objectif et subjectif ; Maniez 2002, 123) et d'examiner les deux pôles principaux de la problématique de l'appariement. Dans le premier chapitre du mémoire, nous reprendrons l'analyse de l'approche centrée sur le document, centrale dans le développement de l'analyse documentaire « classique » (il s'agit de la « thèse forte » en bibliothéconomie). La dimension objective du sujet (section 1.1), le pragmatisme (section 1.2) et la réduction sémantique (section 1.3) seront tour à tour étudiés dans le but de dépister certains fondements incontournables de l'analyse documentaire. Le deuxième chapitre sera consacré à l'approche centrée sur l'utilisateur (à titre d'antithèse), approche théorique qui donne plus d'importance au contexte propre à l'utilisateur et tend à miser sur des outils participatifs. Plusieurs auteurs, comme nous le verrons, placent cette seconde tendance dans une perspective antithétique vis-à-vis des acquis traditionnels de l'analyse documentaire. L'aspect relativiste de la connaissance (section 2.1), la dimension subjective du sujet (section 2.2) et la *folksonomie* (section 2.3) dévoileront certaines des possibilités qu'offre une approche de l'analyse documentaire centrée sur l'utilisateur. Dans le troisième chapitre, nous nous pencherons sur deux aspects particuliers de l'environnement universitaire et leur impact potentiel sur l'analyse documentaire. L'analyse des domaines d'études (section 3.1) et la démarche analytico-synthétique (section 3.2) représentent deux avenues intéressantes pour aborder la problématique de l'appariement dans une perspective qui cible plus directement le milieu universitaire.

Pour conclure cette introduction, nous voulons établir deux dernières précisions sur le plan terminologique. En premier lieu, nous nous proposons d'utiliser le terme « système documentaire » qui nous semble mieux adapté aux diverses facettes de l'analyse documentaire que le terme « système d'information ». Il s'agit d'une question de sémantique qui peut paraître banale, mais le système documentaire, selon nous, décrit plus précisément le lien qui existe entre le document lui-même et sa représentation (*document surrogate*), de même qu'il suggère aussi le chemin qu'emprunte l'utilisateur lorsqu'il va du catalogue à la ressource documentaire comme telle (physique ou virtuelle). Principalement, le système documentaire

utilise comme « panneau indicateur »<sup>5</sup> la notice documentaire (ou notice bibliographique), fiche servant à représenter les différents attributs d'un document. C'est à l'intérieur des limites d'un tel système documentaire que les approches abordées dans ce mémoire peuvent jouer un rôle de premier ordre. En second lieu, la notion de document est fluctuante et peut faire intervenir une vaste typologie, selon l'extension qu'on désire lui donner. Aussi, nous nous limitons dans ce mémoire aux documents textuels, puisqu'ils occupent une place considérable dans les activités des usagers en milieu universitaire ; il est évident, par ailleurs, que divers types de documents font intervenir des problématiques distinctes et requièrent des solutions particulières que nous n'explorerons pas ici. Plusieurs articles et monographies s'intéressent à ces questions, mais celles-ci n'entrent pas dans notre propos.

---

<sup>5</sup> Nous verrons, dans la section 1.1, ce que cette expression veut dire dans le cadre de l'indexation.

# **Premier chapitre**

**Approche centrée sur le document**

Le premier chapitre du mémoire reprend plusieurs éléments bien connus de l'analyse documentaire. Tout d'abord, mentionnons que l'approche centrée sur le document peut aussi être appelée « analyse orientée objet » (*object oriented analysis*), définie de cette manière par Hudon : « Analyse documentaire basée sur le seul contenu du document » (Hudon 2013, 264). Cette définition succincte reflète un parti pris traditionnel de l'analyse documentaire (analyse basée *sur le contenu même du document*) et une tendance objectiviste sous-jacente. Nous aurons l'occasion de mettre en évidence ces particularités dans la mesure où elles nous renseignent sur les fondements de l'analyse centrée sur le document et sur les résultats qu'elle permet d'obtenir.

D'abord, il sera question de la dimension objective du sujet, laquelle implique la représentation indexicale d'un contenu thématique. Nous verrons plus en détails ce que cela veut dire, mais précisons dès à présent qu'il s'agit d'une représentation univoque, objective, neutre (section 1.1). L'analyse documentaire sera ensuite examinée à travers le prisme du pragmatisme, philosophie qui a joué un rôle capital dans son développement (section 1.2). Finalement, nous aborderons la notion de réduction sémantique (section 1.3) en tant que rouage principal d'un type d'analyse qui se pratique à l'aide de langages documentaires traditionnels (répertoires de vedettes-matières, thésaurus et schémas de classification LCC, DDC ou autres). L'objectif de ce premier chapitre est de préciser de quelle manière une approche centrée sur le document, en s'appuyant sur la dimension objective du sujet, oriente l'analyse documentaire. De ce point de vue, nous verrons que la réduction sémantique est au cœur d'une telle approche et qu'elle a un impact certain sur la problématique de l'appariement.

### **1.1) La dimension objective du sujet et l'*aboutness***

Le sujet est l'élément central sur lequel porte l'analyse documentaire, même si une approche disciplinaire appuie souvent le processus de représentation des contenus<sup>6</sup>. Il est donc important d'étudier de quelle manière le sujet est abordé et circonscrit par l'analyste

---

<sup>6</sup> La classification décimale de Dewey, comme ressource courante dans le processus d'analyse documentaire, fournit un bon exemple d'une approche axée davantage sur la discipline (ou domaine d'études) que sur le sujet à proprement parler.

documentaire. À la suite de Maniez (2002, 123), nous partirons de la définition suivante du sujet : « ce sur quoi s'applique la réflexion ». Nous sommes conscient de son niveau de généralité, mais cet énoncé nous permet de situer préliminairement la problématique de l'appariement identifiée dans l'introduction. Dans cette première section, nous tenterons de mieux définir « *ce* » sur quoi s'applique la réflexion, c'est-à-dire de comprendre ce qu'implique cette forme indexicale. Lorsqu'on emploie une formule telle que « voilà (ou *vois-là*) le sujet du document », qu'est-on en train d'indiquer ou « pointer »? Nous verrons qu'une forme d'indication est nécessaire dans l'analyse documentaire telle qu'elle se pratique traditionnellement.

D'une certaine manière, la notion de sujet renferme une dichotomie fondamentale qui est elle-même reflétée dans la structure générale de notre dissertation. Deux composants, l'objectif et le subjectif, doivent être considérés lorsqu'on tente de définir le terme « sujet ». Maniez, en s'inspirant de la définition lexicographique simple citée plus haut (« ce sur quoi s'applique la réflexion »), identifie cette exigence :

Dans sa brièveté, [*cette définition*] fait bien apparaître les deux composantes indissociables du sujet matière : d'une part, un « quelque chose », un objet de connaissance extérieur à l'individu qui s'y intéresse ; d'autre part, l'attention de cet individu en direction de cet objet. (Maniez 2002, 123)

Sans approfondir la réflexion philosophique rattachée à des notions comme « objectif » et « subjectif », il est important de noter que cette opposition, bien qu'elle ait tendance à obscurcir la notion de « sujet », alimente un ensemble de débats dans le domaine des sciences de l'information. Ce dialogue apparaît en filigrane dans ce mémoire et encadre la problématique de l'appariement qui, rappelons-le, concerne la tâche difficile de rapprocher le vocabulaire utilisé dans l'analyse documentaire et celui privilégié par l'utilisateur lors de ses recherches documentaires.

Malgré un certain degré d'indétermination et une terminologie complémentaire foisonnante (« information », « donnée », « contenu »...), la notion de sujet s'affirme comme élément central de l'analyse documentaire et, depuis plusieurs siècles déjà, comme une

dimension importante dans l'organisation de la documentation<sup>7</sup>. En tant qu'objet d'études, le sujet est particulièrement difficile à définir et ses traits essentiels peuvent varier selon les points de vue. Pour l'analyste documentaire, ce niveau de contingence relèverait peut-être du caractère de l'activité qui vise à dévoiler et représenter le sujet : l'analyse documentaire, comme plusieurs commentateurs l'ont souligné, doit être considérée comme un art plutôt qu'une science. Que l'on adopte ou non cette formule, il semblerait que l'analyste documentaire soit confronté à une tâche dont l'objet (la représentation du/des sujet(s) abordé(s) dans la documentation) demeure indéfinissable sous plusieurs aspects.

C'est l'une des raisons pour lesquelles plusieurs notions connexes ont été privilégiées et développées au cours des ans (voir, par exemple, Hjørland 2001). On peut compter parmi celles-ci : thème, concept, domaine d'études, contenu, rhème, information, connaissance, étiquette (*tag*)... Tour à tour, ces notions mettent l'accent sur diverses exigences et favorisent certains procédés analytiques. Dans le contexte du traitement documentaire, ce groupe de termes s'organise entre autres autour du choix difficile entre la nécessité de normaliser la métadonnée sujet (*subject access term*) et celle de répondre aux intérêts divers d'une communauté d'utilisateurs. Naun résume cette tension et souligne le rôle que cette dernière a joué depuis Cutter :

The tension between the requirement for a highly regulated subject access system and the diversity of interests that it must serve is one of the classic problems of librarianship, and the earliest texts in the literature show an awareness of it. This is the problem addressed by Cutter's dictum that the entry should be chosen that "will probably be first looked under by the class of people who use the library". (Naun 2006, 85)

La recherche documentaire probable d'un usager est difficile à prévoir, ce qui montre à nouveau que l'analyse documentaire peut difficilement être considérée comme une science

---

<sup>7</sup>Au dix-septième siècle, dans son *Advis pour dresser une bibliothèque*, Naudé montre l'importance de constituer un catalogue des auteurs et un catalogue des sujets de manière à rendre accessibles les richesses documentaires, puisque les livres comme toutes choses « ne se doivent estimer qu'à l'égal du profit et de l'usage que l'on en tire ». (Naudé 2008, 327-331) Avant cela, au seizième siècle, des « bibliographes » comme Konrad Gesner et Andrew Maunsell avaient de même souligné l'importance de fournir des points d'entrée de type sujet (Taylor 2009, 71-72).

exacte et que des outils polyvalents doivent être développés pour expliciter les intérêts des usagers (nous verrons ceci plus en détails dans les deux derniers chapitres du mémoire).

Comme le suggère Lancaster dans le cas spécifique de l'indexation, une analyse efficace implique au moins deux visées distinctes, dont l'une tend à relativiser les aspects thématiques de la documentation : « Effective subject indexing involves deciding not only what a document is about but also why it is likely to be of interest to a particular group of users. In other words, there is no one "correct" set of index terms for any item » (Lancaster 2003, 9). Une indexation efficace, sans être le résultat d'une démarche parfaitement systématique, peut donc être obtenue par un juste équilibre entre une représentation objective et impartiale<sup>8</sup> du/des sujet(s) du document (« what a document is about ») et une estimation des besoins informationnels des usagers. Comme nous le verrons, il s'agit essentiellement d'une démarche qui implique deux versions de l'analyse documentaire : l'une pragmatiste et objectiviste, l'autre subjectiviste et relativiste. Cette variation paradigmatique deviendra plus évidente à la fin des deux premiers chapitres du mémoire.

Nous abordons dans cette section la dimension objective du sujet, puisqu'elle a largement influencé le développement de l'analyse documentaire. Elle se traduit généralement par une approche centrée sur le document (*document-centered approach*), dont le réquisit est une attention particulière portée aux attributs du document : « A document-centered approach to indexing relies on the indexer's analysis of a number of document attributes as the basis for establishing the document's subject matter » (Mai 2005, 602). Mai poursuit en spécifiant les « attributs » (fréquemment cités) qui servent de point de départ pour l'analyse documentaire : le titre du document ou des chapitres, le résumé, la table des matières, la préface, l'introduction, les références bibliographiques, les illustrations, les tableaux, etc. On retrouve aussi une liste similaire dans la thèse de 2002 de Šauperl, *Subject determination during the cataloging process*, qui repose sur le témoignage de catalogueurs : « The table of contents, introduction, cover, blurb, abstract, and bibliography were cited as other important

---

<sup>8</sup> Naun commente l'importance de l'impartialité et de l'objectivité dans le processus d'analyse documentaire, en précisant que ces valeurs sont consignées dans un document aussi fondamental que l'*American Library Association's Bill of Rights* (Naun 2006, 85).

sources for subject determination by all catalogers » (Šauperl 2002, 119). Ceci fait du document un objet qui peut être décortiqué afin d'identifier les sources les plus fiables pour mener une analyse documentaire neutre.

Pour l'analyste documentaire, les attributs textuels mentionnés plus haut représentent les piliers de la lecture documentaire. Cette activité relègue temporairement le sujet à l'arrière-plan en faveur de l'étude de la structure textuelle : « It is clear then that professional indexers (and especially abstractors) develop an awareness of the structural properties that are inherent in text irrespective of its subject content » (Farrow 1991, 155). La lecture documentaire s'inscrit dans la continuité d'une longue tradition de description bibliographique, que l'on peut faire remonter au moins jusqu'à l'indexation par l'entremise des mots du titre (*title word indexing*<sup>9</sup>). La forme largement conventionnelle du document<sup>10</sup> (introduction, conclusion, division en chapitres, table des matières dans certains cas) permet à l'analyste documentaire d'isoler des éléments textuels, sans tenir compte initialement de leurs caractéristiques sémantiques. Ajoutons aussi cette remarque : réduire ainsi les sources d'information les plus fréquemment utilisées dans l'analyse favorise l'objectivité et la neutralité, ces propriétés étant prisées dans une approche centrée sur le document. Il devient ainsi plus facile de limiter la représentation du contenu aux éléments essentiels, de manière à éviter, par exemple, les problèmes de bruit dans un système documentaire.

Le sujet d'un document, dans une approche objectiviste, est représenté à l'intérieur des limites d'un régime indicatif. Metcalfe, selon nous, est le premier à avoir insisté sur cet élément central, tout en mettant en lumière la difficulté d'éviter le recours à ce qu'il nomme la communication de l'information (*i.e.* divulguer des renseignements supplémentaires, de nature potentiellement subjective, concernant le contenu indexé d'un document) :

---

<sup>9</sup> Une indexation basée sur les mots du titre n'est pas à tout coup objective, puisque son résultat repose en partie sur la terminologie et le « style » de l'auteur (utilisation de l'ironie, de la métaphore et autres tropes à tendance subjective). Nonobstant cette conséquence, ce type d'indexation comporte une composante objective manifeste en prescrivant une source d'information unique et localisée.

<sup>10</sup> Il est évident que tout document ne présente pas nécessairement une structure conventionnelle. Plusieurs articles pourront informer le lecteur qui s'intéresse davantage à l'analyse de documents sonores, visuels ou dont la forme textuelle n'est en rien conventionnelle.

Indexing should stop short of communication of the information indexed, though in practice it may be difficult to draw a line exactly between indication and communication. This may be regarded as a special case of the problem of librarianship in general and of bibliographical research work in particular, of drawing a line between providing the student with the materials of research, the documents, and doing his research for him. (Metcalf 1957, 213)

Si l'on regarde l'évolution de l'analyse documentaire effectuée à l'aide de termes contrôlés, l'approche privilégiée traduit une tendance vers l'objectivité (en limitant, entre autres par la lecture documentaire, le registre de communication), approche à laquelle contribue l'utilisation étendue de l'hyponymie. L'indication est alors le résultat d'un processus de subsumption (placer un contenu particulier, tiré d'un document, sous une classe ou une catégorie générique) dans lequel le sujet est affublé d'un nom générique, contrôlé et de facture neutre (la vedette-matière, le descripteur, un code de classification, etc.). Nous reviendrons plus en détails sur ces éléments à la fin du premier chapitre (section 1.3).

Cette subsumption cautionne l'utilisation de listes de termes, généralement regroupés pour former un langage documentaire, dont l'intérêt repose sur la représentation conventionnelle des sujets présents dans la documentation (fortement liée à une forme de conservatisme ; Richardson 1964, 34). Dans un catalogue bibliographique, cela se traduit par l'utilisation d'index qui contribuent au repérage de l'information (*Information Retrieval*), ceux-ci résultant d'une organisation de l'information sous forme d'indices terminologiques (normalisés dans la plupart des cas). Comme le souligne Maniez, « index » et « indice » partagent une même ascendance étymologique :

Les termes « indice », « index », « indexation » appartiennent à la même famille et les rapports qui les relient entre eux méritent d'être explicités. Indexer, au sens documentaire du terme, c'est accoler sur le document un indice – au sens large de « repère » –, autrement dit une étiquette permettant de repérer son sujet matière. Mais cet indice, pour être facilement retrouvé en cas de besoin, doit être placé dans une liste ordonnée, un index. On passe ainsi de l'apposition d'un *indice* à son versement dans un *index* où l'indice, suivi de l'adresse de son document, joue le rôle de panneau *indicateur*. (Maniez 2002, 144-145 ; l'auteur souligne)

L'indexation a donc comme fonction principale d'*indiquer* l'existence d'un document dont le sujet matière, représenté au sein d'une liste (l'*index*), pourrait intéresser un usager. Cette forme d'organisation est donc intimement liée au repérage de l'information de par son objectif :

« The purpose of indexing is to determine the subject matter of documents and express the subject matter in index terms (e.g. descriptors, subject headings, call numbers, classification codes, or index terms) to make subject retrieval possible » (Mai 2005, 599).

L'exigence d'indication soulignée par Metcalfe vient appuyer l'approche centrée sur le document dans un sens bien précis : l'analyste documentaire, plutôt que de traduire un point de vue, une opinion ou un contexte spécifique d'énonciation (issu d'une démarche scientifique particulière), se limite à un choix de termes génériques susceptibles de représenter objectivement le(s) sujet(s) d'un document. Un indice terminologique, débarrassé des complexités d'une communication en bonne et due forme (dans laquelle pourrait transparaître une opinion, un point de vue, un argument sous-jacent), favorise un repérage efficace de l'information dans la mesure où seulement l'*essence* du sujet matière est exploitée. La forme succincte que prend cet indice contribuerait à la clarté représentative de la notice bibliographique :

En conclusion, le champ « Indexation » apporte à la notice catalographique une plus-value importante. Plus précise que le titre, plus centrée sur l'essentiel que le résumé, la formule d'indexation n'est pas seulement un repérage des sujets ; elle offre aussi une image réduite du contenu des documents, et si elle est intégrée aux notices descriptives présentées à l'interrogateur, elle facilite son choix définitif. (Maniez 2002, 148)

Le terme d'indexation et l'indice de classification servent aussi bien à étiqueter un document qu'à proposer une « image » (ou représentation) de son contenu thématique. Cette seconde fonction, si elle paraît évidente aux yeux du bibliothécaire, doit s'imposer à titre de pré-supposé central de la tendance objective et s'affirmer comme un pôle important dans l'éclaircissement de la problématique de l'appariement<sup>11</sup>. Cette image objective ne concorde pas nécessairement avec celle que l'utilisateur s'est formée à partir d'un contenu thématique similaire, ce qui doit être vu comme l'une des limitations de l'approche centrée sur le document. La dimension objective du sujet favorise nécessairement une « image réduite du contenu des documents » (Maniez 2002, 148) et laisse de côté diverses interprétations qui pourraient émerger dans un registre de communication plus complexe (un résumé ou un compte-rendu de livre par exemple). En

---

<sup>11</sup> Il ne faut pas oublier que cette dernière inclut aussi bien le point de vue de l'utilisateur que celui du professionnel de l'information, puisque cette problématique confronte le vocabulaire utilisé par les deux parties.

abordant la problématique de l'appariement, nous cherchons dans ce mémoire à coupler deux registres de représentation (l'un conventionnel et balisé, l'autre libre et subjectif<sup>12</sup>) au sein d'un même espace bibliographique.

Une analyse documentaire qui favorise la dimension objective du sujet évacue l'accessoire (y compris des éléments contextuels ; voir la section 3.1 pour plus de détails sur la notion de contexte en analyse documentaire) et se concentre sur la représentation d'un « noyau thématique ». Ceci est nécessaire pour baliser le processus d'analyse et forger des conventions durables. Mais, comme nous l'avons déjà dit, le sujet demeure difficile à définir puisqu'il se manifeste de manières diverses dans la documentation. Parmi une multitude de concepts, le terme d'*aboutness* (difficile à traduire en français<sup>13</sup>) représente un compromis intéressant pour transcrire, en termes spécifiquement bibliothéconomiques, la notion de sujet. Fairthorne est généralement considéré comme celui qui a créé ce néologisme (Fairthorne 1969), mais Hutchins est peut-être l'auteur qui a le plus contribué à sa postérité. La définition de l'*aboutness* qui nous intéressera d'abord repose sur une forme de présupposition à propos du niveau de connaissances minimal requis pour lire un texte donné. Il s'agit d'assister une diversité d'utilisateurs, dont le niveau de connaissances peut varier grandement, par l'intermédiaire d'une indication simple, claire et révélatrice du contenu d'un document :

What I am suggesting, therefore, is that for this type of document need an index system should work with a definition of the 'aboutness' of documents which is formulated in terms of the knowledge presupposed by the authors of the texts. If index entries express this kind of document 'aboutness' then the basic conditions mentioned above for reader-document contact should be met: the user of the index is referred to documents on a topic about which he knows roughly what the authors of those documents presuppose of their readers, and he can seek out the documents (or just one of them) with some confidence that he will in fact learn something 'new' about the topic. (Hutchins 1978, 178)

---

<sup>12</sup> Ces deux registres de représentation reflètent la dualité concernant les deux versions de l'analyse documentaire mentionnées plus haut (l'une pragmatiste et objectiviste ; l'autre subjectiviste et relativiste). Ils déterminent chacun un type d'analyse (centrée sur le document / centrée sur l'utilisateur) et délimitent, par conséquent, les deux premiers chapitres du mémoire.

<sup>13</sup> La thématique d'un document, en tant qu'« ensemble organisé de thèmes » (*Trésor de la langue française informatisé*), traduit en partie la notion d'*aboutness*, sans toutefois refléter réellement ses aspects pragmatiste et réductionniste (nous examinons ceux-ci dans la suite de ce premier chapitre).

Hutchins oppose cette définition à celle d'un autre type d'*aboutness* qui résume le sujet d'un document (« summary of its content » ; Hutchins 1978, 173). Cette seconde version concerne plus directement la condensation, mais elle n'est pas étrangère à l'indexation et la classification qui, selon nous, se basent sur une forme de réduction visant à englober (dans la mesure du possible) la somme des contenus d'un document (voir la section 1.3).

En reprenant l'exigence d'indication de Metcalfe et le développement de Maniez concernant l'indice (voir plus haut), l'*aboutness* se définit comme l'indice langagier minimal servant à informer l'utilisateur (à la manière d'un panneau indicateur) de l'existence d'un document dont le contenu pourrait lui être utile. Si l'on propose l'indice « Manuel d'astronomie » à l'utilisateur, celui-ci est en droit de s'attendre à trouver un document qui aborde les fondements théoriques de l'astronomie. L'objectif est en quelque sorte de fournir un point de départ qui devrait convenir à la plupart des utilisateurs, sans procéder à une évaluation de leurs habitudes, connaissances et intérêts particuliers. L'*aboutness* d'un document, tel que nous l'avons défini, incite donc l'analyste documentaire à proposer un/des indice(s), plutôt qu'à produire un (*méta*-)texte (comme c'est le cas avec le résumé documentaire ; Hutchins 1978, 179). La notion d'*aboutness* reconduit les idéaux de neutralité et d'objectivité déjà mentionnés et demande à l'analyste documentaire de prendre en considération seulement ce que l'auteur d'un document exige de son lectorat en termes de connaissances préalables :

The indexers' task is to establish and record the topic or topics which the authors of documents themselves assume their readers should be starting from; they are required only to take account of the authors' own assumptions about their ('ideal') readers; they do not have to formulate their own image of readers. In this way, in theory at least, any 'interference' by the indexer between reader and documents should be minimized. (Hutchins 1978, 179)

Hutchins ajoute qu'une analyse documentaire basée sur cette version de l'*aboutness*, en minimisant les interférences dont serait responsable l'analyste documentaire, est particulièrement profitable à une institution documentaire non spécialisée (Hutchins 1978, 179), telle que la bibliothèque universitaire (en opposition à un centre de documentation dans une entreprise ou une organisation par exemple).

La notion d'*aboutness* favorise aussi un postulat lié à une certaine forme de permanence thématique. Beghtol, en opposant les *meaning(s)* potentiels d'un document à

l'*aboutness*, définit ce postulat : « For the present purposes, however, we may take the general position that texts of all kinds have a relatively permanent aboutness, but a variable number of meaning(s) » (Beghtol 1986, 85). La dimension objective du sujet implique généralement une analyse textuelle monosémique (une seule signification accordée à un énoncé). Ceci est l'une des raisons pour lesquelles les langages documentaires, outils centraux dans une approche centrée sur le document, n'offrent pas la même richesse sémantique que les langues naturelles. Alors que plusieurs significations peuvent être attribuées à tout énoncé lors d'une conversation ordinaire (communication complexe), l'*aboutness*, comme « noyau dur » thématique, se pose comme résultat d'une analyse sémantique à base référentielle unique. Dans ce contexte, le contrôle du vocabulaire est soumis à la nécessité d'associer un seul référent à chaque terme accepté (*i.e.* analyse sémantique à base référentielle unique), de manière à éviter les ambiguïtés. On obtient ainsi une seule signification (monosémie) pour chaque terme, lequel ne renvoie par conséquent qu'à un seul référent. On élimine de cette manière la polysémie présente dans les langues naturelles. Rendue manifeste à travers ce type d'analyse sémantique, la dimension objective du sujet nous permet de relever le caractère objectif et stable de l'analyse documentaire dans sa version traditionnelle.

Nous ne proposons en rien une approche novatrice de l'analyse documentaire dans cette première section, mais nous insistons sur l'importance d'une notion parfois négligée dans le domaine de l'organisation des connaissances. L'*aboutness* souligne une caractéristique différenciatrice stable d'un groupe de documents à l'intérieur d'une collection, tout en mettant en évidence le degré zéro (ou minimal) de la représentation du sujet. Comme le prétend Maniez, cette caractéristique, une fois livrée sous forme d'indice, joue le rôle de panneau indicateur (Maniez 2002, 144-145). Cette manière de décrire l'indice mis à la disposition de l'utilisateur pour repêcher une information pertinente sert de prélude à ce qui suivra dans ce premier chapitre. Nous verrons que l'analyste documentaire se place dans une optique résolument pragmatiste (section 1.2) et que la méthode qu'il privilégie est tributaire d'une réduction sémantique incontournable (section 1.3).

En résumé, lorsque l'analyste documentaire circonscrit *ce sur quoi porte le document*, qu'il prétend essentiellement « voilà le(s) sujet(s) principal/aux du document », il ne fait

qu'indiquer une porte d'entrée thématique (*subject access point*). Comme le fait remarquer Metcalfe, l'analyste documentaire ne peut établir qu'une forme de correspondance (objective mais partielle) entre éléments qui interviennent dans différents contextes : « There are subjects of inquiry, subject headings or classes under which documents on them are indicated, and the subjects of the indicated documents. Obviously there must be correspondence, but they need not agree in specification » (Metcalfe 1957, 222). L'*aboutness* d'un document doit pouvoir s'arrimer à l'objet de plusieurs requêtes d'information (*subjects of inquiry*) aussi bien que se substituer à un nombre indéterminé de significations potentielles issues du contenu de la documentation. La difficulté réside évidemment dans la mesure de correspondance entre ces deux éléments distincts, mesure qui détermine l'appariement potentiel de requêtes d'information (dans une large mesure indéterminées pour l'analyste documentaire) avec la représentation du/des sujet(s) présent(s) dans la documentation.

La dimension objective du sujet permet-elle d'arriver à une solution satisfaisante pour régler ce problème central dans l'analyse documentaire? Ce que nous avons tenté de montrer c'est qu'elle permet de positionner l'*aboutness* dans un système documentaire indexical, c'est-à-dire un système documentaire dont les outils de représentation pointent vers des valeurs constantes et stables<sup>14</sup>. Une analyse documentaire qui favorise une telle démarche permet de proposer des métadonnées qui, dans la perspective d'une approche centrée sur le document, se limitent à la dimension objective du sujet. L'avantage reconnu d'une telle forme d'analyse serait de construire un catalogue proposant des panneaux indicateurs conventionnels, des portes d'entrée sujet (*subject access points*) connues ou du moins reconnaissables pour les usagers. Ceci se réalise généralement par l'entremise du contrôle du vocabulaire, lequel sera abordé plus en détails à la fin de ce premier chapitre. Nous verrons auparavant qu'un point de vue pragmatiste a occupé une place importante dans le développement moderne de l'approche centrée sur le document, en ce qu'il fournit une base intellectuelle pouvant justifier le recours à un système documentaire indexical.

---

<sup>14</sup> Un système documentaire indexical propose des valeurs constantes et stables, qui pourraient être verbalisées comme suit : *LE* sujet de *CE* document est *LE* « calcul différentiel ». Cette valeur, une fois fixée selon les normes en vigueur, ne changera ni dans le temps, ni selon la personne qui contribue au système documentaire (l'analyste documentaire) ou encore celle qui en fait usage. Par opposition, une représentation polysémique permettrait de décrire de diverses manières (susceptibles de changer avec le temps) un même sujet.

## 1.2) L'analyse documentaire en tant qu'exercice pragmatiste

Pour répondre aux exigences de l'organisation de la documentation, l'analyste documentaire se place dans une optique pragmatiste. D'un point de vue philosophique, le pragmatisme peut être résumé par l'attention portée aux effets significatifs, aux gains pratiques qu'une idée (ou un concept) peut avoir sur notre expérience du réel (Lalande 2010, 803-807). Pour le formuler en termes plus spécifiques, le pragmatisme exige d'évaluer le résultat de l'analyse documentaire à raison de l'amélioration de l'expérience de recherche documentaire (pragmatisme documentaire). Même si l'analyse documentaire peut assumer un ensemble de fonctions, comme le remarque bien Hudon, le repérage de l'information s'affirme comme une exigence centrale : « L'analyse documentaire est directement liée au repérage de l'information et de documents, et la qualité de l'analyse est au moins partiellement garante de l'efficacité et de la qualité du repérage » (Hudon 2013, 15). Nous verrons plus bas que l'importante notion d'entrée sujet spécifique (*specific subject entry*), développée par Cutter (1904, 22), rencontre à merveille cette exigence. L'objectif de cette section est de montrer dans quelle mesure une orientation pragmatiste définit l'approche centrée sur le document et quel impact cela peut avoir sur l'analyse documentaire.

Cela peut sembler abusif d'associer une activité professionnelle contemporaine avec un courant philosophique âgé de plus d'un siècle, mais plusieurs éléments justifient ce rapprochement (voir Dousa 2010). Au haut de la liste, nous retrouvons un facteur historique qui a favorisé une orientation pragmatiste dans le développement de l'analyse documentaire. En effet, la naissance de l'analyse documentaire moderne est marquée par une époque qui représente l'âge d'or du pragmatisme (au tournant du vingtième siècle). Les rapprochements sont nombreux entre Paul Otlet, Charles Ammi Cutter, Melvil Dewey, Henry Bliss et des penseurs tels que Charles Sanders Peirce, John Dewey, William James. L'influence de ces derniers s'affirme de plusieurs manières, mais nous voulons surtout montrer que le pragmatisme (avec Peirce tout particulièrement) représente une conception philosophique en adéquation avec les principes fondamentaux de la dimension objective du sujet (section 1.1).

Le pragmatisme, dans sa version classique, est une philosophie qui propose de tendre vers une forme de consensus dans l'évaluation de nos connaissances. C'est en examinant cet

énoncé, dans la suite de cette section, que nous serons en mesure de montrer quel rôle peut jouer ce courant de pensée dans l'orientation que prend l'analyse documentaire. Il sera donc question de notions qui décrivent la posture pragmatiste, mais aussi du lien qui existe entre cette philosophie et le point de vue plus spécifique qu'ont adopté certains auteurs dont les écrits ont façonné l'analyse documentaire moderne (Otlet, Bliss et Cutter surtout).

Le *Traité de documentation* d'Otlet, publié en 1934, peut être vu comme l'un des ouvrages emblématiques des sciences de l'information. L'objectif principal de ce livre, empreint d'un scientisme fort critiqué depuis (voir les écrits de Budd par exemple), est de poser les bases d'une science portant sur le livre et l'organisation de ses contenus (*bibliologie*). La *bibliologie*, ou *documentologie*, prend appui sur une pratique documentaire qui réunit des préoccupations à la fois objectivistes et pragmatiques, comme en témoigne l'ouverture du *Traité de documentation* :

Les buts de la Documentation organisée consistent à pouvoir offrir sur tout ordre de fait et de connaissance des informations documentées : 1° universelles quant à leur objet ; 2° sûres et vraies ; 3° complètes ; 4° rapides ; 5° à jour ; 6° faciles à obtenir ; 7° réunies d'avance et prêtes à être communiquées ; 8° mises à la disposition du plus grand nombre. (Otlet 1989, 6)

Ces huit objectifs mettent l'accent avant tout sur l'amélioration de l'« expérience documentaire » que peut escompter un usager dans ses recherches, par exemple un étudiant à la recherche d'informations sur le courant réaliste en peinture. Pour donner une idée de la complexité à laquelle peut être confronté l'utilisateur, nous pourrions parler de la nécessité de conjuguer les aspects historique, esthétique, politique du réalisme en peinture. L'étudiant peut rapidement être submergé par la quantité de documents qu'une telle recherche peut dévoiler ou encore se voir surpris de ne pas trouver d'information précise sur l'un ou l'autre de ces aspects. Les visées d'Otlet, nous semble-t-il, conviennent parfaitement à l'utilisateur en milieu universitaire, pour qui la recherche documentaire peut représenter une tâche ardue et complexe. Les huit objectifs consignés dans la citation ci-haut visent à rendre la recherche documentaire pratique (commode) et exhaustive, en favorisant une organisation raisonnée et fiable de l'information.

Otlet, comme plusieurs auteurs de son époque et certains de ses prédécesseurs, est sensible au « foisonnement documentaire » et au chaos potentiel que celui-ci peut engendrer :

Le chaos du livre et des documents appelle de nos jours une science qui obvierrait au mal de la documentation devenue désordonnée, répétitive, contradictoire, un mal comparable sous certains aspects à celui des Rhéteurs dont Aristote finit par triompher. Cette science serait pour l'ordre à mettre dans les documents le prolongement de la Logique, qui est la science de l'ordre à mettre dans les idées. (Otlet 1989, 27)

L'ordre s'impose en ayant recours à une méthode et, pour un penseur dans la foulée du pragmatisme classique (pragmatisme du tournant du 20<sup>e</sup> siècle), une méthode valable doit être déterminée par une expérience objective du monde, plutôt que par des croyances contingentes et subjectives : « The world as we observe it, Otlet noted, provides only knowledge of *prima facie* unconnected and chaotically distributed particulars. In order to arrive at true, i.e. scientific knowledge of the world, human beings have to simplify and order multiplicity » (Ducheyne 2009, 224). Cela veut dire qu'organiser la multiplicité (*i.e.* établir des liens raisonnés entre phénomènes particuliers et isolés) est un enjeu épistémique important pour toute démarche qui vise l'avancement des connaissances, incluant une éventuelle *bibliologie*. C'est pourquoi l'organisation de l'information doit reposer sur une méthode à l'épreuve du « chaos du livre », c'est pourquoi la science est le modèle à suivre pour « l'ordre à mettre dans les documents » (Otlet 1989, 27).

Dans la perspective pragmatiste classique, l'organisation de la documentation participe d'une recherche d'informations objectives, misant sur une certaine forme de gain pratique inspiré de la méthode scientifique (et des liens raisonnés qu'elle permet d'établir entre phénomènes particuliers et isolés). En définissant le document en tant qu'objet d'études scientifiques, Otlet élabore une conception de l'organisation documentaire axée sur une information à la fois fiable et facile à diffuser. Il suggère une *bibliologie pure* « édifiée sur la base d'un concept composé des quatre éléments suivants : a) la représentation du monde ; b) par un système de signes ; c) sur des supports pratiques et maniables ; d) donnant lieu à des enregistrements qui puissent être conservés, communiqués et diffusés » (Otlet 1989, 26). Il est facile de voir, surtout avec les points c) et d), que l'accent est mis sur la recherche et la diffusion d'informations, auxquelles peut contribuer la science *bibliologique*. Toutefois, les

points a) et b) soulignent aussi la nécessité de fournir une information fiable (conforme à la réalité) de manière systématique (*i.e.* à l'aide d'un système de signes qui permet d'établir des liens logiques entre phénomènes particuliers et isolés<sup>15</sup>).

La méthode scientifique constitue une sorte d'idéal (« scientisme ») pour plusieurs penseurs pragmatistes du tournant du siècle dernier. Il est assez facile de voir qu'Otlet est tout autant perméable à cette idée, mais nous devons surtout tenter de comprendre comment le système documentaire peut favoriser une expérience de recherche pratique (commode) pour le plus grand nombre. Comme nous y avons brièvement fait allusion dans la section précédente, le conservatisme joue un rôle non négligeable en bibliothèque (Richardson 1964, 34). Si l'on s'inspire de Peirce, initiateur du pragmatisme, il est possible de parler d'habitudes de pensée (« habits of the mind ») dont le rôle est d'orienter objectivement nos croyances et nos actions : « In his view, the beliefs we hold about the world are habits of mind formed on the basis of our experience and regulating our actions vis-à-vis the world » (Dousa 2010, 66). Un apport pragmatiste favorise le recours aux mêmes habitudes de pensée (convergence, régulation objective dans la représentation des sujets) aussi bien lors du processus d'analyse que lors des recherches documentaires subséquentes menées par les usagers. L'approche centrée sur le document puise dans la dimension objective du sujet (voir la notion d'*aboutness* dans la section 1.1) afin d'instituer, aussi bien chez l'analyste documentaire que chez l'utilisateur, des habitudes de pensée similaires (à l'aide de règles comme celles de Cutter apparues en 1876 ; voir Cutter 1904). Un défi important est d'éviter que l'utilisateur ne cultive des croyances fautives (ce qui dépasse le rôle qu'on attribue habituellement au bibliothécaire), mais surtout de lui éviter des actions inutiles telles qu'une recherche à l'aide de termes non pertinents ou imprécis.

Pour plusieurs commentateurs, il semble donc évident que l'influence du pragmatisme classique sur l'analyse documentaire est accompagnée d'une attitude scientiste. Aussi bien chez Bliss et Cutter qu'Otlet, on peut noter des accents qui se rapprochent beaucoup d'une croyance

---

<sup>15</sup> On doit rapprocher un tel système du langage documentaire, dans la mesure où celui-ci permet d'établir des relations sémantiques entre divers termes qui demeureraient isolés sans cela.

aveugle dans la méthode scientifique<sup>16</sup>. Les marques d'une telle conception épistémologique se trouvent une fois de plus chez le fondateur du pragmatisme, Charles Sanders Peirce :

A necessary condition for obtaining a satisfactory result to inquiry is the use of a correct method. This method, in Peirce's opinion, is the "method of science", whose cardinal feature is that through it, "our beliefs may be determined by nothing human, but by some external permanency—by something upon which our thinking has no effect". (Dousa 2010, 66-67)

Ce qui nous importe n'est pas de faire le procès du scientisme, mais d'établir un lien évident avec l'approche centrée sur le document. En effet, le document doit être vu comme un objet sur lequel la subjectivité (de l'analyste documentaire) n'a pas de prise, puisqu'il (le document) est porteur de connaissances déterminées par une permanence externe (« external permanency »). Otlet parle d'ailleurs d'une matérialisation de la pensée par le livre :

Le livre, le document, ont apporté une réalité nouvelle distincte de toutes les autres : la matérialisation de la pensée. Comme la pensée est une image des choses, le livre est venu donner une reproduction, une copie du monde, celui-ci étant tenu comme le modèle. (Otlet 1989, 425)

Cette position pragmatiste classique donne sa légitimité à un processus de réduction hétérogène (section 1.3) qui se base sur le document lui-même, plutôt que sur l'avis versatile des usagers, les connaissances nécessairement limitées de l'analyste ou encore le témoignage d'un auteur dont les réelles intentions demeurent inconnues.

Sur le continent nord-américain, Cutter est certainement l'auteur qui a exercé la plus grande influence et marqué le plus fortement le développement des méthodes de description et d'analyse documentaires. Sa démarche est ancrée dans la philosophie et les préoccupations de son temps, et ses *Règles* (1904) représentent une entreprise axée sur la dimension pratique de la tâche dont s'acquitte l'analyste documentaire<sup>17</sup>. Nous examinerons plus attentivement l'importante notion d'entrée sujet spécifique (*specific subject entry*) pour illustrer la contribution de Cutter et la relation qu'entretient cet héritage avec le pragmatisme. D'abord,

---

<sup>16</sup> Bliss est peut-être l'auteur qui endosse le plus fortement une position scientifique : « Bliss and Peirce thus both envisioned that the body of scientific beliefs ratified by scientific consensus could offer a true account of the way the external world is—a view born of the shared confidence in the efficacy of scientific method » (Dousa 2010, 67).

<sup>17</sup> Rappelons que le pragmatisme, dans la perspective qui nous intéresse (pragmatisme documentaire), est défini par l'impact qu'exerce l'analyse documentaire sur l'expérience de recherche documentaire

précisons ce que Cutter entend lui-même par *specific subject entry* : « *Specific entry*, registering a book under a heading which expresses its special subject as distinguished from entering it in a class which includes that subject » (Cutter 1904, 22 ; l'auteur souligne). Une distinction est donc établie entre le choix d'une entrée spécifique (*specific entry*), représentant plus précisément le sujet, et l'inclusion d'un sujet dans une classe lors du processus de classification. Cutter formalise de cette manière un vocabulaire (langage d'indexation) dont la fonction, au sein du catalogue, est de représenter le/les sujet(s) d'un document avec le niveau de spécificité le plus pertinent pour la recherche documentaire.

L'accès par sujet spécifique privilégie le recours à une syntaxe spéciale, qui s'est longtemps manifestée par l'inversion de l'énoncé « naturel » (ex. : « Philosophie, histoire de » plutôt qu' « Histoire de la philosophie »<sup>18</sup>). En théorie, ceci met l'accent, dans l'organisation d'un index alphabétique, sur le noyau du sujet (« philosophie » dans l'exemple précédent). Le terme anglophone « heading » (dans l'expression *subject heading*) renvoie d'ailleurs à l'expression (l'en-tête) qui détermine l'arrangement alphabétique du terme d'indexation : « *Heading*, the word by which the alphabetical place of an entry in the catalog is determined, usually the name of the author, of the subject, or of the literary or practical form, or a word of the title » (Cutter 1904, 20 ; l'auteur souligne). Pour Cutter, l'élément *heading* doit refléter le concept qui exprime le plus directement le sujet spécifique d'un document, de manière à en suggérer la clé d'accès la plus efficace. Avec ses *Règles* (1904), Cutter place au cœur de l'analyse documentaire la méthode d'organisation du catalogue, espérant ainsi favoriser de bonnes habitudes d'analyse chez le bibliothécaire, mais aussi de bonnes habitudes de recherche chez l'utilisateur. Selon lui, l'accès par sujet spécifique remplit mieux ces deux critères qu'une organisation catalographique basée sur la classification (*classified catalogue*).

Cutter favorise par conséquent, à travers l'indexation, une représentation du sujet qui en accentue l'élément central (le noyau). Ceci requiert de la part de l'utilisateur l'utilisation d'un type particulier d'organisation thématique : en effet, dans un catalogue de type dictionnaire (*dictionary catalog*), le sujet est annoncé par l'expression jugée centrale (*heading*) et non par

---

<sup>18</sup> Il faut cependant rappeler que l'inversion dans la formulation du terme d'indexation n'est plus aussi répandue avec le développement des moteurs de recherche automatisés (Hudon 2013, 108).

la position de celui-ci dans une vaste structure de connaissances (*classified catalog*). Comme nous l'avons montré, une telle démarche convient à des index alphabétiques qui misent sur la représentation spécifique des sujets (*specific entry*). Le postulat est que l'utilisateur doit développer l'habitude de rechercher la documentation en utilisant une terminologie précise, qui laisse de côté ce qui est en périphérie du sujet convoité. Ce n'est pas dire que le seul critère valable pour l'indexation est la précision, mais il s'est rapidement imposé comme le plus important pour favoriser le repérage. Nous pouvons encore aujourd'hui constater l'effet que ce critère exerce sur l'analyse documentaire : « L'exhaustivité de l'indexation pourrait favoriser à la fois le rappel et la précision, mais c'est le degré de spécificité de la représentation qui aura l'influence la plus importante sur le repérage » (Hudon 2013, 45). Hudon mentionne par la suite l'importance de la satisfaction de l'utilisateur en tant qu'élément à considérer dans l'évaluation des méthodes d'indexation. C'est exactement ce qu'avait en tête Cutter lorsqu'il a rédigé ses *Règles* : « General rules, always applicable, for the choice of names of subjects can no more be given than rules without exception in grammar. Usage in both cases is the supreme arbiter,— the usage, in the present case, not of the cataloguer but of the public in speaking of subjects » (Cutter 1904, 69). Le pragmatisme documentaire possède l'avantage de présenter clairement cette exigence et, en même temps, de questionner explicitement la possibilité d'un consensus dans la représentation d'un sujet (comme nous allons le voir dans la suite de cette section).

Plusieurs auteurs ont puisé dans une version ou une autre du pragmatisme pour contribuer à l'avancement des méthodes d'organisation de la documentation, Henry E. Bliss et Paul Otlet comptant parmi les plus connus. Comme nous l'avons souligné, Otlet adhère à une philosophie pragmatiste, axée sur l'amélioration de l'organisation documentaire à partir d'un modèle (voire d'un idéal) scientifique. Avec sa *bibliologie*, il cherche à établir (tout comme Bliss d'ailleurs) une chaîne livre-science-réalité : « On distingue : 1° *Les livres*, dans leur ensemble, ont pour fin la transcription des connaissances scientifiques ; 2° *La Science* a pour but la connaissance de la réalité ; 3° *La Réalité* est l'ensemble des choses corporelles qui existent (monde ou univers) » (Otlet 1989, 373 ; l'auteur souligne). En adhérant à ce modèle, il devient envisageable d'établir une correspondance entre le contenu (scientifique) du livre et une réalité *objective*. Au-delà du scientisme auquel nous avons déjà fait allusion, cette

correspondance évoque aussi la nécessité, dans la tâche qui consiste à représenter l'*aboutness* d'un document, de tendre vers une forme de consensus. Cet idéal traverse l'œuvre de Bliss, souvent négligée en Amérique du Nord, et permet de rapprocher l'analyse documentaire de la pratique scientifique (les principes d'organisation de la documentation développés par Bliss le montrent d'ailleurs très bien ; Bliss 1929).

Le présupposé central est que la science met au jour un ordre présent dans la nature, ordre duquel l'analyse documentaire pourrait tirer un mode d'organisation de la documentation à la fois stable et consensuel. Bliss explique comment cette idée profite à l'organisation de la documentation :

Considering the relativity and adaptability of classifications, the inherent consistency of the several views with the scientific and naturalistic order clearly indicates that this order should prove valid and adaptable to the majority of interests, scientific, philosophic, and pedagogic. For the apportionment of research and for the classification of books in libraries it would make for the *maximal efficiency*. In short, the scientific order is, as we have repeatedly emphasized, the most adaptable to the several consistent views, the most efficient in serving the predominant interests, and the most practical for library classification. (Bliss 1929, 410-411; nous soulignons)

C'est en postulant une efficacité maximale à la fois dans la recherche (recherche scientifique et académique) et dans l'organisation de la documentation que Bliss justifie le recours à un consensus scientifique et éducationnel<sup>19</sup> (voir Sayers et Maltby 1975, 58). Selon nous, il faut considérer ce postulat comme une forme d'idéal, un objectif toujours présent mais paradoxalement inaccessible. L'inclusion ou l'exclusion d'un terme au sein d'une liste d'autorité implique aussi bien l'impossibilité de refléter les préférences terminologiques de chaque usager, que la volonté de favoriser chez les analystes documentaires et les usagers l'utilisation d'une appellation commune.

---

<sup>19</sup> Pour Bliss, l'organisation se manifeste sur les plans social, scientifique, éducationnel et documentaire suivant un même objectif et une logique commune. L'organisation des connaissances devient le reflet de l'organisation de l'expérience, de la pensée, de la volonté, de l'effort, de l'intention (Bliss 1929, 16). Ce postulat, dans un langage qui s'inspire fortement du pragmatisme classique du tournant du 20<sup>e</sup> siècle, représente le fondement chez Bliss de la possibilité même d'une logique de l'organisation documentaire, basée sur l'expérience *commune* des usagers qu'elle prétend servir (étudiants, professeurs, chercheurs, etc.).

En définitive, l'approche centrée sur le document repose avant tout sur une dimension formelle (informations communiquées sous la forme, souvent conventionnelle, d'un document ; voir la section 1.1) qui facilite le repérage de l'information :

Disembodied information, like the grin of the Cheshire cat, cannot be grasped or retrieved. Only when it is delimited, in the sense of being packaged into a document, electronic or otherwise, does it become accessible, and only then is it a candidate for bibliographic description. (Svenonius 2000, 109)

Nous pouvons encore remarquer, avec Svenonius, que l'analyse documentaire centrée sur le document accompagne une méthode de description qui prend comme point de départ la forme physique du document :

In the Anglo-American cataloging tradition, it is a cardinal rule that the starting point for bibliographic description should be the physical form of the item in hand. Applying this rule sets up a one-to-one relationship between documents and the descriptions that are their surrogates. (Svenonius 2000, 108)

Décrire un document et son contenu, si l'on part de cette règle centrale, représente la tentative d'établir une relation directe (« one-to-one relationship ») entre les documents et leurs substituts (« their surrogates »), eux-mêmes déterminés par une permanence externe dans l'optique de l'épistémologie pragmatiste (Dousa 2010, 66-67). Cette relation directe donne une importance accrue au contenu même du document (et à signification « intrinsèque ») vis-à-vis certains facteurs externes (considérés néanmoins dans l'analyse documentaire), tels que le contexte dans lequel œuvre l'institution documentaire, le(s) domaine(s) d'études auquel(s) on associe un sujet ou le point de vue des usagers qui utilisent le système documentaire (ces éléments seront examinés plus en détails dans la suite du mémoire).

Le pragmatisme documentaire est essentiellement la philosophie adoptée pour favoriser de bonnes habitudes aussi bien dans l'organisation de la documentation (comme en témoignent les *Règles* de Cutter) que dans l'utilisation d'un système documentaire (recherche à partir de termes précis ou *specific subject entries*). Ceci obéit à l'idéal *blissien* qui consiste à rechercher une forme de consensus pour l'organisation de la documentation. Nous avons vu que les gains pratiques peuvent être appréciables lorsqu'on privilégie une notion telle que l'entrée sujet spécifique (*specific subject entry*). Il s'agit, selon nous, d'un élément qui occupe

une place importante dans le développement de l'analyse documentaire et qui traduit la nécessité de proposer des résultats précis et tangibles (métadonnées proches du document lui-même en tant qu'objet de l'analyse<sup>20</sup>). L'approche centrée sur le document propose une représentation du contenu susceptible d'être partagée par le plus grand nombre, dans la mesure où les éléments ponctuels, contingents ou subjectifs sont mis de côté et la signification intrinsèque du document est privilégiée :

However, the aim of the indexer is to establish the subject matter independently of any particular use and domain and only place the document in a particular use after the analysis of the document and the establishment of the subject matter has taken place. [...] The focus of the document-centered approach is on the documents' intrinsic meaning and subject matter [*aboutness*] and on the translation of the subject matter into the users' needs. (Mai 2005, 607)

Dans le troisième chapitre, nous regarderons l'analyse des domaines d'études (notion que Mai place au cœur d'une certaine approche de l'analyse documentaire) comme stratégie pour bonifier l'analyse documentaire dans sa version traditionnelle.

Encore une fois, il n'est pas question ici de réinventer ou réinterpréter une approche qui a fait ses preuves depuis plus d'un siècle. Cela ne veut pas dire pour autant que l'approche centrée sur le document soit en mesure, à elle seule, d'éliminer tous les problèmes liés au repérage de l'information. Nous pouvons facilement voir que la problématique de l'appariement serait difficilement résolue par l'entremise d'une analyse documentaire à visée consensuelle. Malgré son importance, il s'agit d'une méthode parmi d'autres pour représenter le ou les sujet(s) présent(s) dans la documentation. Elle est en outre issue d'un parti pris en faveur de la dimension objective du sujet, ce qui laisse de côté en bonne partie le point de vue de l'utilisateur. Nous pouvons par conséquent poser la question suivante : l'*aboutness*, si elle permet d'encadrer la notion de sujet dans sa dimension objective, répond-elle de façon assez précise aux attentes des usagers lorsque ceux-ci entreprennent une recherche documentaire ? Pour approfondir cette question, nous devons introduire la notion de réduction sémantique et examiner comment le contrôle et la structuration du vocabulaire affectent l'appariement.

---

<sup>20</sup> Témoignant de l'*aboutness* du document, ces métadonnées représentent un résultat précis et tangible dans la mesure où elles se limitent essentiellement à la signification « intrinsèque » du contenu du document (plutôt qu'aux multiples interprétations, ou exégèses, qu'un étudiant, un professeur ou un chercheur peut en tirer).

### 1.3) La réduction sémantique

Après avoir abordé le sujet dans sa dimension objective et examiné l'impact du pragmatisme sur le développement de l'analyse documentaire, il est important de regarder de plus près le processus même d'analyse sur le plan sémantique (signification accordée à un terme ou à un énoncé). Faisons auparavant un court rappel. La dimension objective du sujet s'incarne dans la notion d'*aboutness* et s'actualise dans une tendance forte en analyse documentaire, l'approche centrée sur le document. Cette dernière se pratique en accord avec une démarche pragmatiste dont le résultat prend souvent la forme de règles de contrôle et de structuration du vocabulaire (les *Règles* de Cutter étant l'exemple canonique). Tout ceci repose sur la nécessité de servir la cause du repérage de l'information, faisant de l'analyste documentaire une sorte de courtier ou médiateur : « Subject descriptions of documents in library catalogs are provided by catalogers, who are the crucial mechanism linking authors to the readers of documents by developing representations to support information retrieval » (Šauperl 2002, 1). Dans cette section, nous allons tenter de mettre en lumière ce qui nous semble être le mécanisme principal dans une analyse centrée sur le document, la réduction sémantique. Cette notion nous aidera à mieux comprendre le rôle que peuvent encore jouer les langages documentaires par l'entremise du contrôle et de la structuration du vocabulaire.

Reprenons l'analyse à partir de ce qui a été souligné à la fin de la section précédente, à savoir l'importance accordée à la « signification intrinsèque » dans l'approche centrée sur le document. Le caractère immanent de l'analyse (menée presque exclusivement à partir du contenu du document) place celle-ci dans une position particulière. L'approche centrée sur le document, parce qu'elle s'appuie sur la dimension objective du sujet, les caractéristiques formelles du document et une indication minimale du contenu (*aboutness*), peut être dite *autonome*. Certains aspects de l'expérience de recherche documentaire, qui pourraient influencer l'analyste documentaire, sont nécessairement mis de côté en faveur de cette autonomie dans le processus d'analyse. Par exemple, la satisfaction de l'utilisateur (comme utilisateur du système documentaire), mais aussi les intentions de l'auteur sont largement absentes du processus d'analyse, même si le nom de ce dernier peut à l'occasion suggérer à l'analyste une orientation particulière (Šauperl 2002, 131). En amont comme en aval du traitement analytique, une certaine part d'incertitude ne peut être ignorée : « There is no

feedback from users as to whether one or another subject heading is more helpful. There is also no feedback from authors as to whether their intent was correctly interpreted » (Šauperl 2002, 142). Nous pouvons voir ici une limitation importante de l'analyse documentaire lorsqu'elle s'appuie sur des règles dont l'objectif demeure de répondre à des critères bien connus, tels que le rappel et la précision d'un système documentaire<sup>21</sup>, plutôt que de répertorier le point de vue d'un usager ou les intentions d'un auteur.

Il est important de spécifier que la dimension objective du sujet, avec les avantages et les limites qui l'accompagnent, est fortement liée à la réduction sémantique. Elle agit en effet comme base indispensable dans l'approche centrée sur le document qui, comme nous venons de le spécifier, procède d'une analyse *autonome*. Mai suggère deux arguments principaux en faveur de l'approche que nous tâchons de décrire tout au long de ce premier chapitre, la stabilité du résultat de l'analyse et sa dimension tangible (signification intrinsèque) :

The main arguments in favor of the document-centered approach are that: (1) the representation of the document will be based on stable characteristics of the document that are not likely to change over time and that indexing therefore will be of lasting value; and (2) it is difficult to predict how the document will be used in the future and it is therefore better to index the document according to *what it actually contains*. (Mai 2005, 607 ; nous soulignons)

À défaut de favoriser une représentation détaillée du contenu, la dimension objective du sujet permet de circonscrire une assise concrète et durable pour l'analyse documentaire. Comme nous l'avons vu avec la notion d'entrée sujet spécifique de Cutter, le but est de servir une majorité d'utilisateurs en fournissant une représentation claire (sans ambiguïté) du contenu et, dans la mesure du possible, une entrée sujet précise et prévisible (susceptible d'être utilisée par un usager faisant appel au catalogue). La réduction sémantique est au cœur de cette démarche.

Sur le plan épistémologique, l'adjectif « réductionniste » peut être attribué à une méthode scientifique qui établit une séparation claire entre le scientifique (en tant qu'individu) et son objet d'études. L'article de Jones intitulé « Reductionism and library and information

---

<sup>21</sup> La notion d'entrée sujet spécifique de Cutter (*specific subject entry*) constitue une sorte de prélude au critère de précision comme nous l'avons déjà souligné dans la section précédente.

science philosophy » évoque l'influence que cette assertion (fortement objectiviste) a exercée sur le développement des sciences de l'information (imprégnées au moins partiellement, rappelons-le, par le pragmatisme classique) :

One of the most fundamental reductions is the distinction that posits that humans are separate from nature; this then sustains the idea that what science properly studies in its pure form is nature as an objective "thing". The influence of this can be seen at the very heart of LIS [*Library and Information Science*] in the ongoing debate over how we define "information". » (Jones 2008, 492)

Souvent associée au courant positiviste qui a marqué la période 1850-1950, une telle « métathéorie » est qualifiée de réductionniste par plusieurs commentateurs, qui appellent généralement à un renouveau ou au dépassement de cette position éculée<sup>22</sup>. Par la même occasion, une longue tradition d'analyse documentaire est souvent discréditée, de même que la possibilité de mettre en relief certains acquis qui peuvent toujours contribuer au développement de l'analyse documentaire (et des sciences de l'information en général). Nous allons adopter un point de vue opposé et montrer comment la réduction sémantique contribue plutôt à solidifier l'approche centrée sur le document, par le contrôle et la structuration du vocabulaire.

Nous nous intéressons avant tout à la question du réductionnisme parce qu'elle permet d'explorer un lien possible entre les aspects objectifs et subjectifs qui président à l'organisation de la documentation. Par exemple, le contrôle de l'information (plus précisément le contrôle du vocabulaire dans le cas qui nous intéresse) et le repérage de l'information pourraient tous deux être vus comme des démarches complémentaires. Remettre en question une certaine forme de réductionnisme favoriserait le développement d'une approche plus inclusive :

For instance a more inclusive or holistic approach would account for both the "objective", as we find this researched in science, and the "subjective", which is central to the achievements we have made in the humanities. [...]

---

<sup>22</sup> Dans cette perspective, le réductionnisme se présente comme le point de départ d'une réflexion visant aussi bien à comprendre un héritage important en sciences de l'information qu'à faire la promotion d'approches plus récentes et fortement critiques. L'« angle péjoratif » affiche un véritable potentiel polémique, mais ne rend pas très bien compte de l'apport caractéristique du réductionnisme dans le processus d'analyse. Pour plus de détails, voir par exemple les écrits de Budd (2001 et 2004) sur le « scientisme déterministe » et son impact dans le développement des sciences de l'information.

Arguably, for LIS, a major advantage in such a move is the opportunity to adequately account for information control (as in control of that which has become) and information retrieval (as in the engagement with “becoming”), not just in a practical sense but also in our philosophy. Even this brief example shows that if we are set out to create a new philosophy of LIS then we must address the question of reductionism. (Jones 2008, 484)

Pour mieux comprendre la perspective qui est celle de l’analyste documentaire, nous proposons une démarche à la fois pragmatiste (chercher à obtenir des résultats optimaux pour le repérage au sein du système documentaire) et inclusive, se traduisant par une attention portée à la dimension objective du sujet aussi bien qu’aux intérêts des usagers. De manière plus précise, le contrôle de l’information permettrait en quelque sorte de *fixer* la signification de l’objet que l’analyste a « entre les mains » (le document ou « that which has become »), lui donnant licence par le fait même d’utiliser une expression indexicale (ceci est un « Manuel d’astronomie » ; voir la section 1.1). Le repérage de l’information, toujours du point de vue de l’analyste, fournirait un objectif en devenir (« engagement with becoming ») dont l’orientation serait déterminée par les ambitions ou visées des usagers (s’adonner à l’astronomie en amateur, s’astreindre à en étudier les principes, se préparer à écrire un article de vulgarisation scientifique, etc.). Selon nous, la réduction sémantique doit être examinée de plus près pour comprendre les enjeux d’une approche plus inclusive (*i.e.* suggérant une plus grande complémentarité entre la dimension objective du sujet et les raisons qui poussent un usager à consulter la documentation).

Si l’on segmente l’analyse documentaire en tant qu’activité, réductionnisme et subsomption (placer un contenu spécifique, tiré d’un document, sous une classe ou une catégorique générique) semblent à première vue assimilables, notamment en ce qui a trait à la description du noyau thématique du document (*aboutness*). Le balayage du contenu (*i.e.* la lecture documentaire), la sélection des concepts importants et leur représentation sous forme de métadonnées produisent une sorte de succédané du document (*document surrogate*) : « For example, a bibliographic description or a metadata description is a type of reduction that can stand for a text within certain limits ; the key phrase seems to be this one that says “within certain limits” » (Jones 2008, 487). Les limites auxquelles Jones fait allusion sont traditionnellement définies par un ensemble de règles qui ont été développées, comme nous l’avons montré dans la section précédente, dans une optique fortement pragmatiste. La

subsomption contribue fortement, selon nous, à l'équilibre que l'analyste cherche à obtenir entre les mesures de précision et de rappel dans un système documentaire<sup>23</sup>. C'est en bonne partie pour éviter les problèmes bien connus de silence et de bruit que la subsomption intervient, permettant de contrôler à la fois la spécificité et l'abondance des métadonnées utilisées dans l'analyse documentaire.

Sur le plan linguistique, la notion de réduction se rattache avant tout à la subsomption mentionnée plus haut. En d'autres mots, un ensemble de procédés qui vise à placer des contenus spécifiques sélectionnés sous des catégories ou des classes plus générales (subsomption) semblent éminemment réductionniste. La réduction sémantique permettrait d'obtenir, par analyse et synthèse successives, une expression textuelle considérée comme homologue (équivalence dans la signification du contenu d'un document et celle d'un ensemble de métadonnées). Dans un passage qui évoque la lecture documentaire, Beghtol parle d'un processus de subsomption pour arriver à une séquence de « macropropositions » susceptible de résumer le(s) sujet(s) du document :

In this analytic and synthesising process, sequences of sentences with their underlying logical propositions are cognitively compressed and summarised so as to produce a hierarchically-governed sequence of macropropositions (i.e. overriding logical propositions) that express the sum of the meanings of the propositions subsumed under them in the same way a series of co-ordinate sub-classes is assumed to be logically subsumed under its superordinate class. The end product of this continual text reduction is called the discourse topic or the aboutness of the text. (Beghtol 1986, 90)

Nous revenons de cette manière à la notion d'*aboutness*, considérée cette fois comme un résumé qui assure une certaine correspondance entre la *somme* des contenus d'un document (*meanings*) et la description qui en est donnée dans un système documentaire. Aussi, doit-on s'interroger à présent sur la nature et la véritable fonction d'une réduction sémantique qui contribue à la tendance objectiviste de l'approche centrée sur le document.

---

<sup>23</sup> Le contrôle du vocabulaire, comme solution pour éviter les problèmes de synonymie et homonymie, améliorerait le rappel et la précision lors de la recherche documentaire. Pourtant, cette « évidence théorique » n'a jamais été prouvée empiriquement, pas même avec le « projet Cranfield » (Menon 2007, 21).

Le processus de réduction va au-delà d'un simple rapport d'équivalence avec la subsomption. En s'inspirant d'une définition communément admise en philosophie des sciences, il est possible de mettre en lumière de quelle manière le réductionnisme contribue à une approche centrée sur le document. La notion de réduction, dans une optique épistémologique classique, implique qu'une théorie A est en mesure de reproduire et assimiler les explications d'une théorie B, qui est ainsi dite « réduite » (Ludwig in Barberousse, Bonnay et Cozic 2011, 235)<sup>24</sup>. Le processus d'assimilation du contenu cadre bien avec le réductionnisme textuel que Beghtol décrit, quoiqu'avec une certaine réserve. Les particularités de l'analyse documentaire nous obligent à apporter une distinction et à introduire la variante « hétérogénéité » dans le processus de réduction :

Dans d'autres cas, en revanche, il ne va pas de soi qu'on puisse opérer des connexions conceptuelles interthéoriques. On doit alors parler de réduction hétérogène, car le vocabulaire de la théorie à réduire n'est pas inclus dans celui de la théorie réductrice. (Ludwig in Barberousse, Bonnay et Cozic 2011, 236)

On peut s'inspirer de cette mise en garde pour questionner la nature des métadonnées sujet et, surtout, le lien qu'elles entretiennent avec le contenu effectif de la documentation. Ce lien s'exprime au moins en partie à travers la notion de caution bibliographique, centrale dans l'analyse documentaire depuis son introduction par Hulme<sup>25</sup>, dans la mesure où les langages documentaires se développent généralement à partir du vocabulaire présent dans la documentation. L'analyste documentaire, lorsqu'il fait appel à ce type d'outils, fournit des indices (métadonnées sujet) issus d'un processus réductionniste hétérogène qui permet de représenter l'essentiel d'un contenu complexe (tel que celui-ci est communiqué dans la documentation) par l'entremise d'un langage balisé (contrôlé). Il faudrait par conséquent

---

<sup>24</sup> Pour reprendre ceci dans le contexte de l'analyse documentaire, il faut transposer la notion de théorie en des termes quelque peu différents. La théorie A (instance réductrice) serait le langage documentaire, alors que la théorie B (explication réduite) correspondrait au contenu du document. Le rapprochement s'éclairera dans le développement qui suit.

<sup>25</sup> L'approche centrée sur le document, dans son ensemble, repose sur l'important principe que Hulme a placé au cœur de la classification documentaire : « Hulme coined the phrase 'literary warrant' to mean that the schedules of a classification should be moulded to fit published literature, being dictated by what actual (and if possible potential) literature itself merits or warrants » (Sayers et Maltby 1975, 56). La caution bibliographique ajoute donc une dimension supplémentaire à la recommandation qui, dans une approche centrée sur le document, prescrit de se limiter à la signification intrinsèque du document dans l'analyse (voir la section 1.2 et le début de la présente section).

parler d'une réduction inter-idiomatique (plutôt qu'interthéorique), le langage documentaire agissant comme outil pour assimiler et organiser le contenu de la documentation à l'intérieur d'un environnement langagier « exogène ».

Nous pouvons dès maintenant résumer comment le réductionnisme s'exprime dans le contexte de l'analyse documentaire et de quelle manière il contribue à améliorer l'organisation de la documentation. La réduction sémantique décrit la transition du contenu d'un document à son expression indexicale (voir la section 1.1) dans un système documentaire (le plus souvent à l'aide d'un langage documentaire). Dans l'optique d'une approche centrée sur le document, cet indice traduit de manière objective le noyau du sujet (*aboutness*). Rappelons que ce processus, hétérogène puisqu'il assure le passage d'un encadrement (document) vers un autre (système documentaire), escamote les caractéristiques sémantiques trop complexes ou trop particulières pour aller à l'essentiel :

A further step is that of excising irrelevancies and overrefinements (irrelevant and overrefined from the user's point of view), a job very painful to the specialist, for whom the details and refinements are just what causes pride. But excising irrelevancies and overrefinements is at the same time recognizing or creating relevancies and utilities; recognizing relevancies is finding or establishing a connection between a task and a piece of knowledge; and recognizing overrefinement is recognizing the limits of applicability of knowledge. (Wilson 1977, 32)

La nature hétérogène de cette réduction vient de la nécessité d'établir un lien entre un sujet dans sa dimension objective (ce qu'on pourrait rapprocher de l'expression « piece of knowledge » dans la citation précédente) et une tâche (« task ») subjective telle que la recherche documentaire. L'appariement, rappelons-le, est l'objectif visé en ce qu'il actualiserait ce lien sur le plan du vocabulaire et permettrait, en théorie, d'obtenir des résultats pertinents lors de la recherche documentaire.

Rappelons que l'analyse documentaire possède un statut particulier parce que son orientation est déterminée par des considérations pratiques étrangères (au moins en partie) au travail de l'auteur du document analysé (Šauperl 2002, 142). Cette séparation est l'une des raisons pour lesquelles l'analyste documentaire suit une procédure et adopte des outils qui assurent une certaine forme d'autonomie. Plutôt que de proposer un argumentaire sur un sujet

donné (ou énoncer une théorie), l'instance réductrice (langage documentaire) doit proposer des indications pertinentes (généralement contrôlées et structurées) susceptibles d'améliorer la recherche documentaire et de faciliter, par le fait même, les activités des usagers. Le langage documentaire fournit une assistance stratégique dans l'organisation de la documentation, sous la forme d'une structure sémantique :

It [*structuring vocabulary*] must help the searcher construct the best possible strategy in terms of a user's needs (high recall, high precision, or some compromise between the two). In particular, it should prevent a searcher from missing relevant documents through failure to bring together semantically related terms. (Lancaster 1986, 151)

La réduction hétérogène effectuée par l'analyste documentaire doit permettre d'établir des liens sémantiques, entre autres par la subsomption, afin d'offrir la meilleure organisation possible des collections documentaires sur le plan thématique (en choisissant, par exemple, un indice de classification dont la précision reflète bien le contenu du document et son usage potentiel ; voir l'exemple du « Manuel d'astronomie » au début de la présente section). L'objectif est d'appuyer les démarches des usagers tout en anticipant l'impact d'activités telles que la gestion du système documentaire (ex. : une métadonnée affecte différemment les mesures de rappel et de précision à mesure que les collections documentaires évoluent), l'interrogation du catalogue (ex. : visualisation des métadonnées) et le butinage (*browsing*) des collections documentaires.

Pour Svenonius, le contrôle du vocabulaire (*vocabulary control*) est au cœur d'une organisation efficace de l'information :

Vocabulary control is the *sine qua non* of information organization. Information is not organized if it is scattered or if its collocation is cluttered. A natural language cannot be used to organize information effectively because its synonymy and homonymy would cause scatter and clutter. Vocabulary control controls for synonymy and homonymy by normalizing names of attributes and entities. It is the means by which a one-to-one correspondence is established between terms and their referents. The imposition of vocabulary control creates an artificial language out of a natural language. (Svenonius 2000, 89)

Le contrôle du vocabulaire s'appuie sur une sémantique référentielle (*referential semantics* ; Svenonius 2000, 57) que nous avons déjà abordée plus haut (voir la section 1.1). Le

présupposé central sur lequel s'appuie la validité d'un tel contrôle se résume par la valeur *pratique* de la normalisation du vocabulaire, laquelle peut être déduite à défaut d'être prouvée empiriquement (Menon 2007, 21) :

Without normalization there will be recall, precision, and navigation failures in retrieval. *This does not have to be demonstrated empirically*, since it is logically deducible from the fact that natural languages contain synonymy and homonymy at the level of both vocabulary and sentence structure. Ridding a natural language of its surface-level anomalies amounts to reducing the variety exhibited by phenomenal language to its seminal or deep structure. *The reductive analysis pursued in linguistics for the purpose of explaining and generalizing about natural languages, in bibliographic practice, serves the more practical purpose of intelligent information organization.* (Svenonius 2000, 192 ; nous soulignons)

L'analyse documentaire centrée sur le document emprunte un point de vue pratique dont l'objectif principal est de fournir une représentation normalisée des collections. La réduction sémantique est nécessaire pour exercer un contrôle qui, de longue date, a servi à construire une instance langagière (un langage documentaire en d'autres mots) susceptible de pallier les imprécisions des langues naturelles et d'améliorer l'organisation de la documentation.

En parallèle du contrôle du vocabulaire (parfois intégrée à celui-ci), la structure des langages documentaires permet de préciser les liens sémantiques entre les différents termes qui y figurent. Il s'agit d'un aspect important de la réduction sémantique hétérogène que de favoriser un vocabulaire non seulement clair, mais ordonné de manière à favoriser les rapprochements sur le plan sémantique (c'est l'essence même de la classification documentaire que de rapprocher des documents dont le contenu témoigne d'une parenté disciplinaire et/ou thématique). Contrôle et structure sont interdépendants dans le contexte de l'approche centrée sur le document, puisque celle-ci a recours, comme nous l'avons déjà mentionné, à un langage documentaire pour contrôler *et* structurer les métadonnées sujet. Par exemple, la relation « genre-espèce » appuie une démarche réductionniste qui mise sur un système mieux organisé que les langues naturelles, dont l'imprécision pourrait nuire à la création de relations stables entre termes (même une fois ceux-ci soumis à une forme de contrôle sémantique) :

If the vocabulary of a discipline is soft and includes many terms with linguistically indeterminate referents, it is almost useless to attempt to

structure it by genus-species type relationships. Nor does it serve much purpose to force abstract terms like Love, Truth, and Beauty (or for that matter, terms whose meaning are largely dependent on context) into Procrustean genus-species structures. (Svenonius 2000, 166)

Contrôle et structuration sémantiques se renforcent mutuellement, mais dans une perspective qui privilégie grandement l'aspect objectif du sujet. Dans l'optique d'une approche centrée sur le document, une terminologie teintée de connotations et d'usages divers ou ponctuels n'est pas propice à la représentation d'un sujet documentaire, parce qu'elle ne peut être adéquatement contrôlée et ordonnée.

Bien entendu, il faut préciser que différents langages documentaires sont basés sur des structures développées à divers degrés :

Most classifications aim to map a knowledge domain and, in so doing, to show the location of each term relative to all others. This is beyond the ambition of most thesauri. The difference this makes to structure is that the hierarchies in a classification coalesce to form a gigantic upside-down tree, whereas those in a thesaurus resemble more a scattering of small shrubs. (Svenonius 2000, 167)

Malgré une certaine diversité, les langages documentaires « classiques » utilisent des relations sémantiques qui confèrent une certaine cohérence interne dans l'organisation des métadonnées sujet et améliorent la navigation d'un sujet à l'autre. Une fois de plus, un présupposé *pratique*, combiné à une forme de prudence *professionnelle* qui recommande l'usage d'un langage documentaire (Menon 2007, 21), rend la structuration pratiquement incontournable. Le succès des classifications documentaires cautionne au moins en partie une réduction sémantique qui, par une structuration et un contrôle constants, mise sur le conservatisme et la force de l'habitude (habitudes de pensée ; voir la section 1.2) aussi bien chez l'utilisateur que le bibliothécaire (Richardson 1964, 34). La réduction sémantique inter-idiomatique assure la transition (jugée nécessaire selon la thèse forte en bibliothéconomie) vers un système documentaire indexical susceptible d'améliorer l'expérience de recherche documentaire, proposant à l'utilisateur des métadonnées à la fois précises, neutres, monosémiques et stables.

Dans le premier chapitre, nous avons identifié la dimension objective du sujet, le pragmatisme documentaire et la réduction sémantique en tant que piliers de l'approche centrée sur le document. Nous avons choisi comme démarche de mettre en évidence le propos de

commentateurs qui ont su repérer certaines tendances fortes en analyse documentaire. Aussi, nous avons insisté sur des notions (*aboutness*, pragmatisme) qui peuvent paraître moins évidentes ou moins bien implantées dans la littérature en sciences de l'information. Celles-ci nous paraissent néanmoins incontournables pour décrire le point de vue duquel se place l'analyste pour offrir le meilleur service possible aux usagers. La bonification des services techniques (catalogage, analyse documentaire, préparation matérielle des documents, etc.) doit toujours figurer parmi les préoccupations d'une institution documentaire. Nous verrons, dans la section 3.1, qu'une meilleure compréhension du contexte dans lequel se pratique l'analyse documentaire universitaire doit être envisagée pour mieux comprendre les besoins des usagers. Avant cela, nous examinerons dans le deuxième chapitre de ce mémoire l'approche centrée sur l'utilisateur (d'un point de vue antithétique) et ce qu'elle peut apporter à l'analyse documentaire dans le contexte universitaire actuel.

## **Deuxième chapitre**

**Approche centrée sur l'utilisateur**

Notre objectif dans le premier chapitre du mémoire était d'examiner l'approche centrée sur le document et d'en relever les éléments principaux (dimension objective du sujet, pragmatisme documentaire et réduction sémantique). Il s'agissait d'une étape incontournable pour mettre en lumière aussi bien les limites que le potentiel de l'analyse documentaire. Dans le contexte actuel (on peut penser au développement du Web 2.0), examiner l'approche centrée sur l'utilisateur s'impose aussi comme une tâche nécessaire pour comprendre pleinement le rôle que peut jouer l'analyste documentaire. Notre objectif dans ce deuxième chapitre est de souligner les avantages d'une approche centrée sur l'utilisateur pour l'enrichissement des métadonnées sujet dans le milieu universitaire.

Hudon définit l'« analyse orientée usager » de la manière suivante : « Analyse documentaire qui tient compte non seulement du contenu du document, mais également des besoins et des comportements des usagers auxquels le document sera utile » (Hudon 2013, 264). Nous pouvons voir que l'accent est mis tout autant sur les particularités et les attentes des usagers que sur le contenu de la documentation. Pour combler les besoins des usagers, il faut au moins en partie répondre à leurs attentes, si bien que de solliciter leur contribution semble être la voie à privilégier pour faciliter le repérage de l'information. D'ailleurs, cette stratégie, sous diverses formes, semble remporter de plus en plus la faveur des bibliothèques et autres centres de documentation. Cette tendance cadre bien avec la dimension subjective du sujet (section 2.2) et nous invite à repenser le rôle que peut remplir la *folksonomie* (section 2.3) à l'intérieur d'un cadre relativiste (section 2.1). Nous allons par conséquent examiner quelle place peut prendre, sur le plan du vocabulaire en usage, la contribution des usagers dans un système documentaire.

L'approche centrée sur l'utilisateur, plus malléable que celle centrée sur le document, résorbe la dimension objective du sujet de manière à libérer les possibilités de représentation du contenu. Elle peut servir d'antithèse à une démarche plus traditionnelle qui se base sur le pragmatisme documentaire (section 1.2) et la réduction sémantique (section 1.3) de manière à offrir une représentation du contenu partagée par le plus grand nombre. C'est d'abord à travers la notion de relativisme qu'il devient souhaitable de questionner la validité d'un postulat consensuel (voir la section 1.2) et d'établir l'importance d'un langage documentaire relativiste, la *folksonomie*. De ce point de vue, nous devons circonscrire l'approche centrée sur l'utilisateur

de manière plus spécifique que la définition qui est généralement acceptée. Nous entendons par « approche centrée sur l'utilisateur » une analyse documentaire qui se base sur la contribution terminologique des usagers. Il s'agira donc pour nous d'explorer et définir une approche qui répertorie les termes utilisés par les usagers (démarche *folksonomique*) pour témoigner de la signification qu'ils accordent à un texte ou à un énoncé. Notre argumentaire se justifie à raison de la dimension subjective du sujet évoquée par Maniez (2002, 123 ; nous y reviendrons un peu plus bas) et du potentiel qu'offre une *folksonomie* pour prendre en compte les intérêts réels des usagers en milieu universitaire. Par conséquent, cette forme de parti pris nous pousse avant tout à examiner la dimension relative de la connaissance et l'aspect subjectif du sujet, en gardant en tête le milieu universitaire comme toile de fond.

Résumons à nouveau les trois articulations principales de ce second chapitre afin d'apporter une précision supplémentaire. Nous allons établir un cadre théorique à partir de la notion de relativisme (section 2.1) dans le but d'éclairer la dimension subjective du sujet (section 2.2) et définir une *folksonomie* adaptée au milieu universitaire (section 2.3). Il ne s'agit donc pas d'une réflexion épistémologique à proprement parler mais d'un argumentaire qui s'appuie sur une littérature relativement abondante en sciences de l'information. Néanmoins, le sujet de ce chapitre invite une panoplie de questions qui dépassent le cadre de la bibliothéconomie ou même des sciences de l'information. Notre objectif général est de déterminer dans quelle mesure une approche centrée sur l'utilisateur peut orienter la pratique de l'analyse documentaire en milieu universitaire. Par extension, le rôle de l'analyste documentaire, à travers les métadonnées sujet qu'il propose, pourrait être élargi dans le contexte universitaire afin de favoriser une communication (indirecte<sup>26</sup>) plus riche et une interaction toujours croissante avec les usagers.

---

<sup>26</sup> Nous verrons un peu plus bas ce que signifie, dans le cadre de l'analyse documentaire, une communication de type indirect. Par ailleurs, nous avons vu avec Metcalfe (1957, 213) que, dans une approche centrée sur le document, l'analyse documentaire doit se limiter à un registre indicatif plutôt que de s'investir dans une communication complexe. Bien entendu, une approche centrée sur l'utilisateur ne se base pas sur un tel principe.

## 2.1) Le relativisme dans le contexte de l'analyse documentaire en milieu universitaire

Le sujet représente, selon la définition utilisée à la section 1.1, un objet sur lequel s'applique la réflexion d'un individu (Maniez 2002, 123). Cette dimension réflexive nous invite à questionner la place que doit occuper l'évaluation de cette entreprise individuelle dans le processus d'analyse documentaire. S'il était possible de connaître ce qui motive les pensées de chacun, il pourrait être plus facile pour l'analyste documentaire de pressentir l'intérêt que les usagers manifestent pour un document donné. Il s'agit par contre d'une tâche incommensurable. Notre hypothèse est que l'intérêt manifesté par un usager, largement déterminé par le statut particulier qu'occupe la connaissance dans les activités de celui-ci en milieu universitaire, trouve sa source dans des énoncés de recherche, dont le vocabulaire serait le matériau idéal pour le développement d'une *folksonomie*. Avant d'aboutir à une telle conclusion, nous devons examiner ce qui motive notre hypothèse, c'est-à-dire le statut particulier (privilegié) de la notion de connaissance en milieu universitaire ainsi que sa dimension relativiste.

La notion de relativisme étant complexe à traiter, cette section est divisée en trois sous-sections de manière à en faciliter la lecture. Dans la sous-section 2.1.1, nous examinons la connaissance dans une optique relativiste générale. Ensuite, nous introduisons la notion de *knowledge claim* (que nous traduisons par « prétention au savoir ») pour préciser un type de communication qui pourrait grandement intéresser l'analyste documentaire en milieu universitaire (sous-section 2.1.2). Finalement, nous prenons comme exemple l'énoncé de recherche (sous-section 2.1.3) pour illustrer le potentiel d'une position relativiste dans la compréhension de l'intérêt qu'un usager en milieu universitaire manifeste pour un document. Cette section a pour fonction d'introduire des notions (relativisme, *knowledge claim*, énoncé de recherche) susceptibles d'apporter des éléments de réponse supplémentaires à la problématique de l'appariement (laquelle peut difficilement être résolue à l'intérieur d'une approche centrée sur le document). Nous souhaitons établir une base solide à partir de laquelle il sera possible d'éclairer la dimension subjective du sujet (section 2.2) et de proposer un type particulier de *folksonomie* (section 2.3), élaboré en fonction des intérêts de recherche des usagers en milieu universitaire.

### **2.1.1) La connaissance dans une optique relativiste**

Sujet vaste et difficile, la connaissance se situe au cœur de l'analyse documentaire peu importe le milieu dans lequel celle-ci est pratiquée. L'université semble présenter un environnement particulièrement intéressant pour examiner comment se communique et s'exploite la connaissance. Une dimension pragmatiste (voir la section 1.2) ne peut être ignorée, même lorsqu'il est question d'une approche centrée sur l'utilisateur. La dimension relativiste de la connaissance nous invite simplement à regarder plus précisément le statut de la connaissance dans le contexte de l'analyse documentaire. Représente-t-elle une sorte d'édifice immuable (connaissance avérée) ou est-elle plutôt le résultat d'une tâche individuelle et nécessairement subjective ? Lalande donne une définition claire et simple du relativisme : « Doctrine qui admet que toute connaissance (ou toute connaissance humaine) est relative » (Lalande 2010, 914). De manière générale, la connaissance serait relative dans la mesure où elle dépend du point de vue d'un individu ou d'un groupe d'individus. Dans quelle mesure l'analyste documentaire en milieu universitaire peut-il tirer profit de cette assertion pour orienter son travail et répondre aux attentes des usagers ?

La documentation joue un rôle important dans le partage des connaissances, cependant il faut préciser que la communication dans le contexte universitaire dépasse les limites du circuit documentaire : « Scholarly communication is a complex web of connections among scholars and researchers, authors and readers, formal publications and libraries » (Budd 1998, 79). Quel peut être le rôle d'un organe documentaire ? Avec sa plume caractéristique, Bliss nous invite à réfléchir à la place distinctive qu'occupe la bibliothèque au sein de l'université :

A university is sometimes metaphorically called a temple of learning. But the great modern university is more like a city; and its citadel is its library. This we might call a temple of knowledge, or of virtual learning. It is as comprehensive as the university. Its classification should be consistent with the organization of knowledge, with the system of the sciences, and so it should have higher educational value. (Bliss 1929, 115)

La valeur éducative de l'analyse documentaire est déterminée par les indications justes qu'elle peut fournir à l'étudiant qui « déambule dans la cité universitaire ». Rappelons que l'analyse documentaire se place traditionnellement dans une optique pragmatiste et qu'elle cherche à bonifier l'expérience de recherche documentaire. Pour Bliss, l'analyse documentaire est utile

lorsqu'elle se rallie au système des sciences et conséquemment qu'elle adopte, comme critère objectif pour l'organisation des connaissances, le consensus scientifique et éducationnel. Le pragmatisme de John Dewey, influence principale de Bliss, fonde ce critère objectif sur la nécessité de l'action (soit l'apprentissage ou la recherche dans le cas qui nous intéresse) : « The necessity of execution supplies objective criteria. Things have to be sorted out and arranged so that their grouping will promote successful action for ends » (John Dewey cité dans Hjørland 2008a, 97). Néanmoins, il est loin d'être admis que les activités d'apprentissage ou de recherche d'un individu, d'un groupe ou d'une communauté puissent être réduites à un critère objectif et, par conséquent, à un nombre limité d'objectifs partagés. Pour plusieurs, dont le philosophe Michel Foucault, ce type de réduction s'inscrit dans une forme de « fantasmagorie de la bibliothèque » (Radford 1992).

Hérité d'une philosophie pragmatiste classique, le postulat consensuel avancé par Bliss doit être remis en cause dans la mesure où le consensus éducationnel semble de moins en moins atteignable aujourd'hui, ou même souhaitable dans une perspective relativiste. Nous pensons que cette critique se base avant tout sur le constat suivant :

Libraries exist to serve their users, but the user population is increasingly heterogeneous. The days when the academic library could assume that its users were either scholars pursuing their research and teaching interests or full-time undergraduates straight from school have vanished. Today's user may be full- or part-time, may be studying at a distance, perhaps on a course franchised to a local college, may be pursuing funded research as part of an international team or may be a part-time teacher brought in to lead a specific module. (Brophy 2005, 73-74<sup>27</sup>).

Peut-on envisager un langage documentaire dont les caractéristiques apporteraient satisfaction à cet ensemble disparate d'utilisateurs ? D'une manière ou d'une autre, il n'est pas superflu d'adopter un point de vue critique afin de favoriser le développement futur de l'analyse documentaire, surtout à l'ère du Web 2.0 et du catalogue « nouvelle génération » (*next-gen*). La connaissance, telle qu'elle se développe et s'exploite dans le milieu universitaire, requiert une approche de l'analyse documentaire capable d'intégrer divers points de vue (à l'image de la diversité des utilisateurs en milieu universitaire).

---

<sup>27</sup> Voir aussi Budd 1998, 283 et Alemu *et al.* 2012, 334-335.

L'approche relativiste met l'accent sur la connaissance en tant que croyance personnelle (Budd 2001, 126-127). Même si une connaissance peut (évidemment) être partagée par plusieurs, elle demeure le résultat d'une activité cognitive subjective. Cela veut dire qu'un individu se représente, suivant l'intérêt qu'il manifeste pour un sujet, une proposition subjective qu'il considère vraie et fiable. Par exemple, un usager peut formuler l'énoncé suivant dans le but de mieux comprendre un sujet : « la baleine est un mammifère marin ». D'après l'état de nos connaissances, cet énoncé peut être considéré vrai (il reflète une réalité observée) et fiable (il reflète une opinion fondée ou un jugement cohérent, *i.e.* non contradictoire). D'ailleurs, nous avons là, soumises à nombre de débats philosophiques, deux conditions nécessaires (vérité et fiabilité) pour conférer le statut de connaissance à une proposition :

How does the library enable or facilitate an individual coming to some sort of justified true belief? Here again, philosophers debate the means by which claims can be justified. Two particular ways are pertinent as we consider the library's role in knowing—veritism (this term comes from Goldman, but others have picked up on it) and reliability (this idea has been around for some time in connection with individual epistemology). (Budd 2004, 364)

Cette citation nous permet d'entrevoir l'impact que peut avoir la bibliothèque dans le processus d'acquisition de connaissances des usagers. Pour reprendre l'exemple plus haut, la vérité et fiabilité d'un énoncé tel que « la baleine est un mammifère marin » doivent être prises en compte dans la structure d'un schéma de classification lorsque l'état de nos connaissances le justifie (on pourrait appeler cela la « caution épistémique », sous-jacente à la caution bibliographique). Dans une perspective relativiste, un ensemble de métadonnées sujet pourraient bien être soutenu par des propositions subjectives suffisamment répandues pour en justifier l'utilisation lors de l'analyse documentaire<sup>28</sup>. De ce point de vue, nous devons nous poser la question suivante : la bibliothèque universitaire a-t-elle un rôle à jouer dans l'évaluation critique d'une proposition subjective ou d'une croyance personnelle ?

---

<sup>28</sup> Par extension, la pertinence de métadonnées résultant de la contribution des usagers est souvent déterminée par le nombre d'occurrences d'un terme (dans une *folksonomie* par exemple). Les nuages de mots-clés (*tag clouds*) montrent bien que le « poids » d'un terme représente un indice possible de l'intérêt des usagers pour un sujet ou un thème (voir Moulaison 2008, 104)

Pour évaluer l'impact que peut avoir l'approche relativiste dans le milieu universitaire, il faut d'abord examiner la notion de connaissance selon un point de vue général. Nous devrions par la suite être en mesure d'évaluer l'analyse documentaire en fonction d'objectifs propres au milieu universitaire (parmi lesquels nous plaçons l'acquisition, l'évaluation, le partage et le développement des connaissances). Comme nous l'avons suggéré plus haut, deux conditions sont nécessairement requises pour qu'une proposition relève de la connaissance avérée : vérité et fiabilité. Obéissant à un idéal de neutralité (relevant du « positivisme bibliothéconomique » selon Radford 1992, 412), la bibliothèque universitaire ne juge pas de la valeur du contenu de la documentation au-delà du processus de sélection. Elle laisse plutôt aux usagers (étudiants, membres du corps professoral, etc.) le soin de déterminer si un énoncé (« la baleine est un mammifère marin ») est susceptible ou non de contribuer à nos connaissances<sup>29</sup>. Il y a là une distinction majeure qui, dans le contexte de la gestion des langages documentaires, peut s'exprimer par la caution bibliographique : un terme, simple ou complexe, est inclus dans ce type de vocabulaire parce que son usage répandu dans la littérature le justifie et non parce qu'il témoigne de la vérité d'une proposition. En termes bibliothéconomiques, cela veut aussi dire que la connaissance doit être dite subordonnée aux intérêts présumés des usagers, aux exigences de l'organisation des collections et à l'efficacité du système documentaire en place (toujours en lien avec la position pragmatiste examinée dans la section 1.2). Néanmoins, certains aspects épistémiques doivent être considérés pour assurer la qualité des services offerts par la bibliothèque universitaire.

Démontrer la vérité d'un énoncé demeure donc extérieur au mandat du bibliothécaire, l'accès aux documents étant traditionnellement l'objectif principal visé par ce dernier (Radford 1992, 412). Les usagers attendent de leur institution documentaire qu'elle facilite le repérage de l'information (*information retrieval*), ce dont témoignent l'« orientation système » et les mesures bien connues de rappel et précision :

The library is perceived as system- rather than person-oriented. It concerns itself with catalogs, indexes, and other information retrieval technologies that enable it

---

<sup>29</sup> L'inclusion (ou l'exclusion) d'un document dans la collection représente bien une forme de jugement de valeur, mais la sélection documentaire se base davantage sur les besoins de la communauté universitaire (Budd 2004, 365) que sur une méthode scientifique visant à démontrer la vérité d'un ensemble de propositions contenues dans un document.

to represent and retrieve specific texts more effectively, and its success is quantifiable using such standard measures as precision and recall. (Radford 1992, 413)

L'analyse documentaire contribue directement au fonctionnement de ces outils de repérage, mais elle doit aussi être considérée comme une activité intellectuelle, *i.e.* une tâche qui implique une certaine forme de subjectivité dans l'analyse du contenu. De ce point de vue, Budd nous rappelle que l'accès ne peut être complètement dissocié de la notion de contenu :

It is true that the profession of librarianship has concerned itself with matters of access and that access is essential to the services a library offers. Access, however, is not a discrete and separate concept that can be applied without consideration of some other vital factors. When a library provides access, it is providing access to what? Access is an intellectual entity and activity; as such it is inextricably linked to content. It is impossible to provide access to *something* without knowing what that *something* is about. (Budd 1998, 279 ; l'auteur souligne)

L'enjeu d'une analyse relativiste n'est pas simplement de déterminer l'*aboutness* du document, mais bien d'améliorer les possibilités de description du contenu pour assurer un accès moins restrictif. Ceci pourrait faire du repérage de l'information une activité qui nécessite de placer l'appariement au cœur d'une stratégie plus inclusive, attentive aux objectifs que visent les usagers lorsqu'ils consultent la documentation (voir la section 1.3 pour un commentaire sur la possibilité d'une démarche inclusive ; Jones 2008).

Si l'on cherche à favoriser une approche relativiste, la communication doit compter parmi les aspects que cherche à améliorer le bibliothécaire. Au-delà de l'exigence d'indication soulignée par Metcalfe et de la mise en garde concernant la communication (voir la section 1.1), l'analyse documentaire pourrait s'inspirer plus directement de la complexité et du foisonnement caractéristiques des échanges (écrits aussi bien que verbalisés) en milieu universitaire. D'ailleurs, les services documentaires sont établis en fonction des exigences de la communication (directe et indirecte) entre usagers et bibliothécaires au même niveau que la gestion des collections :

In part service is a function of the collections and bibliographic apparatus of the library. In large part it is a function of communication, both direct and indirect, between librarian and user. The indirect communication may come about through the library's catalog, its classification, its physical layout, and its signage. (Budd 1998, 306-307)

La communication en bibliothèque est avant tout interpersonnelle (lors de l'entretien de référence par exemple) et directe (Budd 1998, 307). L'analyse documentaire doit plutôt être rapprochée de la fonction communicative indirecte qu'assume un organe documentaire et, dans sa forme relativiste, elle favorise une forme de liberté dans la représentation des connaissances. Assurer l'accès à la documentation par le biais d'une communication moins balisée favoriserait, nous semble-t-il, le développement des compétences évaluatives des usagers, surtout en ce qui a trait à la fiabilité d'un énoncé (deuxième condition épistémique mentionnée plus haut). L'importance accordée à l'évaluation du contenu d'un document entraîne, dans une perspective relativiste, la nécessité de recourir à une communication qui aurait comme base une variété de jugements subjectifs. Une telle position met l'accent sur le caractère incertain et faillible de la connaissance, amoindrissant du même coup la nécessité de rechercher un consensus (voir la section 1.2). Par conséquent, nous nous plaçons dans une perspective que l'on pourrait appeler faillibiliste (*i.e.* la connaissance est toujours sujette à une révision).

En examinant les bases d'une approche centrée sur l'utilisateur (le relativisme ici), nous nous dirigeons inéluctablement sur le terrain très actuel de la « décentralisation » des moyens de production de contenus (avec le Web 2.0 entre autres). Sur le plan de l'analyse des contenus, l'absence de contrôle d'autorité fait pendant à cette décentralisation et accorde une place importante à la subjectivité dans la production des métadonnées sujet. Cette réalité peut modifier l'expérience des usagers lors d'une recherche documentaire aussi bien que lors de l'évaluation de la pertinence des documents repêchés<sup>30</sup>. Plutôt qu'une autorité centralisée (incarnée dans le langage documentaire, ou liste d'autorité, utilisé par l'analyste documentaire), c'est une diversité de points de vue qui vient alimenter la description des documents, *i.e.* la représentation de leur contenu. Sur le plan épistémique, le relativisme se rapproche du faillibilisme puisqu'il promeut un processus de validation rétroactif, faisant de la connaissance une trajectoire plutôt qu'un point d'arrivée :

---

<sup>30</sup> Les suggestions de liens vers des documents similaires (« You might also like... » ; « People who liked this also liked... ») font maintenant partie intégrante de l'expérience des internautes, parmi lesquels on doit compter les usagers en bibliothèque.

Et premièrement, cet élément crucial : la connaissance est une trajectoire, ou un vecteur qui projette rétroactivement en avant son « pouvoir de validation ». En d'autres termes, nous ne savons pas encore, mais nous saurons, ou plutôt, nous saurons si nous avons su auparavant ou non. La certification rétroactive, ce que Gaston Bachelard appelait « rectification », est l'essence du savoir. (Debaise 2007, 20)

Une conception relativiste de la connaissance se nourrit de l'incertitude épistémique et se réclame de théories qui favorisent une diversité de points de vue. Une « rectification » est toujours possible, ce qui suggère, plutôt qu'un processus de validation fortement balisé (et consensuel), une ouverture épistémique suffisamment large de manière à permettre une multitude de corrections potentielles.

### **2.1.2) Knowledge claims**

La dimension faillible des connaissances façonne la communication de manière particulière, introduisant ce qu'on pourrait appeler des « prétentions au savoir » (*knowledge claims*). L'une des notions les plus intéressantes pour approfondir l'approche relativiste en analyse documentaire, la prétention au savoir est issue d'une épistémologie sociale dont Shera a fait la promotion dès le milieu du 20<sup>e</sup> siècle (Shera 1966, 67). On remarque dans les deux dernières décennies une utilisation plus répandue en sciences de l'information du terme *knowledge claims* (chez Hjørland, Radford et Budd entre autres). Cette démarche théorique souligne l'importance du critère de fiabilité (évaluation de la cohérence d'un énoncé ; voir la sous-section 2.1.1) et invite le plus souvent à analyser la notion de connaissance d'un point de vue faillibiliste, antipositiviste :

We should not talk about knowledge or knowledge organization, but about knowledge claims and the organization of knowledge claims. The implication is that each knowledge claim is supported by and connected with arguments, theories and world views. If this is recognized by the people performing KO [*knowledge organization*], then the activity is not based on “positivism”. (Hjørland 2008a, 98<sup>31</sup>)

On prend en considération un certain niveau d'incertitude dans cette approche, de manière à mettre en lumière le caractère relatif de la connaissance eu égard aux différents arguments,

---

<sup>31</sup> Voir aussi Radford 1992.

théories et points de vue en cause. Le cœur même de la notion de *knowledge claim* se compose d'une dualité manifeste entre tradition et nouveauté : alors que la connaissance (*knowledge*) délimite généralement ce qui est accepté et raisonnable du point de vue des traditions scientifique et académique (pour reprendre le vocabulaire de Bliss), la prétention (*claim*) au savoir est l'élément porteur d'incertitude, de changement, de réaligement ou de transgression<sup>32</sup>.

Dans le contexte de l'analyse documentaire, la continuité prend la forme d'un langage conventionnel (langage documentaire) qui, s'adaptant davantage avec précaution qu'empressement, trace les contours d'une culture savante dans son aspect général (en conformité avec l'approche centrée sur le document). Si l'analyste documentaire veut élargir le spectre de la communication indirecte à laquelle il contribue (à travers le catalogue et autres outils favorisant l'accès à la documentation), il doit aussi porter attention à la dynamique particulière qui contribue au développement des connaissances. Continuité et changement se font face dans la recherche scientifique et académique :

There is continuity because each generation is always taught by the preceding generation of scientists and also because the research questions in which a community is interested are predicated on the current concepts they hold, even when the results of such research might indicate that changes are needed to better adapt the concepts in response to other concepts or other facts about nature. (Hjørland 2008b, 340)

Avec la notion de prétention au savoir, l'évaluation critique des connaissances devient une activité qui implique deux paliers, l'un collectif et l'autre individuel. Budd traduit cette idée dans le langage spécifique de l'épistémologie sociale : « The library can indeed support the critical evaluation of knowledge claims, both individual and collective knowledge growth, and the space where people can engage in the social connections that make knowledge as defined

---

<sup>32</sup> Nous pourrions penser à l'héliocentrisme de Copernic comme exemple canonique ou encore à la théorie de la relativité générale d'Einstein. Avant que Galilée ne soit en mesure de prouver l'hypothèse du premier et que les observations d'Eddington ne valident la théorie du second, il serait plus juste de parler de prétentions au savoir que de connaissances à portée consensuelle. La prétention au savoir met surtout l'accent sur l'incertitude qui accompagne la tentative, portée par un individu ou un groupe de personnes, de reconfigurer ou questionner nos connaissances (en formulant des hypothèses, des suggestions, des prédictions, etc.).

above possible » (Budd 2004, 367). La prétention au savoir décrit le processus d'inclusion de propositions subjectives à l'intérieur d'un bassin de connaissances collectif.

Cette manière de voir la connaissance permet d'appliquer l'approche relativiste à diverses activités et tâches auxquelles contribue la bibliothèque universitaire. Par exemple, la recherche documentaire traduit, pour l'utilisateur, une forme de positionnement épistémique dans un ordre plus large de savoirs :

The activity of conducting literature searches becomes the individual's attempt to locate their knowledge claims within an existing order of knowledge claims. [...] The production of commentaries, critiques, arguments, and debates becomes the act of validating or questioning those locations within the limits of the prevailing arrangement. (Radford 2005, 73)

Dans quelle mesure la bibliothèque peut-elle parrainer ces activités ? Budd croit qu'elle doit au moins considérer certains aspects qui font d'une communauté qu'elle est bien informée : « The library does supply information to its users; the conditions necessary to foster a genuinely informed community (including, perhaps, veritism and reliabilism) must be considered » (Budd 2004, 366). Soergel décrit comment cela pourrait se manifester dans le cas de la bibliothèque universitaire, dont les usagers mènent souvent des recherches complexes :

*Provide the conceptual basis for the design of good research and implementation: assist researchers and practitioners in exploring the conceptual context of a research project, policy, plan, or implementation project and in structuring the problem; support consistent definition of variables/measures for more comparable and cumulative research and evaluation results. (Soergel 1999, 1119 ; l'auteur souligne)*

Nous verrons, dans la sous-section 2.1.3, que la structuration d'une problématique (« structuring the problem ») peut être perceptible à travers la formulation d'un énoncé de recherche. Suivre les objectifs mentionnés par Soergel amène les bibliothécaires, y compris les analystes documentaires, à s'immiscer plus en profondeur dans les activités d'apprentissage et de recherche des usagers.

L'acquisition de connaissances, quelle que soit l'occupation d'un individu, passe d'abord par la recherche d'informations sur un sujet donné. L'utilisateur doit être en mesure de produire des structures conceptuelles à partir des informations recueillies, pour ensuite

préciser sa propre compréhension du sujet, et peut-être reformuler ou bonifier certains concepts. En s'inspirant une fois de plus de Soergel, on peut rapprocher cette activité individuelle des objectifs que cherche à atteindre la bibliothèque universitaire :

*Improve communication and learning: assist writers and readers; support learning through providing conceptual frameworks and challenging students to produce such frameworks; support language learning; and support the development of instructional materials. (Soergel 1999, 1119 ; l'auteur souligne)*

Ces objectifs contribuent à une forme de communication entre usagers et bibliothécaires (communication parfois directe, parfois indirecte) qui pourrait s'avérer être un champ d'action privilégié en milieu universitaire. Si elle répond aux besoins des usagers (pragmatisme oblige), l'analyse documentaire peut contribuer à une communication axée sur l'acquisition, l'évaluation, le partage et le développement des connaissances. Plus spécifiquement, Shera montre comment la documentation et les instruments qui s'y rattachent contribuent à ce qu'il nomme la communication graphique (*graphic communication*) :

*From the foregoing it seems clear that documentation is limited to that aspect of bibliographic organization which treats of the materials and needs of scholars and, hence, is concerned with the scholarly apparatus of bibliographies, indexes, and abstracting services. Therefore, documentation is to be regarded as an essential part of our modern system of graphic communication within the world of scholarship, an instrumental device to expedite the flow of recorded information *within* a group of specialists or *between* various groups of specialists. (Shera 1966, 2 ; l'auteur souligne)*

L'analyste documentaire agit comme médiateur et peut influencer de cette manière sur la circulation de l'information consignée (« the flow of recorded information »). Il favorise indirectement la formulation et l'évaluation de prétentions au savoir (*knowledge claims*) par l'entremise des métadonnées qu'il propose aux usagers. L'attention portée à ce qui motive une prétention au savoir (en tant que positionnement individuel au sein d'une structure de connaissances plus vaste) peut donner une extension plus large à l'analyse documentaire.

L'acquisition de connaissances est en partie déterminée par la relation qu'entretient chaque individu avec la documentation, aussi bien en ce qui concerne l'information qui peut en être extraite que les moyens d'organiser celle-ci. La notion de prétention au savoir se base sur une forme de relativisme qui admet une multitude de points de vue sur un même énoncé (selon

l'approche théorique, l'argumentation ou l'objectif propre à chacun) et qui garde toujours ouverte la possibilité d'un réalignement épistémique. L'analyse documentaire peut, dans une optique faillibiliste, rendre explicite la nécessité d'évaluer et questionner toute connaissance par l'entremise d'une représentation provisionnelle du contenu de la documentation. La prétention au savoir suggère en outre d'établir un rapprochement entre la manière de situer les métadonnées sujet dans une structure plus large (le langage documentaire) et celle, propre à l'utilisateur, de se positionner dans un ordre préexistant de connaissances. Nous verrons plus bas que la *folksonomie* s'avère un outil intéressant pour opérer ce rapprochement (section 2.3). Avant tout, nous allons examiner dans quelle mesure les énoncés de recherche peuvent fournir le matériau nécessaire à une démarche relativiste en milieu universitaire.

### **2.1.3) Les énoncés de recherche**

Avec la notion de prétention au savoir, nous avons abordé la connaissance dans une optique relativiste et nous avons mis la table pour examiner la dimension subjective du sujet. Cette manière de procéder nous permet, avant de prendre acte à nouveau de la notion de sujet (dans une perspective fort différente de celle qui était la nôtre dans la section 1.1), de préciser un exemple de prétention au savoir (*knowledge claim*) pertinente dans le contexte universitaire. Nous avançons que l'énoncé de recherche, défini comme la forme interrogative des diverses problématiques développées sur un sujet<sup>33</sup> (voir l'annexe B pour un exemple), constitue un cas particulier de prétention au savoir (*knowledge claim*). Un énoncé de recherche pourrait fournir une information privilégiée pour comprendre l'intérêt que peut avoir un document aux yeux de l'utilisateur.

Revenons à la question qui sous-tend cette section, c'est-à-dire la contribution du relativisme à la problématique de l'appariement. Notre objectif dans cette troisième sous-section est de montrer qu'un riche bassin d'informations existe dans le milieu universitaire sous la forme d'énoncés de recherche et que ceux-ci pourraient alimenter une approche de l'analyse documentaire centrée sur l'utilisateur. Une perspective relativiste suggère une piste de

---

<sup>33</sup> Il serait peut-être plus exact de parler de « questions de recherche », mais ce terme possède une signification tout autre dans le domaine des sciences de l'information (questions visant à encadrer la recherche documentaire comme telle).

réflexion très actuelle pour l'analyse documentaire : il serait possible de fournir une information fiable, complémentaire à l'enseignement universitaire, en utilisant une représentation plus libre des contenus documentaires. Afin d'en assurer la pertinence, cette information pourrait être issue d'énoncés de recherche produits dans une communauté universitaire, généralement sous la forme d'une question que l'utilisateur souhaite approfondir<sup>34</sup>. Selon Hjørland, la forme interrogative représente un mode de cognition fondamental dans une perspective relativiste :

To say about something that it is informative means that this thing may answer a question for somebody. The informativeness is thus *a relation* between the question and the thing. *No thing is inherently informative. To consider something information is thus always to consider it as informative in relation to some possible questions.* (Hjørland 2007, 1451 ; l'auteur souligne)

Ce point de vue tranche avec l'approche centrée sur le document, qui adopte un registre indexical (« *LE* sujet de *CE* document est *L'*astronomie » ; voir la section 1.1). Le relativisme nous invite à établir une relation entre un document (et l'information qu'il contient) et une interrogation, formulée en l'occurrence par un usager. Dans le milieu universitaire, il pourrait donc être profitable de rapprocher ce point de vue d'une interrogation spécifique et significative<sup>35</sup>, une interrogation qui s'exprime sous la forme d'un énoncé de recherche.

Malgré tout, il peut être hasardeux de se servir d'énoncés de recherche pour orienter l'analyse documentaire. D'une part, ceux-ci peuvent être de nature hautement subjective. D'autre part, ils ne doivent pas être confondus avec la requête d'information ou la question de recherche comme telles (lors d'un entretien de référence par exemple). Néanmoins, un énoncé de recherche, surtout dans le milieu universitaire, précède souvent la formulation d'une requête dans un système documentaire. Quel intérêt pourrait-il y avoir à inclure cet énoncé dans l'évaluation du processus de recherche documentaire ? Un avantage serait de mettre la main

---

<sup>34</sup> L'énoncé de recherche formulé par un usager suggère un ensemble de termes qui interviennent généralement en amont de la recherche documentaire, ce qui en fait une source d'informations unique. Ce champ d'études pourrait se coordonner avec de nombreux travaux qui se penchent, par exemple, sur le vocabulaire utilisé dans les articles scientifiques et les thèses universitaires (titre, sous-titres, résumés, table des matières, etc.).

<sup>35</sup> La forme interrogative, nous semble-t-il, pourrait aider l'utilisateur à formuler (et à formaliser) certains des concepts principaux qui pourraient intervenir dans sa requête d'informations. Dans sa forme succincte, la forme interrogative peut aider l'utilisateur à aller à l'essentiel.

sur une information supplémentaire pour comprendre l'intérêt qu'un usager manifeste pour un document et, par extension, favoriser la production de métadonnées pertinentes (pour faciliter le processus de recherche documentaire). Pour prendre un exemple simple, la question suivante : « Quel système astronomique prévalait avant que ne soit entériné le système héliocentrique de Copernic ? » pourrait aider à comprendre l'intérêt que peut représenter aux yeux d'un étudiant la consultation d'un manuel d'astronomie ou d'une encyclopédie. Il serait possible de postuler un niveau de connaissances de base pour cet étudiant (en tant que lecteur) et un champ disciplinaire bien délimité. En revanche, un étudiant qui travaillerait à partir d'un énoncé de recherche tel que « Comment se pratiquait l'astronomie dans le monde médiéval arabe ? » pourrait ressentir le besoin de consulter une documentation plus variée, ne se limitant pas au domaine de l'astronomie à proprement parler. Cela pourrait être profitable pour un tel étudiant de se voir suggérer des passerelles vers les domaines de l'histoire et de la philosophie à partir d'un mot-clé tel que « médiéval » (pour ce qui est de la dimension historique), mais aussi à partir de la présence simultanée dans un même énoncé des termes « astronomie », « médiéval » et « arabe » (dimension philosophique).

Aussi, Foskett remarque-t-il à juste titre que le lien entre les documents repêchés et la requête d'information formulée par un usager semble incontournable pour déterminer une mesure de pertinence :

If a document is retrieved in answer to a particular request, its relevance may be assessed by a panel of those skilled in the art [*qu'on pourrait traduire par initiés ou experts*], but its pertinence can only be assessed by the originator of the request. In other words, relevance is a consensus judgement, pertinence an individual judgement. (Foskett 1996, 16)

L'énoncé de recherche initial (et la terminologie qu'il mobilise) peut être examiné pour mieux comprendre le jugement individuel (subjectif) qui correspond réellement à la « pertinence » des résultats obtenus lors de la recherche. Par exemple, si les usagers utilisent plus fréquemment le terme « habileté cognitive » que « habileté intellectuelle »<sup>36</sup>, nous pourrions avoir une situation où le consensus potentiel (mesure de *relevance*, que nous pourrions

---

<sup>36</sup> Tiré du thésaurus « Éduthès », cet exemple s'applique très bien, nous semble-t-il, au milieu universitaire. <<http://www.cdc.qc.ca/multites/mtw.exe?k=default&l=91&linkType=term&w=776&n=1&s=5&t=2>>.

rapprocher du pragmatisme documentaire ; voir la section 1.2) serait mieux servi par le premier terme (habileté cognitive), alors que le jugement de pertinence (*pertinence*) d'un usager s'orienterait vers le second terme (habileté intellectuelle). Les énoncés de recherche pourraient révéler la préférence d'un usager ou d'un groupe d'usagers pour un terme plutôt qu'un autre et permettre d'évaluer périodiquement la pertinence des métadonnées sujet associées à la documentation dans le catalogue (la terminologie issue d'un langage documentaire, d'un répertoire de vedettes-matières par exemple, est-elle toujours en vogue?).

Le jugement des usagers, sous la forme d'une mesure de pertinence (*pertinence*) qui ne présuppose aucun consensus (contrairement à la mesure de *relevance*), constitue un indice important en vue de comprendre comment s'oriente la recherche documentaire en milieu universitaire. Davantage que ses semblables, la bibliothèque universitaire doit être attentive au point de vue de ses usagers sur divers domaines de la connaissance pour remplir son mandat (acquisition, évaluation, partage et développement des connaissances). En vue d'approfondir cet aspect de l'analyse documentaire, la position relativiste pourrait s'actualiser dans ce que nous nommerions l'*épistémographie*. Ce néologisme, inspiré de termes tels que « lexicographie », lie l'étude de la connaissance (*épistémè-*) à la nécessité de répertorier les termes (*-graphie*) utilisés par les usagers<sup>37</sup>. L'analyste documentaire, dans son rôle d'*épistémographe*, pourrait gérer une banque de termes issus des énoncés de recherche pour lesquels les usagers manifestent un intérêt. Ceci pourrait se faire en parallèle des langages documentaires « classiques » (répertoires de vedettes-matières, thésaurus et schémas de classification) et permettre d'évaluer la pertinence des termes et des structures qui définissent ces langages issus d'une approche centrée sur le document. Cette entreprise nous amène à considérer le sujet, non pas d'un point de vue objectif comme dans le premier chapitre, mais bien d'un point de vue subjectif (les usagers étant responsables de la production de métadonnées sujet).

Si le lexicographe s'intéresse à l'évolution du lexique selon l'usage qui en est fait aussi bien par le grand public que les auteurs et intellectuels, l'*épistémographe* se pencherait sur la

---

<sup>37</sup> La source directe utilisée serait l'énoncé de recherche plutôt que la documentation elle-même (cette seconde option étant privilégiée lorsqu'on se conforme au principe de caution bibliographique).

terminologie qu'utilisent étudiants, professeurs et chercheurs pour mener à bien leurs travaux. Bien entendu, une évaluation de ces métadonnées potentielles, sur la base de critères de pertinence bien implantés (effet sur les performances de rappel et de précision d'un système documentaire par exemple), serait au cœur du mandat des bibliothécaires en milieu universitaire. Le double rôle « épistémographe-analyste » pourrait modeler la communication indirecte entre analyste documentaire et usagers, et fournir des occasions aux bibliothécaires de s'impliquer concrètement dans le processus d'évaluation de nos connaissances. Le relativisme nous invite à promouvoir un tel canal d'informations et à considérer la connaissance comme un cheminement continu, dont le rouage important pourrait être l'énoncé de recherche en tant que cas particulier de prétention au savoir (*knowledge claim*). Par conséquent, il est souhaitable d'analyser une diversité de points de vue sur un même domaine de la connaissance et de considérer plus attentivement la dimension subjective du sujet. Ceci en vue de renforcer une approche véritablement centrée sur l'utilisateur.

Dans la section 2.1, nous avons examiné tour à tour le relativisme, la prétention au savoir et l'énoncé de recherche. Ceci nous a permis de préciser une tâche (*l'épistémographie*) permettant de répertorier un ensemble de termes, eux-mêmes susceptibles de traduire les divers points de vue qu'il est possible d'avoir sur un même sujet. L'objectif serait de mieux comprendre la *relation* que les usagers entretiennent avec la documentation (de quelle manière veulent-ils insérer une prétention au savoir à l'intérieur d'un bassin de connaissances documentées ?) et rendre explicite leur intérêt pour celle-ci. Ceci offre un éclairage spécifique concernant la problématique de l'appariement et une manière de présenter en parallèle les métadonnées sujet issues d'une approche centrée sur le document et le vocabulaire privilégié par les usagers.

## **2.2) La dimension subjective du sujet**

Nous allons à présent examiner la notion de sujet à partir d'un point de vue relativiste, opposé à la tendance objectiviste adoptée dans le premier chapitre du mémoire. L'objectif final de cette section étant d'éclairer les bases d'une caution usagère, il sera question d'une approche qui prend en compte le caractère subjectif du sujet et, par conséquent, la relation variable qu'entretient l'utilisateur avec la documentation (à notre avis, ceci implique d'abord

d'examiner la notion de lecture). Pour plusieurs commentateurs (Budd, Hjørland, Mai et Frohmann par exemple), l'intérêt d'une telle démarche repose tout simplement sur le fait qu'elle représente l'alternative actuelle la plus intéressante à l'approche classique de l'analyse documentaire, centrée sur le document. Selon nous, l'approche subjectiviste doit être vue comme un élément complémentaire plutôt qu'une autre version de l'analyse documentaire. Elle fournit une information supplémentaire *non contradictoire* (même dans le cas où il s'agit d'une information *contraire*) par rapport aux métadonnées sujet issues de l'approche centrée sur le document. Finalement, on pourrait ajouter, avec Alemu *et al.*, que le défi principal du bibliothécaire est de déterminer quelle place doit occuper le point de vue de l'utilisateur dans les décisions concernant l'organisation de la documentation :

It has also become apparent that the primary challenge facing librarians today is, to first of all place users at the center of all information organization decisions, and to that effect, ensure that metadata systems reflect the world view of their users in terms of representing various perspectives and interpretations. (Alemu *et al.* 2012, 338)

Ceci est peut-être plus facile à formuler qu'à mettre en pratique, mais il semblerait qu'un tel défi est inévitable dans le présent contexte informationnel. Une abondance de ressources et d'outils disponibles placent la perspective de l'utilisateur au cœur de nouvelles méthodes d'organisation de l'information.

Comme nous l'avons déjà souligné dans l'introduction à propos de l'organisation du mémoire, l'approche centrée sur l'utilisateur sert d'antithèse (fonction argumentative) à la démarche pragmatiste classique que nous avons placée au cœur du développement de l'analyse documentaire (premier chapitre). La perspective subjectiviste confère une dimension supplémentaire à la notion de sujet et place celle-ci dans le cadre d'une activité personnelle et contingente. Cette section nous permettra d'aborder la lecture en tant qu'élément supplémentaire pour justifier le recours au relativisme (sous-section 2.2.1). Ensuite, il sera possible de décrire un environnement de recherche documentaire (limité aux ressources langagières) qui favoriserait une interaction accrue entre utilisateurs et analystes documentaires (sous-section 2.2.2). Finalement, nous verrons qu'une part de subjectivité doit aussi être admise chez l'analyste documentaire, une caution usagère permettant toutefois d'encadrer l'analyse du sujet dans sa dimension subjective (sous-section 2.2.3).

### 2.2.1) *La lecture en milieu universitaire*

Afin de réintroduire la notion de sujet et de délimiter la perspective qui sera la nôtre dans l'ensemble de cette section, reprenons ce que Maniez dit à propos de ce concept pivot : « Si notre conception du sujet est exacte, il s'agit d'un objet mental à la croisée de l'objectif et du subjectif. Il est donc insuffisant de caractériser le sujet d'un document par sa base référentielle » (Maniez 2005, 157). Le référent du sujet, comme nous avons tenté de le démontrer dans le premier chapitre du mémoire, peut être limité à un seul élément (univocité) par l'entremise d'un processus analytique réductionniste (section 1.3). Au-delà de cette constatation, il faut aussi prendre en considération l'aspect subjectif du sujet, lequel est informé par une diversité de points de vue :

Nous avons défini un sujet comme une parcelle de l'univers connaissable à laquelle s'attache l'intérêt d'un individu. À la différence des autres clés [*de recherche*], cette notion n'est pas entièrement objective, l'intérêt de deux individus pour le même objet de connaissance pouvant être motivé par des points de vue divers. (Maniez 2005, 171)

Une approche à tendance subjectiviste n'engendre pas nécessairement un traitement particulariste du sujet (*i.e.* représenter les traits distinctifs d'un sujet), mais elle entretient un rapport certain avec le relativisme. Rappelons la définition concise du relativisme donnée par Lalande : « Doctrine qui admet que toute connaissance (ou toute connaissance humaine) est relative » (Lalande 2010, 914). Le sujet, de ce point de vue, doit être analysé à la lumière d'une diversité de lectures possibles, à raison des multiples significations qui peuvent être attribuées à un énoncé (celui-ci décrivant une « parcelle de l'univers connaissable »). Une approche subjectiviste s'intéresse avant tout à la *relation* qu'entretient l'individu avec un domaine de la connaissance et avec la documentation qui en est le véhicule.

La dimension relativiste de la connaissance admet, *a priori*, une variété d'interprétations beaucoup plus grande qu'une approche qui vise à rendre objective et neutre l'analyse documentaire. Ceci résulte de la volonté d'inclure un plus grand nombre de points de vue sur un sujet donné, de même que de considérer différentes lectures pour un même document. Une démarche subjectiviste qui s'attarde à la lecture se réclame en partie de la

théorie de la réception, dont l'une des caractéristiques principales, comme le fait remarquer Mai, est l'aspect relatif de la signification textuelle :

In this understanding of texts and their interpretation, a text does not have meaning in and by itself. The meaning of the text is created by the reader as the text is read and used. A reader does not *respond to* the meaning of a text. The reader's response *is* the meaning of the text. (Mai 2005, 604 ; l'auteur souligne)

Nous voyons que l'approche centrée sur l'utilisateur peut s'inspirer de cette théorie de manière à donner une dimension supplémentaire au document, dans la mesure où celui-ci devient le reflet d'une lecture particulière. L'expérience de lecture définit les significations subjectives que donnent les usagers au contenu de la documentation, une approche centrée sur l'utilisateur permettant de rapprocher cette expérience subjective du processus de recherche documentaire. Nous verrons que la lecture en milieu universitaire, qui implique souvent une forme d'exégèse<sup>38</sup> chez l'étudiant ou le professeur, a le potentiel de souligner fortement la problématique de l'appariement, puisqu'elle s'oppose au régime indicatif qui définit traditionnellement une analyse documentaire centrée sur le document (voir la section 1.1).

Banou, Kostagiolas et Olenoglou soulignent l'importance de la lecture en milieu universitaire, activité dont la promotion n'est pas nécessairement encadrée par une politique académique formelle :

The differences in reading behaviour between the students and the general population illustrate that the academic environment, including the library, promotes and improves reading behaviour. This is not, however, necessarily a result of a "formal" reading policy decided by the administration of the University and/or the Academic Library. (Banou, Kostagiolas, Olenoglou 2008, 499)

Dans un contexte où les compétences informationnelles sont de plus en plus valorisées, la question corrélative de la lecture doit être étudiée par le bibliothécaire en milieu universitaire. Si ce n'est fait à travers une politique décidée conjointement par l'université et la bibliothèque,

---

<sup>38</sup> Pour utiliser une terminologie plus actuelle, nous pourrions aussi parler d'une « lecture critique » pour laquelle l'utilisateur s'adonne à une interprétation du texte ou de l'énoncé qu'il a sous les yeux. Bien qu'il s'agisse d'un exercice qui obéit à certaines règles, à certaines traditions méthodologiques et théoriques, il s'éloigne grandement de la lecture documentaire, fortement balisée. La lecture critique initie souvent une forme de communication complexe, laquelle ne représente pas traditionnellement le champ d'action de l'analyste documentaire.

les habitudes de lecture des usagers pourraient néanmoins être prises en compte par l'analyste documentaire. Idéalement, cet exercice exigerait une participation aussi bien des membres du corps professoral que des étudiants. Dans un contexte favorable, les habitudes de lecture des usagers pourraient faire l'objet d'études académiques visant à dégager une information importante pour la bibliothèque universitaire :

Research on the reading behaviour of students is indeed very important and may express the needs and expectations of the users-students, "giving voice to them". The academic libraries and the academic institutions need that information in order to develop their policies and strategies. (Banou, Kostagiolas, Olenoglou 2008, 500)

Du point de vue de l'analyse documentaire, ce type de recherche suggère indirectement une orientation plus libérale dans les procédures d'indexation et de classification<sup>39</sup>, pour inclure une terminologie plus vaste et plus malléable (tels que les *tags*, sur lesquels nous reviendrons à la section 2.3).

Parce qu'elle n'est pas balisée (contrairement à la lecture documentaire), une lecture subjective du document s'éloigne du processus qui vise à dévoiler l'*aboutness* du document telle que nous l'avons définie dans le premier chapitre du mémoire. Ceci met l'accent sur une forme d'analyse contingente qui, sans être exempte de règles, n'obéit pas aux mêmes codes que l'analyse documentaire classique (centrée sur le document). Sans le bénéfice d'un *feedback* constant (Šauperl 2002, 142), les habitudes de lecture des usagers constituent une variable inconnue, un facteur d'incertitude pour l'analyste documentaire. C'est pour cette raison que nous avons introduit, dans la section précédente, l'énoncé de recherche en tant que point de départ potentiel pour l'analyse documentaire. La terminologie issue d'un ensemble suffisamment large d'énoncés de recherche pourrait révéler des tendances fortes, des liens interdisciplinaires, des structures conceptuelles (similaires à celles que l'on retrouve dans un thésaurus), voire même des incohérences ou des conflits.

---

<sup>39</sup> La classification est généralement soumise à des règles strictes pour favoriser la régularité dans l'organisation des collections. Néanmoins, une orientation plus libérale pourrait s'exprimer, dans une collection de documents numériques, par la « polyclassification », alors que celle-ci est généralement proscrite pour les collections physiques (Hjørland 2008b, 340).

De manière générale, une vision relativiste de l'analyse documentaire, par l'importance qu'elle accorde à l'expérience du lecteur, pose l'usage pour arbitre discrétionnaire au détriment de l'application de règles générales prédéfinies. Ceci peut paraître paradoxal, mais une telle position transparait jusque dans les célèbres règles proposées par Cutter pour la description des sujets :

General rules, always applicable, for the choice of names of subjects can no more be given than rules without exception in grammar. Usage in both cases is the supreme arbiter,— the usage, in the present case, not of the cataloguer but of the public in speaking of subjects. (Cutter 1904, 69)

L'« usage » avoisine, chez Cutter, des mots tels que coutume ou habitude, mais il peut aussi suggérer une multitude de perspectives et d'orientations théoriques spécifiques. En ce sens, déchiffrer le point de vue d'usagers dont les intérêts divergent (en raison, par exemple, des exigences particulières d'un domaine d'études) s'avère un défi intéressant pour la discipline de l'analyse documentaire. C'est pourquoi devant un tel mandat, nous avons choisi de limiter notre étude au milieu universitaire, tout en examinant les développements récents dont pourrait bénéficier la discipline de l'analyse documentaire dans son ensemble.

Notre objectif n'est pas de détailler une approche novatrice de l'analyse documentaire en milieu universitaire, ni de transposer un outil analytique spécifique dans l'univers bibliothéconomique. Néanmoins, nous voulons souligner, par le biais d'auteurs qui se sont intéressés à cette question, la dimension supplémentaire que peut apporter à l'analyse documentaire une orientation théorique subjectiviste. Un document ne pouvant, selon le point de vue relativiste (ou constructionniste<sup>40</sup>), transmettre une information de manière parfaitement objective, une attention particulière doit être portée aux aptitudes cognitives de l'utilisateur. En tant que lecteur, celui-ci doit être considéré comme une sorte d'agent causal, apte à « construire » un sens spécifique (subjectif) à partir d'une information donnée :

---

<sup>40</sup> Une distinction existe, en philosophie des sciences, entre la position constructionniste qui met l'accent sur l'aspect subjectif de la cognition et la position constructiviste (héritée de Kant) qui inclut à la fois des éléments objectifs et phénoménologiques (observation des phénomènes par un sujet). Pour notre propos, le terme « relativisme », bien qu'il laisse de côté cette distinction, donne une idée juste de l'aspect subjectif du sujet lorsqu'il est porté à l'attention d'un individu à travers le médium textuel.

Advocates of the user-oriented approach like Dervin (1983) cast doubt on the mechanistic and mundane understanding of information as the direct communication of messages between senders and receivers. The emerging cognitive and constructivist theories underlined that the information user is not a passive information processing system but actively makes sense of the surrounding reality and attaches personal meanings to information. (Savolainen 2009, 195)

La difficulté est d'anticiper les points de vue divers des usagers ainsi que l'usage varié qui peut être fait de la documentation, en vue de refléter ces dispositions dans l'analyse documentaire. Pour l'instant, nous pouvons souligner l'apport potentiel de cette dimension subjective dans l'utilisation des systèmes documentaires et le défi qui consiste à accommoder cette diversité de points de vue au moyen d'une terminologie pertinente.

Rappelons que Maniez associe fortement l'aspect subjectif à l'intérêt qu'attache un individu à un sujet donné (Maniez 2005, 171). Nous pouvons ajouter que cet intérêt évolue au fil du travail intellectuel et de la lecture qui composent en bonne partie les activités des usagers. L'aspect subjectif du sujet, dans un contexte universitaire à plus forte raison, entretiendrait donc un lien direct avec l'acquisition de connaissances et l'évaluation de la pertinence de l'information pour accomplir une tâche académique. Toutefois, le rôle du bibliothécaire étant de faciliter l'accès aux documents plutôt que d'évaluer les connaissances des usagers (voir la sous-section 2.1.1), la lecture peut être considérée comme le prolongement de l'intérêt pour un énoncé de recherche, tel que celui-ci s'inscrit dans un bassin de connaissances plus vaste (voir la sous-section 2.1.3). Il est donc souhaitable de s'attarder à l'environnement de recherche documentaire qui encadre ce processus et à l'impact que peut avoir l'analyse documentaire sur ce dernier. Cet angle d'approche nous permettra de passer de l'utilisateur en tant que lecteur à l'utilisateur en tant que « fureteur ».

### ***2.2.2) L'environnement de recherche documentaire***

Dans le présent contexte, l'environnement de recherche documentaire est défini comme l'ensemble des ressources langagières<sup>41</sup> offertes dans un système documentaire pour permettre

---

<sup>41</sup> Dans le contexte du présent mémoire, il s'agit à la fois d'indications provenant de langages documentaires classiques (voir le premier chapitre du mémoire) et de communications initiées par les usagers (principalement le vocabulaire issu des énoncés de recherche).

aux usagers de trouver diverses sources d'information (documents). Cette définition, d'une portée intentionnellement restreinte, nous permet de nous occuper avant tout de la problématique de l'appariement<sup>42</sup> et d'évaluer le résultat de l'analyse documentaire par rapport aux ressources et aux fonctions qu'offrent certains outils de recherche largement utilisés. En effet, plusieurs études récentes indiquent que les usagers en bibliothèque universitaire préfèrent un environnement de recherche qui se rapproche des bases de données en ligne et des moteurs de recherche (plutôt qu'un environnement qui s'inspire des catalogues traditionnels) :

The repercussion of the advances in indexing is also well known in the information behavior of users, as reported by the research users and undergraduate students participating in this research, who prefer subject retrieval resembling a database or a Web metasearch engine. (Fujita *et al.* 2010, 32)

Ces usagers peuvent ignorer les acquis d'une longue tradition de catalogage dont les avantages sont largement reconnus (Fujita *et al.* 2010, 32-33), mais ils possèdent généralement une expérience non négligeable dans la recherche d'informations en ligne (surtout, semble-t-il, chez les *digital natives*). Si l'analyste documentaire veut approfondir sa compréhension du processus de recherche d'information, s'il veut par le fait même pouvoir répondre aux attentes des usagers, toute occasion d'initier une interaction doit être considérée attentivement afin d'enrichir l'environnement de recherche documentaire proposé aux usagers.

Cette interaction peut être obtenue de diverses manières, directes et indirectes : entretien de référence, entrevue, sondage, analyse des habitudes de recherche documentaire, etc. Dans le contexte qui est le nôtre, cela veut surtout dire fouiller les significations que les usagers peuvent attacher à un texte donné et mieux comprendre comment celles-ci peuvent influencer l'utilisation du système documentaire. Ces objectifs se combinent à celui d'assister le mieux possible les usagers dans leurs recherches documentaires à l'aide de systèmes plus performants :

One of the current challenges of Information Retrieval Systems (IRS) is to develop tools able to integrate more semantics [*meanings*] in their treatment.

---

<sup>42</sup> Une fois de plus, il s'agit de favoriser l'appariement entre une ressource langagière proposée dans un système documentaire (métadonnée sujet) et le vocabulaire utilisée par l'utilisateur dans ses propres recherches.

The aim is twofold: 'understanding the content of documents' and 'understanding the user needs' and to be able to link them. (Bouramoul 2011, 425)

Prendre acte d'une diversité de significations (« integrate more semantics ») dans le traitement analytique, plus spécifiquement, atteindrait donc des objectifs en tous points semblables à ceux identifiés par Bouramoul : bonifier la compréhension du contenu d'un document et lier celle-ci aux besoins des usagers. C'est dans une telle optique que la dimension subjective du sujet s'impose comme un pôle important de la problématique de l'appariement. La solution se traduit-elle nécessairement par une liberté et une variété plus grande dans la terminologie utilisée pour représenter un éventail de significations plus vaste ? Sans abroger sa contribution spécifique (contrôle et structuration des métadonnées sujet), nous pensons que l'analyste documentaire peut détailler l'environnement de recherche spécifique à un usager ou un groupe d'usagers en utilisant, comme point d'ancrage, l'information contenue dans les énoncés de recherche.

Nous admettons que l'environnement de recherche documentaire est une notion qui peut se décliner de plusieurs manières, mais nous voulons avant tout souligner la tendance actuelle qui vise à actualiser les langages documentaires et à offrir aux usagers un environnement langagier détaillé (avec la contribution d'analystes documentaires aussi bien que d'usagers). Dans cet ordre d'idées, il paraît nécessaire de regarder la possibilité d'offrir une plus grande variété de termes pour, entre autres, multiplier les possibilités d'interaction bibliothécaires/usagers (communication indirecte ; voir la sous-section 2.1.1) et rendre compte de divers points de vue sur la documentation. À défaut de côtoyer quotidiennement les usagers, l'analyste documentaire peut jouer un rôle actif dans l'expansion du vocabulaire utilisé pour la recherche documentaire. À titre d'exemple, le développement des collections numériques et l'accès à distance à la documentation représentent des tendances qui pourraient profiter de l'utilisation de métadonnées enrichies, métadonnées liées à la terminologie utilisée par les usagers :

This behavioral pattern [*accès aux documents à distance*] highlights the importance of making electronic library resources discoverable, perhaps with the provision of richer metadata. The idea of providing richer metadata is quite important because, if the metadata is not built accurately (e.g., in a manner that employs users' terminologies) then information objects, could

remain undetected or obscure, and hence, inaccessible, even though they could be present in library databases or on their shelves. (Alemu *et al.* 2012, 338)

Si l'on veut réellement inclure le lecteur dans l'équation, il faut placer celui-ci en amont du processus d'analyse. L'approche centrée sur l'utilisateur nous invite à réfléchir aux possibilités d'amélioration des langages documentaires (par exemple, en favorisant un projet *folksonomique* en parallèle ; voir la section 2.3), de manière à les rendre plus dynamiques sans nécessairement changer leur nature ou leur rôle.

Essayons par conséquent de déterminer l'impact positif que l'orientation subjectiviste pourrait exercer sur le développement des langages documentaires. Comme nous y avons déjà fait allusion, l'approche centrée sur l'utilisateur ne doit pas être vue comme un désaveu de la tradition bibliothéconomique, mais comme une orientation qui nous invite à poser un regard critique sur les moyens et outils que nous utilisons pour organiser la documentation. Dans un esprit similaire, Hudon et Mustafa El Hadi suggèrent certaines modifications qui pourraient être apportées aux schémas de classification pour en assurer la pertinence dans les années à venir :

Puisque les forces et les faiblesses des schémas de classification documentaires sont connues, il est possible de décrire assez précisément les modifications à leur contenu et à leur structure qui pourraient leur assurer un rôle de premier plan sur les réseaux d'information. Il faut développer des mécanismes qui permettront de suivre de plus près l'évolution des savoirs et d'intégrer rapidement non seulement les nouveaux sujets, mais également les nouvelles façons de présenter les sujets existants. Il faut moderniser et populariser les termes utilisés pour représenter les sujets. Il faut prévoir plus de flexibilité et des structures alternatives qui permettent la représentation des besoins particuliers de certains usagers, en fonction de différences culturelles, notamment. (Hudon et Mustafa El Hadi 2010, 28)

Cette citation met l'accent sur les améliorations à apporter aux schémas de classification existants, plutôt que sur l'élaboration de nouveaux langages documentaires. De même, la dimension subjective du sujet s'inscrit dans un renouveau ou une réévaluation des outils déjà en place, avec pour objectif de mieux adapter ceux-ci aux mécanismes de développement des savoirs (en milieu universitaire par exemple).

Dans la section 2.1, nous avons fait allusion à la position critique qu'adoptent fréquemment les auteurs qui se réclament d'une certaine forme de relativisme au détriment d'une démarche objectiviste. L'analyse documentaire, pense-t-on généralement dans un tel cas, peut bénéficier d'informations provenant des usagers eux-mêmes, révélant la dimension subjective du sujet (en marge du document en tant qu'objet soumis au registre indexical ; voir la section 1.1). Dans le milieu universitaire, donner une place importante à la dimension subjective du sujet peut certainement permettre « de suivre de plus près l'évolution des savoirs » et faciliter la « représentation des besoins particuliers de certains usagers » (Hudon et Mustafa El Hadi 2010, 28). Mettre en place des outils dynamiques augmente les occasions d'interactions entre usagers et bibliothécaires et enrichit le cadre de travail de l'analyste documentaire. Cette démarche munit celui-ci d'un tableau plus détaillé du vocabulaire privilégié par les usagers, un fonds terminologique évolutif et « excentré » (du point de vue du contrôle d'autorité). En retour, l'analyste documentaire peut rendre compte d'habitudes de lecture, de points de vue spécifiques sur la documentation (selon un domaine ou un niveau d'études par exemple) de manière à enrichir l'environnement langagier dans lequel les recherches documentaires sont menées par les usagers. Il est bien connu qu'une telle démarche peut affecter (négativement) les mesures de rappel et de précision dans un système documentaire. Une forme de caution particulière doit par conséquent être envisagée pour circonscrire la dimension subjective du sujet.

### ***2.2.3) La caution usagère***

En définitive, la dimension subjective du sujet nous oblige à admettre une part de contingence dans l'analyse textuelle (*i.e.* attribuer des significations variées et variables à un énoncé ou à un texte dans son ensemble). Cela n'est d'ailleurs pas étranger à tout type d'analyse documentaire, qu'elle soit centrée sur le document ou sur l'utilisateur. À l'instar des activités de l'utilisateur, le travail de l'analyste documentaire contient une part de subjectivité, peu importe le nombre et la qualité des règles d'indexation et de classification utilisées, peu importe la précision et l'exhaustivité de la politique documentaire en vigueur. Mais rend bien compte de cette réalité en mentionnant les intérêts implicites ou explicites de l'analyste :

An important and often overlooked aspect in indexing is the indexers' implicit or explicit interests, perhaps since it is expected that the indexers will assume a neutral role or no role at all. However, since analysis in indexing is a context-dependent interpretation and decision making process, the indexer unavoidably assumes a certain position. An analysis of the indexer's role is needed to make it explicit. (Mai 2005, 608)

Dans une perspective subjectiviste, le rôle de l'analyste documentaire doit être façonné en considérant le contexte dans lequel il exerce ses fonctions. Dans le cas qui nous intéresse, il s'agit d'un milieu défini par les activités d'apprentissage et de recherche des usagers.

L'orientation à donner à l'analyse documentaire peut venir de l'intérêt qu'un document représente aux yeux de l'utilisateur dans l'accomplissement d'une tâche académique. L'un des avantages d'une analyse documentaire centrée sur l'utilisateur, de ce point de vue, repose sur l'exploitation d'une caution usagère. Ce terme n'est peut-être pas couramment utilisé, mais nous le retrouvons chez Lancaster dès 1986, en contraste avec la notion plus connue de caution bibliographique :

Deriving terms from the type of literature that is to be indexed clearly satisfies the requirement for literary warrant (i.e., an index term is justified only if it is known to occur in the literature). There is another requirement that is frequently overlooked, however. This can be referred to as *user warrant*: A term is justified for inclusion in an index only if it is of interest to the users of the information service. (Lancaster 1986, 26 ; l'auteur souligne)

Suivant cette définition de la caution usagère, l'inclusion d'un terme dans un index serait justifiée par l'intérêt démontré par les usagers d'une institution documentaire pour ce même terme. Ce type de caution représente le prolongement du relativisme souligné dans la section 2.1, puisqu'ils mettent tous deux l'accent sur le point de vue soutenu (de façon implicite ou explicite) par l'utilisateur.

Similaire dans son fonctionnement à la caution bibliographique, une caution usagère pourrait fournir des indices probants quant au vocabulaire en vogue (à rapprocher de l'*épistémographie* ; voir la sous-section 2.1.3). En outre, nous pourrions la définir par l'entremise du vocabulaire issu des énoncés de recherche : un terme est inclus dans le langage documentaire (telle qu'une *folksonomie* ; voir la section 2.3) lorsqu'il traduit un intérêt de recherche pour un utilisateur. Le double avantage serait de préciser la portée d'une caution

usagère et d'encadrer la dimension subjective du sujet en mettant l'accent sur un type d'intérêt plus précis, c'est-à-dire l'intérêt de recherche en milieu universitaire. Faisant appel au vocabulaire issu des énoncés de recherche, l'analyste documentaire peut examiner ce que signifie, pour une communauté d'utilisateurs, d'associer tels termes à un contenu et, faisant appel aux langages documentaires classiques, expliciter les liens conceptuels (liens associatifs par exemple) qui existent entre ces termes eux-mêmes. Un utilisateur pourrait, par exemple, formuler un énoncé de recherche qui suggère un lien sémantique entre le réalisme en peinture et le réalisme en littérature. Ces deux sujets pourraient être rapprochés par l'analyste documentaire en mettant en parallèle deux indices de classification qui témoignent de ce que Savage a appelé les « distributed relatives » (sujets sémantiquement liés mais séparés dans un plan de classification ; voir Savage 1946). Cet apport de l'analyste documentaire représente un gain potentiel pour l'utilisateur dans ses recherches documentaires, de même qu'une manière d'augmenter la visibilité des collections documentaires.

Comme nous l'avons suggéré dans la section 2.1, le vocabulaire utilisé dans les énoncés de recherche formulés par les utilisateurs constitue une source d'informations riche de significations. Nous pensons qu'une caution usagère qui s'en inspire contribue à décrire le contexte dans lequel les utilisateurs s'acquittent de leurs tâches académiques, y compris la recherche documentaire, et à inclure plus activement les préoccupations qui en émanent dans le travail de l'analyste documentaire. La dimension subjective du sujet serait restreinte principalement aux intérêts de recherche des utilisateurs en milieu universitaire, de manière à favoriser un canal d'informations suffisamment étroit (pour limiter la variété et l'abondance de l'information). La caution usagère, agissant au sein d'un projet *folksonomique* (dont il sera question dans la section 2.3), peut entraîner au moins deux mesures de pertinence : l'une associée au « poids » d'un terme (son utilisation fréquente pour décrire un concept ou pour subsumer un énoncé ou un texte quelconque), l'autre associée à l'utilisation d'un terme (ex. : un terme inclus dans le syllabus d'un cours ou d'un séminaire pour décrire un concept jugé important ; ce terme pourrait décrire des intérêts de recherche convergents<sup>43</sup>).

---

<sup>43</sup> Par ailleurs, il serait possible d'arguer qu'un syllabus, représentant une sorte de plan de travail en milieu universitaire, est composé d'un ensemble d'énoncés de recherche.

Nous avons parlé amplement de l'énoncé de recherche parce qu'il permet de solliciter la participation des usagers tout en demeurant centré sur leurs intérêts. La section 2.2 nous aura permis d'établir une caution usagère qui met l'accent sur une terminologie susceptible d'enrichir l'environnement de recherche documentaire, tout en limitant les biais potentiels liés à l'analyste documentaire lui-même et ceux qui pourraient influencer la formulation d'un énoncé de recherche. L'usage fréquent d'un terme ou son utilisation spécifique (ex. : inclusion dans un plan de cours) peuvent témoigner d'intérêts de recherche et ainsi fournir une information privilégiée à l'analyste documentaire dans le cadre de son travail. Une interaction analystes/usagers adaptée au milieu universitaire peut être favorisée par la participation des usagers à un type précis de communication indirecte (voir la sous-section 2.1.1), dont la *folksonomie* est le rouage principal. C'est un tel projet que nous allons maintenant examiner.

### **2.3) La *folksonomie***

Nous allons à présent montrer que la *folksonomie* peut s'avérer être un outil intéressant pour l'analyste documentaire, puisqu'elle constitue une occasion supplémentaire d'interagir avec les usagers (communication indirecte). En conformité avec le principe de caution usagère suggéré par Lancaster, la collaboration des usagers peut être vue comme un moyen efficace d'améliorer un système documentaire : « If one can get the users of a system to cooperate fully (as one might, for example, if developing a thesaurus for use by a single organization), another approach is possible: Have the users select terms from the literature. If practical, this is probably the best approach of all » (Lancaster 1986, 28). Dans le même ordre d'idées, la production de métadonnées associée au Web 2.0 s'affirme comme un paradigme intéressant qui intervient dans la foulée du relativisme : « In a socially constructed metadata paradigm, users not only search/browse and access content but also participate in its production and description » (Alemu *et al.* 2012, 313). Dans son ensemble, cette section vise à montrer que ce paradigme relativement récent doit être considéré attentivement dans le contexte de l'analyse documentaire. Nous allons d'abord faire un bref retour sur la problématique de l'appariement (sous-section 2.3.1). Ensuite, nous préciserons dans quelle mesure une *folksonomie intégrée*, librement orientée par la caution usagère, peut contribuer à l'analyse documentaire en milieu universitaire (sous-section 2.3.2).

Avant de poursuivre, il est important de préciser deux aspects de la *folksonomie* telle que nous l'aborderons ici, le premier étant spécifique à la réflexion menée jusqu'ici dans le deuxième chapitre du mémoire et le second étant d'un ordre plus général. D'abord, elle doit être vue comme le résultat de la description d'intérêts de recherche (par le biais de l'énoncé de recherche ; voir la sous-section 2.1.3) et non comme le résultat de la description d'un document préalablement consulté par l'utilisateur. Ensuite, l'avantage de la *folksonomie* reposerait « sur le système de *crowd sourcing*, c'est-à-dire l'alimentation des plateformes par les données issues de la foule » (Le Deuff 2012, 59). Comme nous y avons brièvement fait allusion au début de la sous-section 2.1.1, la fréquence d'utilisation d'un terme est un élément communément souligné pour en mesurer la pertinence et l'intérêt potentiel aux yeux de l'utilisateur (le nuage de mots-clés, ou *tag cloud*, représente bien ce principe ; voir Anfinnsen, Ghinea et de Cesare 2011, 65). Nous aurons l'occasion de suggérer une forme de pertinence spécifique au milieu universitaire (la terminologie utilisée dans un plan de cours), mais il est certain qu'une participation active des usagers est un élément incontournable de la *folksonomie*.

### **2.3.1) La problématique de l'appariement et la dimension subjective du sujet**

La problématique de l'appariement est, à sa base, largement déterminée par des préoccupations terminologiques. Comme nous l'avons déjà souligné, l'absence de *feedback* venant des usagers (concernant leur appréciation d'un langage documentaire par exemple) représente une faiblesse que les catalogueurs eux-mêmes soulignent (Šaupperl 2002, 142). De plus, l'impossibilité de répertorier toutes les motivations susceptibles de donner lieu à une recherche documentaire empêche d'attribuer au vocabulaire contrôlé une cote d'efficacité liée aux intérêts des usagers<sup>44</sup>. La perspective relativiste, décrite dans la section 2.1, remet en question l'évaluation objective des métadonnées sujet, sans toutefois rejeter *a priori* la sujétion de celles-ci au contrôle d'autorité ni l'impact positif qu'une telle méthode peut avoir sur les

---

<sup>44</sup> Si une telle cote d'efficacité pouvait être déterminée (ex. : la correspondance entre un terme retenu et la préférence d'une majorité d'utilisateurs sondés pour ce même terme), alors nous pourrions envisager une forme de caution usagère objective. Mais, nous pensons qu'une telle approche (évaluation terminologique objective à partir de données subjectives, telles que les opinions, suggestions ou commentaires des usagers) court toujours le risque d'être trop réductrice pour être vraiment efficace.

mesures classiques de rappel et de précision. Nous pensons que ces mesures possèdent cependant un caractère trop général pour indiquer si une requête d'informations est réellement représentative des intérêts et exigences de l'utilisateur. Dans un tel contexte, le vocabulaire contrôlé peut souffrir, aux yeux de l'utilisateur toujours, d'un manque de précision difficile à déceler pour l'analyste documentaire :

Furthermore, if requests are always made in terms selected from the controlled vocabulary, there is little opportunity to improve the vocabulary by making it more responsive to user requirements. In other words, one is given no evidence that the vocabulary needs to be more specific. (Lancaster 1986, 152)

Une indication à l'effet que le vocabulaire doit être précisé peut provenir d'une approche subjectiviste, permettant à l'utilisateur d'indiquer la terminologie spécifique qui lui serait utile dans ses recherches documentaires (à quoi peut contribuer, comme nous le verrons, la *folksonomie*). Le plus souvent, l'utilisateur se détourne du vocabulaire contrôlé pour favoriser une recherche par mots-clés (Badke 2012a), ou encore se satisfait d'un résultat qui atteint un niveau de pertinence variable, dont le seuil d'acceptabilité est certes difficile à fixer. Dans quelle mesure l'utilisateur est-il conscient des possibilités supplémentaires que l'analyse documentaire peut lui offrir ?

Il est important de noter que la problématique de l'appariement résulte aussi du développement de la bibliothéconomie et, plus précisément, des exigences particulières de l'analyse documentaire. Nécessaire pour représenter adéquatement le contenu de la documentation dans les systèmes documentaires (surtout dans le cas de vastes collections comme celles des bibliothèques universitaires), l'approche centrée sur le document<sup>45</sup> est issue d'une tradition catalographique méconnue, dont plusieurs des caractéristiques principales sont héritées de Cutter. Les procédés descriptifs et analytiques, avec tous leurs développements internes (règles de catalogage, d'indexation et de classification) et externes (développement des outils informatiques par exemple), impliquent des compétences informationnelles que la plupart des utilisateurs ne sont pas prêts à acquérir :

---

<sup>45</sup> Il est important de se rappeler que ce type d'analyse comporte un ensemble de règles, souvent implicites, pour délimiter le cadre thématique général d'un document (*aboutness*). De son côté, l'approche *folksonomique*, permettant une plus grande expressivité dans la notation, tend à conserver (ou restituer) une part plus importante de la complexité sémantique de la documentation (voir Feinberg 2011).

Cataloging rules are very complex because they are intended and designed to describe and provide access to complex entities. These systems were founded by and for librarians and information specialists. Extensive training and hands-on experience are required in order to gain proficiency and expertise in applying them. The majority of users of a library or website does not have the time or interest in learning how to use them extensively. (Lawson 2009, 580)

Les langages documentaires ne sont pas toujours faciles à exploiter, même pour les bibliothécaires et autres spécialistes des sciences de l'information. Cette difficulté a été périodiquement relevée par plusieurs commentateurs (Petee 1946, 46, Harris 1970, 13-20 et Badke 2012a). Ceci expliquerait au moins en partie l'attrait que peut représenter pour les usagers un vocabulaire libre tel que la *folksonomie* : malgré certains inconvénients évidents (par exemple : limitation ou encore absence complète du processus de contrôle du vocabulaire), elle présente aux usagers une plus grande variété terminologique et une plus grande souplesse dans l'utilisation. C'est dans cette perspective que l'on peut examiner la dimension subjective du sujet pour approfondir la problématique de l'appariement.

### **2.3.2) La folksonomie intégrée**

Le terme *folksonomie*, néologisme lié au développement du Web 2.0, est généralement défini par l'utilisation de *tags* (étiquettes) pour organiser l'information. Cette pratique s'inscrit dans un changement d'orientation vis-à-vis des méthodes habituelles d'indexation :

Traditionally, information resources are described, organized thematically, and classified either by experts, or by their creators. Using social bookmarking sites, users produce new sets of metadata by describing resources using generally uncontrolled keywords, referred to as tags. (Kakali et Papatheodorou 2010, 192)

Par contre, si l'on regarde ce phénomène avec un certain recul, le *tagging* (étiquetage) se veut une méthode très ancienne, voire primitive, d'organiser l'information :

Tagging is one of the oldest and simplest concepts of structuring our daily world. For instance, tags are used to remember which sort of jam was filled in a particular jar; you tag your luggage before the check-in at the airport; some people tag buildings with graffiti; and others tag resources on the web. Tagging means to attach a specific set of information – some sort of personal meta-data – to some object of our daily life. (Derntl *et al.* 2011, 1460)

Le *tagging* représente une activité dont la caractéristique principale est la liberté dans le choix du vocabulaire lors de l'indexation. Souvent présentées dans une liste ou un nuage de mots (Lawson 2009, 574), ces métadonnées personnelles (*personal meta-data*) composent le vocabulaire de ce qui, à l'intérieur d'un système documentaire, peut être appelée une *folksonomie*.

L'absence de contrôle, ou l'emploi très restreint des règles qui servent traditionnellement à la conception et l'utilisation des langages documentaires, représente donc un aspect qui distingue la *folksonomie* de la liste de vedettes-matières, du thésaurus ou encore du schéma de classification. Il s'agit à la fois d'une force et d'une faiblesse sur le plan de la recherche documentaire et de l'organisation de l'information (Le Deuff 2012, 65). Si la diversité terminologique occupe une place importante dans la *folksonomie*, reflet sur ce plan du vocabulaire diversifié des usagers<sup>46</sup>, c'est dire que cette dernière est aussi caractérisée par une terminologie « bruyante » (synonymes, homonymes, acronymes, etc.; voir Cantador, Konstas et Jose 2011, 2 ; Derntl *et al.* 2011, 1462). Il est bien connu que les termes d'indexation sous forme de *tags*, s'ils sont plus « flexibles » sur le plan de la représentation, ne permettent généralement pas d'atteindre une grande précision dans la recherche documentaire.

En contrepartie, l'un des avantages dans l'utilisation d'une *folksonomie* est l'élargissement de l'image ou l'affranchissement de la représentation d'un domaine de la connaissance :

Additionally, traditional system can be enhanced by contributions from article authors and taggers who provide a more expansive picture of the relevance and contributions of articles to a field of knowledge and to related fields. While the collection of author keywords and tags for supplementing descriptors would have been expensive in the past, the increasing move to electronic journal articles can actually be beneficial for subject indexing as it

---

<sup>46</sup> L'aspect suggestif des *tags*, surtout lorsque ceux-ci sont présentés à l'intérieur d'un nuage de mots ou d'un outil de navigation quelconque, est certainement un avantage de ce type de terme d'indexation pour la recherche documentaire. Il aide par conséquent à limiter ce que nous pourrions appeler librement le syndrome de la « boîte blanche », décrivant l'inconvénient manifeste de ne pas savoir par quel terme commencer une recherche documentaire.

provides access to more information such as author keywords and tags which can enhance the process of knowledge discovery. (Kipp 2011, 260)

Ce supplément d'information, qu'il vienne d'un auteur ou de ses lecteurs, peut améliorer la recherche documentaire et les activités liées à ce qui est souvent appelé *knowledge discovery*, en tant que paramètre de plus en plus mis de l'avant avec l'implantation des catalogues 2.0. En conformité avec le relativisme, une image plus développée d'un domaine de la connaissance contribue à la communication indirecte (voir la sous-section 2.1.1) qui naît de l'utilisation des métadonnées dans un contexte universitaire. Les usagers peuvent ainsi profiter de contributions catalographiques provenant de sources diverses et collaborer, à travers leurs propres prétentions au savoir (voir la sous-section 2.1.2), à une communauté partageant un ensemble d'objectifs convergents (sans nécessairement être consensuels).

La bibliothèque universitaire, si elle veut demeurer au cœur des activités académiques et combler le retard qu'elle accuse dans les services offerts en ligne (Ross et Sennyey 2008, 151), pourrait être bien servie en « mettant la main » sur les annotations de sa communauté d'usagers pour en organiser le contenu. Cet enjeu se trouve au cœur de l'approche centrée sur l'utilisateur dans la même mesure que la place occupée par le bibliothécaire, dans la conjoncture actuelle, est l'objet de questionnements et de débats. La contribution potentielle de l'analyste documentaire reste incertaine et les technologies 2.0, de leur côté, demeurent souvent en marge des services traditionnels offerts par la bibliothèque : « It is argued, the adoption of Web 2.0 technologies in general is very limited in libraries. When implemented, standards-based and socially constructed metadata approaches have been considered in isolation » (Alemu *et al.* 2012, 314). Les auteurs poursuivent en mentionnant qu'un cadre conceptuel (*conceptual framework*) plus général devrait pouvoir inclure les deux approches. De manière plus utilitaire, l'interaction entre l'analyste documentaire et les usagers constitue, encore une fois, l'axe à privilégier. Mesurée en termes d'annotations servant à formuler un/des énoncé(s) de recherche, la participation des usagers à un *projet folksonomique* favorise une plus grande collaboration, partant des conditions favorables aux interactions souhaitées. Il s'agit d'un « projet » puisque le but est d'intéresser graduellement les usagers à ce type d'outil d'organisation des connaissances, tout en y apportant une contribution liée à la tradition bibliothéconomique.

Cette participation usagère pourrait se traduire par la conception de catalogues personnels ou, pour utiliser un terme plus fréquent dans le milieu universitaire, par la production de bibliographies. Feinberg nomme ce type de groupement, lorsqu'il est constitué d'annotations (*tags*) choisies par un usager, une « bibliographie expressive » et elle soutient que cette dernière peut non seulement exprimer plus adéquatement le point de vue d'un lecteur mais aussi encourager le dialogue (Feinberg 2011). Elle argue par ailleurs que l'assistance fournie par le bibliothécaire dans la promotion de ce type d'analyse bibliographique est conforme à la vision d'Otlet, pour qui la bibliothéconomie devait encourager avant tout l'association entre personnes pour faciliter l'apprentissage (vision que partageaient, selon nous, d'autres fondateurs comme Cutter, Bliss, Shera et Ranganathan). Une analyse documentaire centrée sur l'utilisateur pourrait profiter de l'évaluation des exposés (*stories* ou *narratives*) sous-jacents aux bibliographies expressives :

I think, though, that we might make more progress toward these worthy visions [*i.e. idéaux inspirés d'Otlet*] not by trying to find and integrate all the perspectives that we can, but by figuring out what it is to tell a good story with a knowledge organization scheme and to help people, as both individuals and groups, tell their own good stories. Expressive bibliographies, in which audience members become authors, aggregating, annotating, and manipulating citations, can provide a means for many such stories to flourish, to engender, potentially, wide-ranging and spirited dialogue around a set of resources. (Feinberg 2011, 133)

Cette description de la bibliographie expressive se rapproche beaucoup du travail de l'étudiant ou du professeur qui analyse, annote et classe les documents qui contribuent à ses travaux. Parce qu'elle contribue à l'approche centrée sur l'utilisateur à travers une forme de « dialogue académique », la notion de bibliographie expressive s'applique avec pertinence, selon nous, dans le contexte universitaire. Un langage documentaire construit à partir d'un nombre suffisamment élevé de bibliographies expressives, elles-mêmes une extension des énoncés de recherche<sup>47</sup>, pourrait s'affirmer comme *folksonomie intégrée*, c'est-à-dire une *folksonomie* qui incorpore une terminologie liée directement aux intérêts de recherche des usagers.

---

<sup>47</sup> L'énoncé de recherche peut être vu comme un point de départ pour l'utilisateur ou, pour le dire autrement, ce qui motive sa recherche documentaire. Nous pensons que la bibliographie expressive peut incorporer non seulement cette terminologie de départ, mais aussi les termes et les concepts qui s'ajoutent au fur et à mesure que l'utilisateur consulte la documentation. D'autres énoncés de recherche peuvent s'ajouter et le champ d'investigation s'élargir.

La conclusion à laquelle aboutit Feinberg dans son article sur la bibliographie expressive est d'un intérêt considérable pour nous, car elle suggère le rôle de médiateur (ou *courtier documentaire*) que peut jouer le bibliothécaire:

If we, as a discipline, can discover and teach people how to effectively design, interpret, and criticize individual, limited, contingent knowledge organizations schemes, such as these expressive bibliographies, then we might find ourselves, as facilitators of empowerment, proceeding a little farther towards that ultimate goal of enlightenment. (Feinberg 2011, 133)

Profitant d'une tradition plus que centenaire dans le domaine bibliographique, la discipline de l'analyse documentaire peut jouer un rôle dans la promotion de bibliographies expressives en milieu universitaire. La *folksonomie intégrée*, dans sa version universitaire, pourrait désigner le résultat d'une analyse documentaire soumise à la caution usagère (*i.e.* orientée par les intérêts de recherche des usagers), et dont le matériau est intimement lié aux activités académiques des usagers (les énoncés de recherche principalement).

Pour arriver à un résultat probant, pour rendre compte des bibliographies expressives développées par les usagers, l'analyste documentaire doit faire ressortir à la fois les préférences terminologiques des usagers et les relations sémantiques engendrées (pour établir des liens hiérarchiques et associatifs au sein de la *folksonomie intégrée*). Cette dernière fonction, que l'on pourrait rapprocher de l'« épistémographie », peut s'appuyer sur la structure existante des langages documentaires utilisés. Dans la section 2.1, nous avons suggéré, à travers le terme « épistémographe », un rôle que pouvait jouer l'analyste documentaire dans une perspective relativiste. Avec la *folksonomie intégrée*, ce rôle se précise : favoriser la médiation entre le vocabulaire exprimé par les usagers<sup>48</sup> et celui que met en évidence une analyse documentaire centrée sur le document, de manière à résoudre la problématique de l'appariement ou, pour le moins, éclairer les enjeux épistémiques sous-jacents. À l'aide de cet

---

<sup>48</sup> Le plan de cours serait, dans le contexte de l'*épistémographie*, une source importante pour déterminer quels liens peuvent exister entre différents termes (ex. : « l'influence de la peinture réaliste sur les romanciers français au dix-neuvième siècle ») et l'évolution de ces liens à travers le temps (suivant les intérêts de recherche des professeurs). De plus, la bibliographie incluse à la fin d'un plan de cours fournit une liste de documents pour lesquels une analyse centrée sur l'utilisateur pourrait être particulièrement intéressante. Un lien cette fois entre un document mentionné dans la bibliographie et un ou plusieurs termes utilisés dans le plan de cours pourrait, en suggérant une représentation supplémentaire du contenu, être d'une grande aide pour un usager qui s'intéresse aux thématiques abordés dans un cours donné (même dans un contexte « hors programme »).

outil supplémentaire, nous croyons que l'utilisateur peut comparer plus aisément ses propres structures conceptuelles avec celles développées dans les langages documentaires, facilitant à la fois d'éventuelles recherches documentaires, à la fois l'atteinte de l'objectif pédagogique avancé par Soergel (la production de *conceptual frameworks*<sup>49</sup> ; Soergel 1999, 1119).

En conformité avec le relativisme, une telle *folksonomie intégrée* contribue donc à l'évaluation de prétentions au savoir (*knowledge claims* ; voir la sous-section 2.1.2), ces dernières s'affirmant à travers l'élaboration de bibliographies expressives (insertion d'un schéma conceptuel personnel au sein d'une structure de connaissances préexistante). Comme nous y avons fait allusion plus tôt, nous utilisons le terme « intégrée » pour caractériser une *folksonomie* qui prend sa source à même les énoncés de recherche des usagers. L'avantage manifeste réside dans la prise en charge d'intérêts et de besoins spécifiques à un groupe de personnes (selon nous, le domaine d'études devrait déterminer le regroupement le plus souvent privilégié ; voir la section 3.1). L'émergence d'approches portées davantage vers les besoins d'une communauté particulière, avec ses savoirs et ses expertises propres, semble être dans l'air du temps (*Zeitgeist*), manifestant la reprise d'un engouement similaire à celui affiché pour le thésaurus il y a quelques années :

Such a tendency perhaps represents the confluence of certain KO traditions— cf. the production of special classifications, indexes, and subject bibliographies geared towards particular user communities— with a postmodern *Zeitgeist* that both endorses a view of knowledge as formed by active interaction between people and the world rather than as resulting from a human program of cognitively “mirroring” the world in thought and valorizes multiple perspectives on what counts as knowledge while rejecting notions of an absolute truth. (Dousa 2010, 69 ; l'auteur souligne)

Nous revenons ainsi à une vision relativiste de la connaissance, comme fondation de l'approche centrée sur l'utilisateur, et à la nécessité de promouvoir les interactions entre divers acteurs au sein de l'université (étudiants, professeurs, chercheurs et bibliothécaires). Une *folksonomie intégrée* doit participer d'un foisonnement sémantique qui encourage le dialogue,

---

<sup>49</sup> Voir la sous-section 2.1.2.

la critique<sup>50</sup> et la recherche documentaire (avant tout en tant qu'inclusion d'une prétention au savoir dans une structure épistémique préexistante, tel qu'un langage documentaire).

La réussite d'une *folksonomie intégrée*, à l'échelle de l'université, doit être appuyée par une stratégie adéquate et non être motivée par le désir de simplement suivre la vague 2.0 :

It has also been suggested that existing standards-based metadata systems be re-evaluated, especially in light of socially constructed metadata approaches. Whilst, it has been remarked that these two metadata approaches do not stand in opposition to each other, interviewees have recommended that, libraries should embrace Web 2.0 technologies strategically, rather than adopting it for the sake of the technology. (Alemu *et al.* 2012, 338)

Une fois de plus, le recours aux énoncés de recherche permet d'entrevoir une *folksonomie intégrée* qui répond à cette exigence : elle pourrait éventuellement intervenir *en amont* de la recherche documentaire et, par conséquent, être pleinement incorporée dans les activités académiques des usagers. Ces derniers pourraient observer, en ajoutant un énoncé de recherche dans le système documentaire, de quelle manière se positionnent les éléments conceptuels mobilisés au sein des langages documentaires en vigueur<sup>51</sup>, de même qu'au sein d'une *folksonomie* façonnée à partir d'énoncés de recherche similaires. On pourrait de cette manière proposer deux indices d'utilisation complémentaires pour chaque terme, suivant son usage au sein de la communauté universitaire (présence ou non dans les énoncés de recherche des étudiants selon le cycle et/ou le domaine d'études par exemple) et sa présence dans les vedettes-matière ou descripteurs du catalogue (en précisant si le terme est utilisé pour décrire monographies, articles, documents audio-visuels, etc.).

Dans cette section, nous avons décrit un outil (sous-section 2.3.2) qui, dans le contexte d'une approche centrée sur l'utilisateur, serait en mesure d'éclairer certains éléments de la problématique de l'appariement (sous-section 2.3.1). La *folksonomie intégrée* peut enrichir l'environnement dans lequel l'utilisateur mène ses recherches documentaires, permettant un

---

<sup>50</sup> On pourrait donner, comme exemple simple, le professeur qui remet en question la place qu'occupe un concept dans un énoncé de recherche formulé par l'étudiant.

<sup>51</sup> L'avantage premier serait de constater si un terme est accepté ou non dans le langage documentaire dont se sert la bibliothèque. Il serait aussi possible de voir si le terme utilisé est subsumé sous une catégorie plus générale, ou encore, s'il peut être précisé davantage pour améliorer la stratégie de recherche documentaire.

double *mapping* de la terminologie qu'il souhaite employer dans la formulation de ses énoncés de recherche. Elle apporterait aussi à l'analyste documentaire un outil d'évaluation supplémentaire (le vocabulaire contrôlé adhère-t-il à la terminologie privilégiée par les usagers de la bibliothèque universitaire ?). Un tel projet repose, comme nous l'avons dit, sur la participation des usagers, mais aussi sur la capacité de promouvoir et d'utiliser l'expertise des bibliothécaires pour bonifier l'expérience universitaire. Quelle forme peut prendre la contribution spécifique de l'analyste documentaire? Cette question nous permettra, dans le troisième chapitre, d'entreprendre une réflexion portant sur certaines particularités de l'utilisation de la documentation en milieu universitaire et l'impact que ces particularités peuvent avoir sur l'analyse documentaire.

Dans le deuxième chapitre du mémoire, nous avons vu que l'approche centrée sur l'utilisateur se place dans une optique relativiste, voulant que toute connaissance humaine soit relative (Lalande 2010, 914). Une telle position cadre bien avec la dimension subjective du sujet telle qu'elle est décrite par Maniez (2005, 171) et promeut le recours à une caution usagère dont le matériau, en milieu universitaire, pourrait bien être les énoncés de recherche. Nous avons présenté le rouage principal de l'approche centrée sur l'utilisateur comme une *folksonomie intégrée* (susceptible de représenter les intérêts de recherche des usagers) et nous avons tenté de montrer quelle valeur celle-ci pouvait ajouter à une approche plus traditionnelle centrée sur le document. Il nous reste encore à préciser davantage comment la problématique de l'appariement se manifeste dans le milieu universitaire et de quelle manière l'analyste peut suggérer à l'utilisateur des indices pertinents pour mener avec succès ses recherches documentaires. Nous verrons, toujours dans la perspective de la problématique de l'appariement, que l'analyste doit tenter de se rapprocher du vocabulaire de l'utilisateur (exigences liées aux domaines d'études ; voir la section 3.1) et, qu'inversement, l'utilisateur peut devoir se rapprocher d'une certaine forme de vocabulaire contrôlé (exigences liées au système documentaire ; voir la section 3.2). Le troisième chapitre du mémoire sera consacré à ces deux aspects dont l'intérêt, pour l'analyse documentaire, repose sur certaines particularités de l'analyse des contenus en milieu universitaire.

## **Troisième chapitre**

**Particularités de l'analyse des contenus en milieu universitaire**

Dans ce troisième chapitre, nous voulons mettre en lumière certaines particularités de l'analyse des contenus en milieu universitaire et leur potentiel du point de vue des approches que nous avons examinées dans les deux premiers chapitres du mémoire. L'analyse des contenus, dans ce contexte, fait référence aussi bien à l'analyse documentaire (centrée sur le document et centrée sur l'utilisateur) qu'aux activités académiques des usagers (en particulier : recherche, lecture et analyse de la documentation offerte par la bibliothèque). Après avoir présenté deux approches générales de l'analyse documentaire, notre objectif est de repérer certaines particularités qui pourraient rapprocher le travail de l'analyste documentaire de celui des usagers en milieu universitaire et ainsi bonifier la communication indirecte (voir la sous-section 2.1.1) à laquelle contribue l'analyse documentaire. Dans la section 3.1, nous examinerons l'impact de l'analyse des domaines d'études sur le travail de l'analyste documentaire en milieu universitaire. Dans la section 3.2, nous décrirons comment une orientation analytico-synthétique peut diriger la contribution des usagers, rapprochant ainsi l'approche centrée sur l'utilisateur et l'approche centrée sur le document. Bien entendu, nous gardons en tête l'objectif central du mémoire qui est de proposer des pistes de réflexion pour élucider la problématique de l'appariement. Nous croyons que certaines particularités propres au milieu universitaire représentent des éléments de réponse intéressants.

### **3.1) Analyser les domaines d'études**

Dans cette section, nous tâcherons de montrer quel intérêt peut avoir pour l'analyste documentaire l'étude suivie du développement des disciplines universitaires. Il s'agit d'une fonction complémentaire aussi bien à l'approche centrée sur le document que celle centrée sur l'utilisateur, fonction qui permet de mieux comprendre le contexte dans lequel l'utilisateur fait appel à la documentation. D'ailleurs, l'analyse des domaines d'études (*domain analysis*) se trouve à mi-chemin entre un point de vue relativiste (deuxième chapitre du mémoire) et une approche objectiviste (premier chapitre du mémoire), comme le montre cet extrait tiré d'un article de Hjørland : « The information value of a specific access point *is relative to the conventions used in a specific domain or tradition* » (Hjørland 2002, 446 ; nous soulignons). Nous estimons que les disciplines universitaires représentent un ensemble de domaines spécifiques (modélisés le plus souvent sur la base de traditions académiques et/ou théoriques) agissant comme points d'ancrage pour circonscrire le contexte dans lequel un utilisateur a recours à la

documentation (dimension subjective). Nous n'aborderons pas l'analyse des domaines d'études en tant qu'approche particulière (voir Hjørland 2002 et Mai 2005 pour un tel propos), mais comme une orientation profitable pour l'analyse documentaire en milieu universitaire.

Un domaine d'études peut avoir une extension plus ou moins étendue, allant d'un thème large comme la géographie à un thème plus précis comme le développement durable. Nous pensons que les intérêts de recherche dont fait la promotion un département ou une faculté (à travers le corps professoral, les chercheurs et/ou les étudiants) représentent des unités signifiantes pour les usagers (par exemple, la géographie culturelle<sup>52</sup>). Nous pouvons constater qu'à l'intérieur d'une discipline telle que la géographie, divers champs d'études plus spécifiques se développent et s'imposent avec leurs conventions particulières (par exemple, le « développement durable » qui s'est grandement développé ces dernières années), alors que d'autres sujets d'études s'éteignent avec le temps (l'exemple bien connu de l'eugénisme et la phrénologie nous viennent à l'esprit). Le domaine d'études traduit bien, selon nous, la dynamique entre tradition et nouveauté mentionnée dans la sous-section 2.1.2 (consacrée aux prétentions au savoir ou *knowledge claims*, sur lesquelles nous reviendrons à la fin de cette section). Le domaine d'études est une notion qui, répétons-le, peut suggérer une certaine forme de complémentarité entre les deux approches examinées au préalable dans ce mémoire, entre les aspects objectifs du sujet (suivant un certain nombre de conventions) et ses aspects subjectifs (selon les intérêts de recherche de chacun et la manière de développer un énoncé de recherche).

L'analyse documentaire tient compte traditionnellement de caractéristiques appartenant à chaque domaine d'études (particularités de l'analyse documentaire dans un domaine comme les arts visuels par exemple), même si d'autres critères peuvent être jugés plus déterminants (nommément, l'impact de l'utilisation d'une vedette-matière sur les mesures de précision et de rappel). Qu'il s'agisse de développer l'armature d'un schéma de classification ou d'évaluer des questions lexicales et syntaxiques dans la formulation de termes d'indexation (vedettes-

---

<sup>52</sup> De manière à préciser les significations que les usagers peuvent donner à cette expression, le lien entre le terme « géographie » et le terme « culturelle » peut être examiné dans la perspective de la discipline plus générale qu'est la géographie. Cet exemple est tiré du site du département de géographie de l'université de Montréal. Voir le lien suivant : <http://geographie.umontreal.ca/recherche/expertises-de-recherche/>.

matière ou descripteurs), l'étude des domaines de la connaissance joue un rôle non négligeable. Lorsqu'on rapproche les notions de domaine de la connaissance et d'analyse documentaire, on pense généralement aux thésaurus consacrés à un domaine professionnel ou à un champ d'études. Si elle ne se traduit pas par une terminologie et une organisation aussi développées dans le cas des langages documentaires à vocation encyclopédique, l'étude disciplinaire n'en demeure pas moins un composant important de toute analyse documentaire.

Selon Talja, Tuominen et Savolainen, l'approche centrée sur les domaines d'études (*domain analysis*) peut être subsumée sous la catégorie des approches « collectivistes », lesquelles tendent à privilégier l'étude d'éléments socio-épistémologiques dans l'usage qui est fait de la documentation :

Collectivist approaches, especially domain analysis, adopt a sociological-epistemological view of information practices and relevance, rather than a user-subjective view. The basic assumption of domain analysis is that scientific domains have different languages, relevance criteria and ontological and epistemological commitments. (Talja, Tuominen et Savolainen 2005, 88)

Sans adopter une position subjectiviste (contrairement à l'approche souvent privilégiée dans les *user studies* ; Hjørland 2002, 430-432), l'analyse des domaines de la connaissance propose de relever les particularités épistémologiques, conceptuelles, langagières (langages spécifiques ou *languages for special purposes*) et terminologiques pour chaque secteur du vaste univers de la connaissance. Cette orientation théorique considère le domaine comme étant composé non seulement de connaissances à proprement parler, mais aussi des pratiques qui orientent le développement des savoirs : « Collectivist approaches are oriented toward a deeper understanding of the practices of professional groups and scientific domains, and the tacit knowledge underlying these practices » (Talja, Tuominen et Savolainen 2005, 88). Dans le contexte universitaire, ceci peut être étudié par le biais d'une orientation disciplinaire, dans la mesure où celle-ci met l'accent à la fois sur les aspects thématiques et pratiques des domaines d'études.

L'orientation centrée sur les domaines d'études contribue, à sa manière, aux critiques adressées à l'approche centrée sur le document (examinée dans le premier chapitre du

mémoire). Elle tend à nier que le sujet d'un document peut être extrait directement de ce dernier pour en alimenter l'analyse :

The subject matter does not exist in the document, ready to be discovered and pulled out by the indexer. The indexer will necessarily include knowledge of or guesses about the document's potential use and the users' needs in his/her analysis of the document. In other words, a true document-centered approach is problematic, and indexing will necessarily involve context in the indexer's final decisions about subject matter and index terms. (Mai 2005, 603)

Selon Mai, une approche centrée sur le document n'est pas en mesure de démontrer l'autonomie ou la neutralité complète de l'analyste documentaire et elle ne peut rendre compte adéquatement d'un aspect subjectif inhérent à l'indexation : « However, the document-centered approach does not address the inherent subjective interpretative nature of indexing and fails to show how the indexer might establish the subject matter independently of any context or use » (Mai 2005, 607). Le contexte dans lequel est effectuée l'analyse documentaire, tout autant que celui dans lequel les métadonnées résultantes sont utilisées, représente l'axe central de l'approche centrée sur les domaines d'études. Dans le milieu universitaire, l'analyste documentaire peut se servir à la fois d'une tradition objectiviste (premier chapitre du mémoire), à la fois de la contribution des usagers (deuxième chapitre du mémoire) pour se rapprocher le plus possible des exigences liées à l'analyse des domaines d'études<sup>53</sup>.

Comme dans tout milieu, l'analyste documentaire en milieu universitaire doit se demander si le fruit de son activité professionnelle répond aux attentes de sa communauté d'usagers. Sur ce point, Šauperl émet un constat plutôt négatif dans son étude de 2002 :

We do not understand yet how to describe the users for the cataloguer and how the cataloguer would tailor subject description for a user. [...] The other focus would be to study what demands different audiences have in terms of cataloging work. Should subject description be different for an academic or public library and what should the differences be? (Šauperl 2002, 164)

---

<sup>53</sup> Nous faisons allusion aussi bien aux exigences liées au caractère de la documentation dans un domaine d'études précis (les arts visuels par exemple) qu'aux attentes des usagers selon leurs intérêts de recherche. Comme nous l'avons souligné dans la section 2.3, la *folksonomie* et les langages documentaires traditionnels, une fois mis en parallèle, peuvent fournir une information importante concernant les métadonnées les plus pertinentes (ou les plus signifiantes) pour aider l'utilisateur dans ses recherches documentaires.

Comme question complémentaire, on peut se demander si le milieu académique offre des possibilités distinctives pour bonifier l'analyse documentaire. Une avenue intéressante pourrait bien être la classification documentaire, qui semble trop souvent négligée malgré ce qu'elle peut apporter à l'utilisateur en milieu universitaire. Parmi les divers types de langages documentaires, le schéma de classification est celui qui accorde le plus de place à l'approche disciplinaire, ce qui pourrait fort bien servir une orientation centrée sur les domaines d'études.

Plus qu'un langage permettant de situer « physiquement » la documentation, le schéma de classification permet d'organiser nos connaissances de manière systématique. La structure classificatoire, propice à la navigation sur une base disciplinaire, peut améliorer une recherche documentaire basée sur la catégorisation thématique :

An additional approach to providing subject access to the catalog would be through a hierarchical directory of resources arranged by category. To use such a directory, a library patron would begin by clicking on one of a number of very broad subject terms and then work through the hierarchy to a very specific term. Because catalog records already include Library of Congress classification numbers, this scheme could easily serve as the backbone of such a directory, although the terminology of this scheme would have to be adapted for use by the public. (Carstens et Buchanon 2004, 43)

Le schéma de classification éclaire une dimension importante de l'analyse documentaire, à savoir l'environnement disciplinaire et intellectuel dans lequel est positionné un sujet : « A third important point, not previously introduced, is that arrangement of documents or entries should take into account not only the basic, 'core' subject of a book, but also *the subject discipline or intellectual environment to which the basic subject belongs* » (Sayers et Maltby 1975, 55 ; les auteurs soulignent). Au-delà du noyau thématique (« core subject ») d'un document, le contexte disciplinaire ou intellectuel fournit une information supplémentaire pour organiser la documentation.

En milieu universitaire, l'analyse des disciplines est considérée, soit de manière implicite ou explicite (à travers, par exemple, l'appellation d'emploi « bibliothécaire disciplinaire »), comme une condition préalable pour améliorer la communication entre le bibliothécaire et les usagers. Dans le deuxième chapitre du mémoire, nous avons amplement parlé de l'impact positif qu'une communication indirecte plus détaillée pourrait avoir sur

l'analyse documentaire. La présente section aborde plutôt l'orientation centrée sur les domaines d'études pour mettre en lumière l'importance de facteurs contextuels propres à chaque secteur du vaste univers de la connaissance. La structure thèse/antithèse adoptée dans les deux premiers chapitres du mémoire nous invite à regarder le milieu universitaire dans une perspective plus nuancée, de manière à satisfaire un ensemble d'exigences complexes (liées à la documentation aussi bien qu'aux intérêts de recherche des usagers ; nous y reviendrons plus bas).

Modelées par une dimension contextuelle bien définie, les métadonnées sujet peuvent augmenter leur degré de pertinence et rendre compte de particularités disciplinaires. De manière générale, porter une attention plus grande aux valeurs et objectifs qui définissent un domaine d'études peut donner à l'analyste documentaire une perspective supplémentaire, de manière à opérer la synthèse entre les aspects objectifs (suivant un certain nombre de conventions en vigueur dans un domaine d'études) et subjectifs du sujet (selon les intérêts de recherche et les diverses manières de développer un énoncé de recherche). Une discipline tend généralement vers une forme de consensus, qui peut justifier l'analyse du sujet dans sa dimension objective, sans toutefois éliminer toute forme de contribution individuelle et originale :

Different scholars approach their work in different ways. Some are conservative, while others are radical, constantly pushing the boundaries. But this in no way negates the fact that disciplines have overarching conventions, consistent ways in which they live and advance, based upon what they commonly value and believe. (Badke 2012b, 107)

Cette citation reconduit, dans le contexte de la recherche universitaire, la dualité objectif/subjectif qui se manifeste dans la notion de sujet. Alors qu'une convention objective peut fortement encourager l'étudiant à se documenter pour éclairer les éléments d'un énoncé de recherche (ce qui participe de valeurs universitaires répandues), rien ne l'empêche d'user d'originalité dans le traitement analytique subséquent (en faisant appel, par exemple, à un document traitant d'un domaine d'études peu développé). L'utilisateur en milieu universitaire, par la nature des diverses tâches académiques, est le plus souvent confronté à une telle dualité lorsqu'il s'intéresse à un sujet.

Les particularités de chaque discipline modifient les modes de production documentaire et d'évaluation des savoirs, en partie parce que la nature de ceux-ci diffère selon le domaine d'études. À titre d'exemple, la communication scientifique est généralement plus structurée dans les « sciences pures » que dans les sciences humaines (*humanities*) :

We should be interested in investigating whether there are essential causes that can explain structures and differences. Such essential causes may be related to the nature of different domains. The scientific communication system may, for example, be more formalized and structured compared to the humanities because there are more objective criteria for knowledge. It is much more difficult in the humanities to outsource information retrieval and reviewing, for which reason the development of specialized database services, reviews, etc. should be expected to be less specialized in relation to the production of primary literature. (Hjørland 2002, 448)

La communication en milieu universitaire emprunte différentes formes (dissertation, compte-rendu, rapport de laboratoire, examen, communication orale, etc.), mais il n'est pas dans notre propos de les dénombrer. Nous voulons plutôt étayer l'utilité d'une approche disciplinaire en regard de la dichotomie objectiviste/subjectiviste inhérente à l'analyse documentaire (dichotomie présente à travers la notion de sujet). En gardant toujours en tête la problématique de l'appariement, il semblerait que le traitement du sujet oscille toujours entre des considérations liées aux attributs thématiques de la documentation et la présence incontournable de facteurs subjectifs. Une dichotomie analogue peut être repérée dans l'évolution des domaines d'études, puisqu'ils reposent à la fois sur une « information consignée » soumise à un processus objectif de validation (comme nous le verrons dans le paragraphe suivant), à la fois sur des activités critiques qui favorisent la réinterprétation continue des contenus documentaires et, par conséquent, l'émergence de nouvelles significations. Il nous semble important d'examiner cette dichotomie plus en détails, puisqu'elle a un impact sur notre compréhension des domaines d'études.

L'information, lorsqu'elle est consignée dans la documentation en milieu universitaire, participe d'une métaphore objectiviste bien implantée qui codifie l'accès aux collections. Il s'agit d'une métaphore confrontant l'utilisateur non seulement à l'imposante organisation physique de la documentation en bibliothèque, mais aussi à son pendant épistémique (l'organisation de la connaissance) :

Moreover, the subject organization of scientific documents has an objective correlative in the ordered arrangement of books and periodicals on the shelves of science libraries, which offers an impressive material metaphor of a massive and robustly stable organization of scientific knowledge. (Frohmann 2004, 247)

Cette « robuste organisation » de connaissances scientifiques est aussi à l'œuvre dans ce que Frohmann appelle nos pratiques documentaires et nos disciplines d'écriture :

Insofar as science mandates pursuit of universalising and objectifying projects, the disciplines of writing, the masses of documents, and their institutionalized apparatuses of organization and subject analysis become crucial to its historical and cultural configuration. Writing and documentary practices are central to it in ways that information and communication are not. (Frohmann 2004, 260)

Sur les plans historique et culturel, les pratiques documentaires et, en ce qui nous concerne plus particulièrement, l'analyse documentaire exercent une influence non négligeable sur le développement des domaines de la connaissance, suggérant un cadre de travail bien défini et appuyant un mandat à tendance universaliste et objectiviste. La notion d'*aboutness* (voir la section 1.1) n'est pas étrangère à un appareillage (*apparatuses*) dont le rôle est d'organiser la documentation de manière neutre, stable et pérenne. La tendance objectiviste dans l'analyse documentaire institue plusieurs mécanismes (lecture documentaire et réduction sémantique en particulier) qui circonscrivent l'analyse et lui confèrent une certaine stabilité et durabilité (par exemple, en évitant les effets de mode). Est-ce toutefois suffisant pour produire une image complète et à jour du développement des domaines d'études ?

Comprendre le développement des domaines d'études est une tâche complexe qui implique, entre autres, des aspects historiques, médiatiques (médias utilisés pour la transmission de l'information), épistémologiques (Hjørland 2002 ; Mai 2005, 608). Pour l'analyste documentaire, une perspective plus concrète peut être envisagée de manière à éclairer le contexte d'analyse le plus favorable. Une série de questions, posées en continuité, donne à l'analyse des domaines d'études sa spécificité et lui confère une orientation théorique définie :

The domain-centered approach will ask a series of questions, starting with questions about the domain, moving on to questions about the users, then to questions about the indexers, and finally will pose questions about the

document to be indexed. The indexer will address questions about the domain, users, and indexers in preparation for analyzing a document and ultimately indexing it. (Mai 2005, 608)

En posant ces questions, l'analyste documentaire recherche une information supplémentaire<sup>54</sup> dans le but de préciser davantage les aspects pertinents d'un sujet :

The indexer asks where the users fit into the larger domain to identify special topics that are of potential interest to the users, or if there is a specific bias in which the users are interested, or if there is a particular level of specificity that the users need. The indexer needs this information to identify the specific subject matter of the document that might interest or be useful to the users. (Mai 2005, 608)

Dans une certaine mesure, ces questions ont de tout temps influencé le développement des langages documentaires, mais elles doivent aussi servir à établir une correspondance plus grande, sur le plan de l'analyse des contenus, entre l'analyste documentaire et les usagers en milieu universitaire. L'objectif serait à la fois d'élargir l'éventail sémantique (quelles sont les significations possibles d'un énoncé de recherche selon le domaine d'études ?) et de préciser la profondeur sémantique exigée de l'analyse documentaire. Un *projet folksonomique*, tel que décrit dans la section 2.3, peut grandement aider à formaliser l'étude des disciplines universitaires de ce point de vue (en ajoutant une source d'informations qui obéit à la caution usagère ; voir la sous-section 2.2.3). Cette approche subjectiviste, développée au moyen des énoncés de recherche des usagers, peut informer l'analyste documentaire à propos des objectifs visés dans un domaine d'études (ex. : développer des modèles de gestion environnementale pour favoriser le développement durable) et d'un ensemble de valeurs partagées (ex. : encourager la biodiversité).

Dans tous les cas, l'analyse des domaines d'études est condition d'une certaine forme de parti-pris contextuel (culturel, épistémique, axiologique, etc.), l'analyste documentaire se trouvant à cette occasion dans une position qui exige de « dépasser » la dimension objective

---

<sup>54</sup> En consultant d'autres catalogues, l'analyste documentaire peut s'interroger, par exemple, sur le choix des métadonnées sujet utilisées par ses confrères en milieu universitaire. Il peut s'interroger sur son propre rôle dans le développement de guides disciplinaires ou encore dans le choix d'un système intégré de gestion de bibliothèque (par exemple, avec ou sans module permettant d'accueillir une *folksonomie* telle que décrite dans la section 2.3). L'information supplémentaire obtenue peut être utilisée dans le contexte d'un domaine d'études afin d'identifier certaines exigences dans l'analyse (ex. : l'indexation est-elle adaptée à la réalité des étudiants en pharmacologie? en histoire? en relations industrielles?).

du sujet. Il doit se demander, comme le laisse voir Mai dans la citation plus haut (2005, 608), si certaines particularités liées à la documentation peuvent intéresser les usagers œuvrant dans un domaine d'études. Ainsi, il peut justifier une analyse<sup>55</sup> dont le résultat doit être considéré comme un énoncé argumentatif en faveur de certains aspects du contenu :

The main concern is that the indexer identifies the domain's overall goals and purposes, its historical development, and its current research traditions. The ultimate goal of this analysis is to gain an understanding of the domain that enables the indexer to position an argument within the domain. (Mai 2005, 608)

En se basant sur la notion de *knowledge claim* examinée dans la sous-section 2.1.2, on peut interpréter la contribution de l'analyste documentaire comme une forme spécifique de prétention au savoir. Nous pensons qu'une telle prétention au savoir, parce qu'elle prend une forme argumentative (en-dehors du cadre traditionnellement neutre de l'analyse documentaire ; voir le premier chapitre du mémoire), doit se baser sur une caution usagère pour refléter les intérêts de recherche des usagers (voir la sous-section 2.2.3). Par conséquent, il est souhaitable de considérer l'apport de l'analyse des domaines d'études dans le contexte de la *folksonomie intégrée* telle qu'elle a été définie plus tôt (voir la section 2.3). Il s'agit d'un échange discursif entre analystes documentaires et usagers qui prend sa source dans les intérêts que ces derniers expriment pour telle ou telle notion (par l'entremise d'énoncés de recherche). De ce point de vue, le discours de l'analyste documentaire se situe à la croisée d'une approche centrée sur le document (dont la terminologie provient le plus souvent de langages documentaires traditionnels) et d'une approche *épistémographique* (voir la sous-section 2.1.3) susceptible de répertorier la terminologie utilisée par les acteurs d'un domaine d'études.

Dans cette section, nous voulions montrer qu'une orientation centrée sur les domaines d'études peut amener l'analyste documentaire à préciser son propre rôle dans le milieu universitaire de même qu'à déterminer dans quelle mesure les particularités d'un domaine d'études (traditions, objectifs, valeurs, développements récents) doivent être prises en compte

---

<sup>55</sup> Nous avons déjà vu que Mai considère que l'approche centrée sur le document, à elle seule, n'est pas en mesure de justifier adéquatement sa propre méthode d'analyse (Mai 2005, 603 et 607).

dans l'analyse documentaire. Parce que les domaines d'études sont relativement bien segmentés et les objectifs visés par les usagers sont défendus à travers un ensemble de discours structurés et argumentés (dissertations, rapports de stage, examens, syllabus, etc.), une occasion intéressante s'offre à l'analyste documentaire en milieu universitaire afin d'organiser la documentation suivant l'évolution raisonnée d'une discipline (à l'aide de métadonnées enrichies ; voir les sections 2.2 et 2.3). Les énoncés de recherche représentent, selon nous, le matériau le plus propice (sur les plans sémantique et stratégique<sup>56</sup>) pour l'approfondissement de la communication indirecte entre analystes documentaires et les usagers. Une orientation centrée sur les domaines d'études représente une expertise importante, un outil pour mieux comprendre les exigences spécifiques à chaque discipline universitaire. Une autre particularité de l'utilisation de la documentation en milieu documentaire met en lumière la nécessité de fournir à l'utilisateur les moyens de formaliser le résultat de ses propres analyses. C'est ce que nous allons examiner dans la section suivante.

### **3.2) Favoriser une démarche analytico-synthétique**

Dans cette section, nous tenterons de montrer qu'une orientation analytico-synthétique, tout comme l'analyse des domaines d'études (*domain analysis*), peut bonifier la communication indirecte entre l'analyste documentaire et les usagers, de même qu'encourager la participation à une *folksonomie intégrée*. Si, dans la section 3.1, la notion de domaine d'études a été utilisée pour permettre à l'analyste documentaire de mieux répondre aux exigences particulières des usagers en milieu universitaire, une orientation analytico-synthétique pourrait aider l'utilisateur, de son côté, à répondre aux exigences de l'organisation de la documentation. Nous voulons dire par là que l'utilisateur pourrait profiter d'une orientation analytico-synthétique, issue d'une longue tradition (parfois méconnue), pour orienter ses recherches documentaires et mieux comprendre les règles (syntagmatiques par exemple ; nous y reviendrons plus bas) et les choix qui gouvernent l'analyse documentaire<sup>57</sup>. Parallèlement,

---

<sup>56</sup> Ceux-ci sont propices sur le plan stratégique parce qu'ils initient, répétons-le, une communication sur la base des intérêts de recherche des usagers et parce qu'ils impliquent le plus souvent la poursuite d'une tâche académique ciblée.

<sup>57</sup> Voir la sous-section 2.3.1 concernant la complexité des règles d'analyse documentaire et les difficultés que cela engendre pour l'utilisateur.

une telle orientation pourrait favoriser chez l'utilisateur le développement de schémas conceptuels en accord avec les recommandations d'individus qui, dans le milieu universitaire, occupent des fonctions similaires ou complémentaires aux siennes.

Nous demandons au lecteur de faire preuve de patience concernant une définition du complexe notionnel « analytico-synthétique », celui-ci pouvant être expliqué plus clairement une fois placé dans une tradition d'analyse documentaire. En effet, pour bien comprendre ce qu'implique une démarche analytico-synthétique, il faut se rapporter au développement des langages documentaires au tournant du vingtième siècle. À cette époque, Kayser construit un langage d'indexation basé sur l'utilisation de catégories fondamentales : « Kaiser, who was the first to design such a language, named his categories *concretes*, *processes*, and *localities*. *Concretes* and *processes* correspond roughly to the subjects and predicates of sentence grammar » (Svenonius 2000, 129 ; l'auteur souligne). Si l'analyse documentaire sur une base catégoriale a donné une impulsion à l'approche analytico-synthétique (on doit aussi penser à Ranganathan bien sûr), celle-ci se réfère à une activité intellectuelle plus fondamentale encore. Portant bien son nom, elle se compose de deux étapes complémentaires (analyse et synthèse) dans l'étude d'un sujet, quel qu'il soit.

La théorie moderne de la classification, comme l'explique Chan, repose en partie sur un tel processus (analyse suivie d'une synthèse) dont l'avantage principal est d'isoler les diverses « facettes » d'un sujet :

Modern classification theory, on the other hand, places emphasis on *facet analysis* and *synthesis* - the *analysis* (or breaking up) of a subject into its component parts and the *synthesis* (or reassembling) of those parts as required by the document to be represented (Chan 2007, 312 ; l'auteur souligne<sup>58</sup>)

Une démarche analytico-synthétique présuppose donc la possibilité de décomposer un sujet documentaire en divers composants, puis de recomposer, à partir de ceux-ci, un/des énoncé(s) (condensation), un code (classification) ou un ensemble de termes (indexation) susceptibles de représenter adéquatement le contenu d'un document. Une définition précise est difficile à

---

<sup>58</sup> Voir aussi Broughton 2004, 257.

donner pour une telle démarche parce qu'elle est en quelque sorte diffuse dans le processus d'analyse documentaire lui-même. Comme nous l'avons expliqué avec la réduction sémantique (voir la section 1.3), analyse et synthèse sont nécessaires pour arriver à représenter l'*aboutness* d'un document (voir la section 1.1 pour la notion d'*aboutness*). L'objectif de cette section est plutôt de montrer en quoi une démarche analytico-synthétique (décomposition d'un contenu et recombinaison d'un sujet) peut améliorer la contribution des usagers au système documentaire (à travers la *folksonomie* ; voir la section 2.3) et favoriser une recherche documentaire à partir de schémas conceptuels bien structurés.

Dans sa forme classique, la démarche analytico-synthétique s'inspire largement de la linguistique<sup>59</sup>, en ce qu'elle vise à formaliser des éléments appartenant aux langues naturelles. Elle permet de recomposer le ou les sujets dont traite un document selon les éléments paradigmatiques sélectionnés et les règles syntagmatiques privilégiées<sup>60</sup>. En d'autres mots, et pour le dire simplement, l'analyste détermine, par exemple, si le sujet d'un document est mieux représenté par le terme « cuisine » ou le terme « peinture » (choix paradigmatique) et s'il convient d'introduire une forme de conjonction (selon des règles syntagmatiques prédéfinies) pour préciser davantage la thématique (« cuisine libanaise » ou « peinture impressionniste française au 19<sup>e</sup> siècle »). Pour donner une idée de la parenté profonde entre l'approche analytico-synthétique et l'évolution de l'analyse documentaire, on peut ajouter que la logique booléenne possède un mode d'opération similaire<sup>61</sup> et que celle-ci a joué un rôle primordial dans le développement des systèmes documentaires.

Le champ d'action le plus souvent associé à la démarche analytico-synthétique demeure la classification. Selon Broughton, la classification sous forme de facettes joue un

---

<sup>59</sup> Dans la citation que nous avons présentée au début de la section, Svenonius (2000, 129) fait allusion à l'héritage grammatical qui a caractérisé le travail de Kayser. Ceci n'est pas très surprenant dans la mesure où il s'agit d'un travail *sur* le langage (même si le résultat peut être un langage de nature « artificielle » et contrôlée). D'ailleurs, cette caractéristique est tout autant évidente dans la syntaxe de la *Colon Classification* de Ranganathan. Le RVM, les LCSH, la DDC, l'UDC, et même la LCC dans une moindre mesure, nécessitent aussi l'utilisation d'une syntaxe, une sorte de « phraséologie » pour combiner les éléments analysés de manière logique et constante.

<sup>60</sup> Pour une explication plus détaillée de la dualité syntagmatique/paradigmatique, voir Maniez 2002, 213-218.

<sup>61</sup> Ceci est vrai, en particulier, des connecteurs logiques « or » et « and » qui obéissent, respectivement, à un mode de fonctionnement paradigmatique et syntagmatique.

rôle important dans les méthodes actuelles de repérage de l'information (*information retrieval*) :

It is clear that faceted classification in some form or another now plays an integral part in most methods of information retrieval. It is well established as a method of construction in classification schemes and thesauri, and has affected the development of even the most conservative of systems in the area of traditional document description and organisation. It is popular as a navigational tool for web sites of all sorts, helping to structure all manner of objects and information about them, from children's shoes to the artistic output of the High Renaissance. It is beginning to be taken up by researchers in the fields of automatic indexing and the semantic web as a conceptual tool to assist in the understanding of the most complex relationships between objects. (Broughton 2006, 68)

Une panoplie d'outils et d'applications sont dérivés de l'approche analytico-synthétique, faisant de celle-ci un rouage important d'un grand nombre de systèmes documentaires. Dans le contexte des bibliothèques, l'utilisation de facettes pour améliorer les options de recherche dans les catalogues 2.0 représente l'indice le plus patent de l'influence (directe ou indirecte) exercée par des innovateurs comme Kayser, Ranganathan, Otlet, Lafontaine et les protagonistes du *Classification Research Group* (Frické 2011, 496).

Dans le jargon associé à la démarche analytico-synthétique, on parle souvent de schémas à facettes (*faceted scheme*) pour désigner un certain type de langage documentaire, dont l'un des avantages est de représenter clairement les relations entre concepts : « A second major feature of the faceted scheme is clarity in the expression of the relationships between concepts, both the intra-facet relationships (semantic relationships) and the inter-facet relationships (syntactic relationships) » (Broughton 2006, 54). Cette dénomination (*i.e. faceted scheme*) n'est pas des plus heureuses puisqu'elle invite la confusion entre les termes « catégorie » et « facette » (si l'on reprend le propos de Broughton ci-haut, il faudrait plutôt parler, selon nous, de relations intra-catégoriales et inter-facettes). Dans le contexte de l'analyse documentaire, les relations paradigmatiques sont de type intra-catégoriales (sélection, à l'intérieur même d'une catégorie, d'un concept plutôt qu'un autre) alors que les relations syntagmatiques s'expriment effectivement entre diverses facettes (appartenant ou non à la même catégorie). Cela vient du fait qu'une facette doit être considérée comme l'actualisation ou la réification d'une catégorie lors du processus d'analyse (ex. : le « Canada » est une

actualisation de la catégorie « lieu » ou « espace », et représente une facette d'une multitude de sujets). Elle doit être vue comme une qualité intrinsèque du sujet documentaire<sup>62</sup> :

In concluding, it must be emphasized that facets are not qualities of class numbers alone; nor are they peculiar to the scheme of classification used. On the contrary, facets inhere in the subjects themselves: that is, they exist in the subjects themselves, whether we sense them or not. Subjects will be helpfully featured, and their arrangement will be made filiatory and helpful within any scheme of classification if such a scheme is based upon facet analysis, and if the class numbers reflect the facets properly. (Ranganathan 1962, 81)

C'est pourquoi, dans l'approche analytico-synthétique issue de la tradition de Ranganathan, les facettes d'un sujet peuvent être variables, mais les catégories doivent être stables<sup>63</sup>.

La précision des termes est importante pour notre propos et notre objectif : nous voulons décrire une orientation théorique et pratique susceptible d'encourager la participation à un *projet folksonomique* (à travers, par exemple, la formulation des divers éléments conceptuels, ou facettes, qui composent un énoncé de recherche) tout en harmonisant l'analyse des contenus par l'entremise de catégories constantes. Une communication indirecte peut s'établir sur cette base : les analystes documentaires, de concert avec les bibliothécaires disciplinaires et (idéalement) les membres du corps professoral, établissent diverses catégories de manière à baliser l'analyse des contenus et, par extension, l'expression des facettes thématiques identifiées par les usagers (voir l'annexe B pour un exemple). L'approche centrée sur le document, parce qu'elle se restreint à la dimension objective du sujet, constitue un bon point de départ pour limiter le « bruit » dans la communication et pour mettre l'accent sur le noyau d'un sujet donné. L'utilisation de catégories générales constantes permet, de son côté, de représenter et d'organiser les thématiques à l'aide de propositions structurées et claires<sup>64</sup>. À

---

<sup>62</sup> Comme analogie, on peut penser aux « facettes de la personnalité ». Bien qu'il ne s'agisse pas de qualités tangibles, on peut dire qu'elles appartiennent à l'individu. Il est intéressant de noter que la catégorie principale dans la théorie de Ranganathan se nomme « Personnalité » et qu'elle identifie le noyau d'un sujet documentaire, même si cela est parfois accompagné d'une certaine ambiguïté (Ranganathan 1962, 83-84).

<sup>63</sup> Ces catégories stables (et universelles) sont : Personnalité, Matière, Énergie, Espace et Temps. Pour prendre un exemple simple, la catégorie « Temps » peut permettre de repérer, isoler et combiner par la suite la facette « 20<sup>e</sup> siècle » dans l'énoncé du sujet d'un document.

<sup>64</sup> Dans la suite de cette section, nous fournirons des exemples de catégories potentielles, mais celles-ci ne doivent pas être interprétées comme les seules possibles. Nous pensons que les usagers en milieu universitaires (étudiants, professeurs, chercheurs) se trouvent dans la situation idéale pour établir ces catégories et en évaluer périodiquement la pertinence (avec la contribution des bibliothécaires bien sûr).

titre d'exemple (*à noter* : nous identifions les catégories à l'aide de crochets), nous pourrions parler de la corrosion [action] de l'acier [matériau] causée par une solution saline [agent]. Cet énoncé se base sur des connaissances qui font consensus (conformément à l'approche centrée sur le document) et s'organise selon des catégories stables potentiellement utiles pour l'utilisateur. Elles peuvent fournir une information supplémentaire par rapport à un énoncé énumératif du type « corrosion et acier et solution saline », en précisant la nature des liens sémantiques entre concepts.

Nous venons de voir que les diverses facettes d'un sujet, une fois exprimées sous une forme propositionnelle (ex. : corrosion [action] de l'acier [matériau] causé par une solution saline [agent]), sont organisées de manière plus significative que dans une simple énumération. Dans le contexte d'une approche centrée sur le document, on obtient un tel résultat par l'entremise d'un processus récursif qui subsume une diversité de significations particulières sous une « macro-proposition » générale (Beghtol 1986, 90; voir la section 1.3). L'*aboutness* du document se révèle ainsi clairement, bien qu'à un niveau de généralité élevé :

This recursive process eventually yields an expression of the aboutness of the text in the form of the text's highest appropriate macroproposition, which is then used to create a statement of the text topic in the surface verbal elements available in the particular natural language. (Beghtol 1986, 91)

L'approche analytico-synthétique peut appuyer ce processus en isolant les diverses facettes selon une structure prédéfinie. En lien avec cette structure prédéfinie, une logique particulière se développe dans l'approche analytico-synthétique, logique que l'on pourrait appeler « catégoriale ». Comme nous l'avons souligné dès le début de cette section, il est nécessaire de rapprocher la démarche analytico-synthétique de l'étude de catégories générales pour l'organisation et la recherche documentaire.

Ce que nous appelons librement logique catégoriale a été examiné attentivement par de Grolier au début des années 60. Dans un rapport soumis à l'UNESCO en 1962, de Grolier énonce ce qui doit motiver une étude portant sur l'utilisation de catégories pour organiser l'information. Il s'agit de dépasser une expression langagière (naturelle ou artificielle) qui se base sur l'énumération de termes (une suite du type « corrosion, acier, solution saline »), pour préciser une multitude de relations logiques entre concepts :

Nous chercherons, d'une manière plus générale, à examiner l'ensemble des procédés par lesquels, dans une codification utilisée pour la recherche des informations contenues dans les documents, on peut éviter la lourdeur et la complexité croissantes de simples *énumérations* de termes, pour marquer les *relations* entre ces termes. (Grolier 1962, 16 ; l'auteur souligne)

Comment développer une codification (*i.e.* un langage artificiel) qui soit en mesure d'éclairer les relations qui unissent les complexes terminologiques (indices de classification et termes d'indexation) plutôt que de simplement énumérer ceux-ci ? Ces relations, comme nous l'avons déjà souligné, sont à la base de nature paradigmatique (cuisine *ou* peinture) et syntagmatique (cuisine *et* française). L'avantage de cette démarche fortement structuraliste (Grolier 1962, 158), du point de vue d'une approche centrée sur l'utilisateur cette fois (voir le deuxième chapitre du mémoire), est d'encourager la schématisation des énoncés de recherche de manière à questionner aussi bien l'appartenance d'un concept à une catégorie que les manières de combiner les facettes d'un sujet entre elles. Nous pensons que cela peut favoriser une convergence entre l'énoncé de recherche formulé par l'utilisateur, la terminologie la plus pertinente dans un domaine d'études<sup>65</sup> et un langage documentaire utilisé par la bibliothèque (en permettant d'identifier aussi les termes rejetés qui y sont répertoriés).

Le dynamisme de l'approche analytico-synthétique est un autre aspect qui en fait un outil intéressant, capable du point de vue de la classification d'accommoder tous sujets émergents (ce qui peut être plus difficile dans un schéma à tendance énumérative). Svenonius y voit l'une des forces de la *Colon Classification* de Ranganathan : « As a classification, it is both dynamic and parsimonious. Consisting of the constituent parts of knowledge, it is capable of expressing all subjects presently existing and those that might be evolved [*sic*] in the future » (Svenonius 2000, 175). L'utilisateur, s'il a l'occasion de développer ses propres schémas conceptuels sur une base analytico-synthétique, peut toujours y ajouter des concepts sans le souci d'une réorganisation constante. De plus, il peut garder une trace des relations sémantiques identifiées dans l'analyse initiale (ex. : troubles de l'apprentissage [noyau du sujet

---

<sup>65</sup> L'utilisation répandue d'un terme (mesure importante dans une *folksonomie*) de même que la fonction des usagers qui utilisent ce même terme constituent des indices qu'il importe de prendre en considération (voir la sous-section 2.2.3). La terminologie utilisée par un professeur dans son plan de cours pourrait, par exemple, avoir un poids plus important dans une *folksonomie* et représenter un ensemble de métadonnées associées à des documents jugés pertinents pour la réussite d'un cursus universitaire (documents présents très souvent dans la bibliographie du plan de cours).

ou « Personnalité » dans le jargon de Ranganathan] chez les enfants [entité] en situation de pauvreté [agent]) et confronter ses analyses subséquentes en isolant chaque élément, ainsi que l'appartenance de celui-ci à une catégorie donnée (l'utilisateur pourrait questionner si le terme « en situation de pauvreté » appartient à la catégorie [agent] ou encore à la catégorie [environnement ou milieu], lui permettant de développer un argument en faveur de l'une ou l'autre conception<sup>66</sup>). Décomposer des structures conceptuelles complexes et clarifier les relations sémantiques obtenues comptent parmi les forces de l'approche analytico-synthétique :

Facet analysis is significant for the clarity of the light it shines upon the relationships between objects and entities, and abstract concepts and their associated labels. It gives a rational, scientific, methodology for the construction of systems; it enables the full and precise description of objects of considerable structural complexity and of multi-dimensional semantic composition; it provides a flexible syntactical apparatus for the combination and ordering of concepts where this is required. (Broughton 2006, 69)

Analyser des contenus complexes et combiner par la suite une diversité de concepts comptent parmi les activités qui touchent aussi bien la recherche documentaire que les activités académiques des usagers en général. C'est l'une des forces de la démarche analytico-synthétique que de rendre ceci explicite, aussi bien lors de l'analyse de contenus que lors de la recherche documentaire comme telle.

Nous pouvons voir qu'une démarche analytico-synthétique apporte certains éléments de réponse à la problématique de l'appariement. Elle donne la chance à l'utilisateur d'organiser les concepts qui l'intéressent (extraits d'un énoncé de recherche par exemple) et de confronter le vocabulaire qu'il privilégie à celui présent dans un langage documentaire utilisé par la bibliothèque. L'utilisateur, lorsqu'il soumet un terme de recherche dans le système documentaire (ex. : « gestion de l'information »), peut visionner la liste des documents répertoriés sous ce terme dans le catalogue (le terme étant répertorié, par exemple, dans l'index des sujets). Les facettes d'un sujet, une fois identifiées et coordonnées à l'aide d'une syntaxe qui précise

---

<sup>66</sup> Ceci peut appuyer une prétention au savoir, telle que cette notion a été abordée dans la sous-section 2.1.2.

certaines des relations sémantiques qui existent entre elles<sup>67</sup>, apportent une dimension supplémentaire qui suggère à l'utilisateur d'examiner ses prétentions au savoir (voir la sous-section 2.1.2) dans le cadre d'une démarche structurée<sup>68</sup>. Créées à partir de contenus produits par professeurs et chercheurs, des catégories générales pour chaque domaine d'études peuvent, en clarifiant les concepts (à quelle famille appartiennent-ils?) et les relations qui les définissent, engendrer un dialogue plus riche sur la base d'une *épistémographie* telle que nous l'avons décrite dans la sous-section 2.1.3. Par extension, elles pourraient donc servir à organiser une *folksonomie* en milieu universitaire (voir la section 2.3), mais elles ne devraient pas, selon nous, modifier directement la structure d'un langage documentaire classique telle que le RVM.

Ultimement, l'approche analytico-synthétique permet, par l'entremise d'une structure prédéterminée, de dériver un texte à la fois expressif et réducteur à partir du contenu de la documentation :

For this reason [*un schéma de classification offre une version dérivée du contenu d'un document*], a classification system cannot be used to create original texts, but must be reflective and reductive in nature. From this perspective, attempts to increase notational hospitality and expressiveness by means of facet analysis and synthetic notations may be seen as attempts to increase our ability to use the classification system to create new texts. (Beghtol 1986, 99)

Ces « textes nouveaux » rejoignent la notion de bibliographie expressive (Feinberg 2011) abordée dans la section 2.3. L'objectif visé est de proposer aux usagers une méthode d'organisation (analytico-synthétique) propice à la recherche documentaire, sur la base de catégories qui leur sont familières ([matériau] pour un étudiant en chimie ou [groupe social] pour un étudiant en sociologie), tout en leur permettant de consigner le résultat de leurs propres analyses de contenu. La mise en relation de ces deux pôles est, nous semble-t-il, un excellent moyen de résoudre la problématique de l'appariement.

---

<sup>67</sup> Ces relations pourraient aussi inclure, par exemple, l'hyponymie et la synonymie selon ce que prescrit le langage documentaire utilisé.

<sup>68</sup> Une fois de plus, le plan de cours peut jouer un rôle pour encadrer ce processus. Il suggère à l'étudiant une démarche structurée (selon les thèmes qui segmentent le déroulement du cours par exemple) et conforme, de prime abord, aux exigences du professeur.

Pour conclure, une démarche analytico-synthétique favorise le dialogue et la critique à travers l'évaluation (et la réévaluation) de la pertinence d'un concept, d'une catégorie, d'un argument, etc. Parce qu'elle permet la « dissection » méthodique d'un sujet, elle sert aussi bien l'analyste documentaire que l'utilisateur pour qui l'analyse des contenus constitue une tâche académique. Nous croyons que l'intérêt de la démarche analytico-synthétique, de ce point de vue, est de permettre à l'utilisateur d'une part, de créer des schémas conceptuels sur la base de catégories stables jugées pertinentes par un grand nombre d'intervenants (professeurs, étudiants, chercheurs, bibliothécaires), d'autre part de comparer ces schémas (et la terminologie qui les compose) avec le vocabulaire contrôlé utilisé pour organiser la documentation en bibliothèque universitaire. C'est pourquoi une démarche analytico-synthétique est de nature complémentaire et qu'elle doit favoriser chez les utilisateurs une meilleure compréhension (même implicite) des exigences liées aux systèmes documentaires. Elle doit être vue comme un outil supplémentaire pour organiser de manière raisonnée une diversité de contenus.

Dans ce troisième chapitre, nous avons examiné les avantages que procurent une orientation centrée sur l'analyse des domaines d'études et une démarche analytico-synthétique. La première peut jouer un rôle quant à la problématique de l'appariement en plaçant l'analyse documentaire dans le contexte d'un domaine d'études et en identifiant le vocabulaire en vigueur dans celui-ci. La seconde permet à l'utilisateur de produire des schémas conceptuels raisonnés (conformes aux exigences d'un professeur par exemple) et pertinents pour la recherche documentaire, dans la mesure où ceux-ci peuvent être confrontés à la structure des langages documentaires utilisés en bibliothèque. Dans les deux cas, l'importance première est accordée au vocabulaire utilisé par l'utilisateur, le vocabulaire contrôlé servant plutôt d'appui pour encadrer deux stratégies susceptibles d'enrichir et compléter une analyse documentaire classique (centrée sur le document). Nous croyons que l'analyse documentaire peut s'orienter vers ce type de complémentarité en faisant un usage judicieux des nouvelles technologies (pour gérer une *folksonomie* par exemple), mais aussi en mettant à contribution une longue tradition. Un juste équilibre pourrait bien donner des résultats intéressants pour quiconque s'intéresse à la problématique de l'appariement.

## Conclusion

La problématique de l'appariement est accompagnée d'un questionnement qui met en cause la place traditionnelle qu'occupe l'analyse documentaire. Nous avons misé dans ce mémoire sur l'intérêt que peut susciter un réalignement plus ou moins important des procédés classiques d'analyse documentaire. Hudon et Mustafa El Hadi résumant ce questionnement en précisant certains aspects de la conjoncture actuelle :

Sous l'influence des réseaux sociaux et participatifs en pleine expansion, devons-nous bientôt constater la désuétude de tout processus d'organisation systématique des connaissances et de la documentation tel que mené depuis des générations ? Saurons-nous au contraire tirer profit de la tradition et de la systématisation pour créer des modèles et des outils d'organisation capables de s'ajuster aux besoins en constante évolution et à un appareillage technologique de plus en plus imposant ? (Hudon et Mustafa El Hadi 2010, 10)

Est-il possible donc de conserver les acquis d'une approche plus traditionnelle de l'analyse documentaire tout en exploitant le potentiel d'outils participatifs dans la mouvance du Web 2.0 ? Nous avons tenté de répondre à cette importante question en examinant deux facettes incontournables de l'analyse documentaire : le document et l'utilisateur.

En examinant la problématique de l'appariement, nous avons abordé l'analyse documentaire dans une perspective dichotomique. Ce choix vient de la dualité qui se situe au cœur de la notion même de sujet, laquelle possède à la fois une dimension objective et une dimension subjective. Les deux approches principales (document/utilisateur) ont par conséquent été présentées selon un modèle thèse/antithèse, de manière à souligner leurs différences ainsi que leur apport respectif à l'analyse documentaire (premier et deuxième chapitres du mémoire). À notre avis, le rapprochement de ces deux tendances générales représente un enjeu important pour bonifier les métadonnées sujet proposées en bibliothèque universitaire. C'est aussi pourquoi nous avons proposé par la suite deux orientations complémentaires susceptibles de faciliter la communication (indirecte) entre bibliothécaires et utilisateurs en milieu universitaire (troisième chapitre du mémoire).

Le premier chapitre du mémoire, consacré à l'approche centrée sur le document, a servi en quelque sorte de mise au point concernant l'analyse documentaire telle qu'elle est le plus souvent pratiquée. Trois grands axes ont été privilégiés : la dimension objective du sujet, le pragmatisme classique et le réductionnisme sémantique (hétérogène). Notre démarche se voulait avant tout exploratoire, dans la mesure où nous cherchions à mettre en lumière les fondements d'une activité intellectuelle hautement balisée. Il est reconnu qu'un outil tel que le langage documentaire, s'il se construit sur la base du contrôle du vocabulaire, peut assister avantageusement l'analyste documentaire dans son travail. Cette dimension doit, encore aujourd'hui, être mise de l'avant pour assurer le caractère pragmatique des métadonnées sujet et renforcer la performance des systèmes documentaires.

Dans le second chapitre, nous avons décrit une approche centrée sur l'utilisateur qui se concrétise dans une *folksonomie intégrée*, c'est-à-dire un outil terminologique ancré dans les énoncés de recherche des usagers en milieu universitaire. Le vocabulaire utilisé par les usagers, s'il est lié aux tâches académiques dont ceux-ci s'acquittent, peut représenter une source d'information significative pour l'analyste documentaire. Comme nous l'avons vu avec les dimensions relativiste de la connaissance et subjective du sujet, la documentation peut être utilisée pour répondre à des objectifs divers et soutenir des points de vue variés. Suivre étroitement l'acquisition, l'évaluation, le partage et le développement des connaissances s'affirme comme un objectif important pour la bibliothèque universitaire du 21<sup>e</sup> siècle. La *folksonomie intégrée* pourrait être au cœur d'une stratégie qui vise à élargir la communication indirecte entre analystes documentaires et usagers.

Avec la problématique de l'appariement, nous avons touché à un élément central de l'analyse documentaire : la correspondance entre les métadonnées sujet choisies par l'analyste et la terminologie dont l'utilisateur se sert dans ses recherches documentaires. L'utilité de l'indexation et de la classification est conditionnée au moins en partie d'une communication efficace aussi bien dans un sens (de l'analyste documentaire vers l'utilisateur) que dans l'autre (de l'utilisateur vers l'analyste documentaire). C'est pour cette raison que nous avons examiné, dans le troisième chapitre, deux orientations complémentaires dont le potentiel dans ce domaine nous semble apparent. D'une part, l'analyse des domaines d'études fournit à l'analyste

documentaire une image détaillée des objectifs que cherchent à atteindre les usagers en milieu universitaire. Elle permet en outre d'orienter avec une plus grande précision la formulation et le choix des métadonnées sujet susceptibles de répondre aux exigences de ceux qui contribuent directement à un domaine d'études. D'autre part, l'approche analytico-synthétique représente un outil puissant dont peut se servir l'utilisateur pour organiser les concepts qui caractérisent ses intérêts de recherche. La participation conjointe du corps professoral et des bibliothécaires représente un facteur important dans la recherche de catégories générales ayant pour but d'orienter les usagers dans la production de schémas conceptuels. Nous pensons que la combinaison de ces deux approches pourrait, dans les prochaines années, assurer un rôle prééminent à la bibliothèque universitaire et, plus particulièrement, à l'analyste documentaire.

Pour finir, nous voulons réaffirmer l'importance de questionner la place que peut occuper une approche centrée sur l'utilisateur dans le milieu universitaire. Qu'elle se concrétise dans la mise en place d'une *folksonomie* ou qu'elle s'intéresse plus directement aux stratégies qu'utilisent les usagers dans leurs recherches documentaires (en faisant usage de mesures statistiques ou de sondages par exemple), une telle approche, même si elle demeure « générale », peut apporter une information vitale à l'analyste documentaire. Nous pensons aussi que l'approche centrée sur le document et sa longue tradition peuvent illustrer certaines exigences liées à tout système documentaire et, encore aujourd'hui, influencer positivement le développement de l'analyse documentaire. En outre, nous avons examiné la problématique de l'appariement, parce qu'elle touche au cœur même de la dualité document/utilisateur et qu'elle sera encore longtemps la préoccupation de tout spécialiste (bibliothécaire ou autre) pour qui l'organisation efficace de la documentation représente un défi stimulant. Il est probable qu'un bon nombre de systèmes, théories, arguments, approches soient développés et oubliés, alors que la problématique de l'appariement évoluera en arrière-plan avec chaque essai, chaque erreur, chaque nouvelle idée.

## Bibliographie

- Alemu, Getaneh *et al.* 2012. The social space of metadata: perspectives of LIS academics and postgraduates on standards-based and socially constructed metadata approaches. *Journal of Library Metadata* 12, no 4 : 311-344.
- Anfinnsen, Svein, Gheorghita Ghinea et Sergio de Cesare. 2011. Web 2.0 and folksonomies in a library context. *International Journal of Information Management* 31 : 63-70.
- Badke, William B. 2012a. Save the subject heading. *Online* 36, no 6 : 48-51.
- Badke, William B. 2012b. *Teaching research processes: the faculty role in the development of skilled student researchers*. Oxford (UK) : Chandos Publishing.
- Banou, Christina, Petros A. Kostagiolas et Anna-Maria Olenoglou. 2008. The reading behavioural patterns of the Ionan University graduate students: a reading policy of the Greek academic libraries. *Library Management* 29, nos 6-7 : 489-503.
- Beghtol, Clare. 1986. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation* 42, no 2 : 84-113.
- Bliss, Henry Evelyn. 1929. *The organization of knowledge and the system of the sciences*. New York : Holt.
- Borlund, Pia. 2003. The concept of relevance in IR. *Journal of the American Society for Information Science and Technology* 54, no 10 : 913-925.
- Bouramoul, Abdelkrim. 2011. The semantic dimension in information retrieval, from document indexing to query reformulation. *Knowledge Organization* 38, no 5 : 425-437.
- Breeding, Marshall. 2010. *Next-gen library catalogs*. New York : Neal-Schuman Publishers.
- Brophy, Peter. 2005. *The academic library*. 2<sup>e</sup> éd. Londres : Facet Publishing.
- Broughton, Vanda. 2004. *Essential classification*. New York : Neal-Schuman.

- Broughton, Vanda. 2006. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings* 58, nos 1-2 : 49-72.
- Budd, John M. 1998. *The academic library: its context, its purpose, and its operation*. Englewood (Colo.) : Libraries Unlimited.
- Budd, John M. 2001. *Knowledge and knowing in library and information science: a philosophical framework*. Lanham (MD) : Scarecrow Press.
- Budd, John M. 2004. Academic libraries and knowledge: a social epistemology framework. *The Journal of Academic Librarianship* 30, no 5 : 361-367.
- Carstens, Timothy and Heidi Buchanan. 2004. The future of the catalog: a user-friendly academic search engine. *Technical Services Quarterly* 22, no 2 : 37-47.
- Chan, Lois Mai. 2007. *Cataloging and classification: an introduction*. 3<sup>e</sup> éd. Lanham ; Toronto ; Plymouth : Scarecrow Press.
- Contador, Iván, Ioannis Konstas et Joemon M. Jose. 2011. Categorising social tags to improve folksonomy-based recommendations. *Web Semantics: Science, Services and Agents on the World Wide Web* 9, no 1 : 1-15.
- Cutter, Charles Ammi. 1904. *Rules for a dictionary catalog*. 4<sup>e</sup> éd. (révisée). Washington : Government Printing Office.
- Cutter, Charles Ammi et Francis L. Miksa. 1977. *Charles Ammi Cutter, library systematizer*. Littleton (Colo.) : Libraries Unlimited.
- Debaise, Didier (coord. scientifique). 2007. *Vie et expérimentation : Peirce, James, Dewey*. Paris : Librairie philosophique J. Vrin.
- Derntl, Michael *et al.* 2011. Inclusive social tagging and its support in Web 2.0 services. *Computers in Human Behavior* 27 : 1460-1466.
- Dervin, B. 1983. *An overview of sense-making research: concepts, methods, and results to date*. Paper presented at the Annual Meeting of the International Communication Association, Dallas, Texas.

- Dousa, Thomas M. 2010. Classical pragmatism and its varieties: on a pluriform metatheoretical perspective for knowledge organization. *Knowledge Organization* 37, no 1 : 65-71.
- Ducheyne, Steffen. 2009. "To treat of the world". Paul Otlet's ontology and epistemology and the circle of knowledge. *Journal of Documentation* 65, no 2 : 223-244.
- Fairthorne, Robert A. 1969. Content analysis, specification, and control. *Annual Review of Information Science and Technology* 4 : 73-109.
- Farrow, John F. 1991. A cognitive process model of document indexing. *Journal of Documentation* 47, no 2 : 149-166.
- Feinberg, Melanie. 2011. Expressive bibliography: personal collections in public space. *Knowledge Organization* 38, no 2 : 123-134.
- Foskett, Antony Charles. 1996. *The subject approach to information*. 5<sup>e</sup> éd. Londres : Library Association Publishing.
- Frické, Martin. 2011. Faceted classification: orthogonal facets and graphs of foci? *Knowledge Organization* 38, no 6 : 491-502.
- Frohmann, Bernd. 2004. *Deflating information: from science studies to documentation*. Toronto : University of Toronto Press.
- Fujita, Mariângela Spotti Lopes, Vera Regina Casari Boccato et Milena Polsinelli Rubi. 2010. The indexing context for cataloging books using a sociocognitive approach. *Brazilian Journal of Information Science* 4, no 2 : 20-37.
- Grolier, Éric de. 1962. *Étude sur les catégories générales applicables aux classifications et codifications documentaires*. Paris : UNESCO.
- Harris, Jessica Lee. 1970. *Subject analysis: computer implications of a rigorous definition*. Metuchen (N.J.) : Scarecrow Press.
- Hjørland, Birger. 2000. Documents, memory institutions and information science. *Journal of Documentation* 56, no 1 : 27-41.

- Hjørland, Birger. 2001. Towards a theory of aboutness, subject, topicality, theme, domain, field, content... and relevance. *Journal of the American Society for Information Science and Technology* 52, no 9 : 774-778.
- Hjørland, Birger. 2002. Domain analysis in information science: eleven approaches – traditional as well as innovative. *Journal of Documentation* 58, no 4 : 422-462.
- Hjørland, Birger. 2005. Empiricism, rationalism and positivism in library and information science. *Journal of Documentation* 61, no 1 : 5-10.
- Hjørland, Birger. 2007. Information: objective or subjective/situational? *Journal of the American Society for Information Science and Technology* 58, no 10 : 1448-1456.
- Hjørland, Birger. 2008a. What is knowledge organization (KO)? *Knowledge Organization* 35, nos 2/3 : 86-101.
- Hjørland, Birger. 2008b. Core classification theory: a reply to Szostak. *Journal of Documentation* 64, no 3 : 333-342.
- Hudon, Michèle et Widad Mustafa El Hadi. 2010. Organisation des connaissances et des ressources documentaires : de l'organisation hiérarchique centralisée à l'organisation sociale distribuée. *Les Cahiers du numérique* (numéro thématique : Organisation des connaissances et Web 2.0) 6, no 3 : 9-38.
- Hudon, Michèle. 2013. *Analyse et représentation documentaires : introduction à l'indexation, à la classification et à la condensation des documents*. Québec : Presses de l'Université du Québec.
- Hufford, Jon R. 2013. Can the library contribute value to the campus culture for learning? *The Journal of Academic Librarianship* 39, no 3 : 288-296
- Hutchins, W. J. 1978. The concept of 'aboutness' in subject indexing. *Aslib Proceedings* 30, no 5 : 172-181.
- Jones, Bonna. 2008. Reductionism and library and information science philosophy. *Journal of Documentation* 64, no 4 : 482-495.

- Kakali, Constantia et Christos Papatheodorou. 2010. Exploitation of folksonomies in subject analysis. *Library and Information Science Research* 32, no 3 : 192-202.
- Kipp, Margaret E. I. 2011. Tagging of biomedical articles on CiteULike: a comparison of user, author and professional indexing. *Knowledge Organization* 38, no 3 : 245-261.
- Lalande, André. 2010. *Vocabulaire technique et critique de la philosophie*. 3<sup>e</sup> éd. « Quadrige ». Paris : Presses universitaires de France.
- Lancaster, Frederick Wilfrid. 1986. *Vocabulary control for information retrieval*. 2<sup>e</sup> éd. Arlington (Virg.) : Information Resources Press.
- Lancaster, Frederick Wilfrid. 2003. *Indexing and abstracting in theory and practice*. 3<sup>e</sup> éd. Champaign : University of Illinois, Graduate School of Library and Information Science.
- Lawson, Karen G. 2009. Mining social tagging data for enhanced subject access for readers and researchers. *The Journal of Academic Librarianship* 35, no 6 : 574-582.
- Le Deuff, Olivier. 2012. *Du tag au like : la pratique des folksonomies pour améliorer ses méthodes d'organisation de l'information*. Limoges : Fyp éditions.
- Lugwig, Pascal. 2011. Réduction et émergence. In *Précis de philosophie des sciences*, sous la direction d'Anouk Barberousse, Denis Bonnay et Mikaël Cozic. 232-263. Paris: Vuibert.
- Mai, Jens-Erik. 2005. Analysis in indexing: document and domain centered approaches. *Information Processing and Management* 41 : 599-611.
- Mai, Jens-Erik. 2011. Folksonomies and the new order: authority in the digital disorder. *Knowledge Organization* 38, no 2 : 114-122.
- Maniez, Jacques. 2002. *Actualité des langages documentaires : fondements théoriques de la recherche d'information*. Paris : ADBS.
- Menon, Bruno. 2007. Les langages documentaires : un panorama, quelques remarques critiques et un essai de bilan. *Documentalistes-Sciences de l'information* 44, no 1 : 18-28.

- Metcalf, John W. 1957. *Information indexing and subject cataloging: alphabetical: classified: coordinate: mechanical*. New York : Scarecrow Press.
- Moulaison, Heather Lea. 2008. Social tagging in the Web 2.0 environment: author vs user tagging. *Journal of Library Metadata* 8, no 2 : 101-111.
- Naudé, Gabriel. 2008. *Advis pour dresser une bibliothèque*. Paris : Klincksieck.
- Naun, Chew Chiat. 2006. Objectivity and subject access in the print library. *Cataloging and Classification Quarterly* 43, no 2 : 83-95.
- Otlet, Paul. 1989. *Traité de documentation : le livre sur le livre : théorie et pratique*. Liège : Centre de lecture publique de la Communauté française de Belgique.
- Pettee, Julia. 1946. *Subject headings: the history and theory of the alphabetical subject approach to books*. New York : H.W. Wilson.
- Radford, Garry P. 1992. Positivism, Foucault, and the fantasia of the library: conceptions of knowledge and the modern library experience. *Library Quarterly* 62 : 408-424.
- Radford, Garry P. 2005. Structuralism, post-structuralism, and the library: de Saussure and Foucault. *Journal of Documentation* 61, no 1 : 60-78.
- Ranganathan, Shiyali Ramamrita. 1962. *Elements of library classification: based on lectures delivered at the University of Bombay in December 1944 and in the schools of librarianship in Great Britain in December 1956*. 3<sup>e</sup> éd. New York : Asia Publishing House.
- Richardson, Ernest Cushing. 1964. *Classification, theoretical and practical: together with an appendix containing an essay towards a bibliographical history of systems of classification*. 3<sup>e</sup> éd. Hamden (Conn.) : Shoe String Press.
- Ross, Lyman and Pongracz Sennyey. 2008. The library is dead, long live the library! The practice of academic librarianship and the digital revolution. *The Journal of Academic Librarianship* 34, no 2 : 145-152.
- Šauperl, Alenka. 2002. *Subject determination during the cataloging process*. Lanham (Md.) : Scarecrow Press.

- Savage, Ernest Albert. 1946. *Manual of book classification and display: for public libraries*. Londres : George Allen and Unwin : The Library Association.
- Savolainen, Reijo. 2009. Epistemic work and knowing in practice as conceptualizations of information use. *Information Research* 14, no 1 : paper 392. [<http://InformationR.net/ir/14-1/paper392.html>]
- Sayers, W.C. Berwick et Arthur Maltby. 1975. *Sayer's Manual of classification for librarians and bibliographers*. 5<sup>e</sup> éd. Londres : Deutsch.
- Shera, Jesse Hauk. 1966. *Documentation and the organization of knowledge*. Londres : Crosby Lackwood.
- Soergel, Dagobert. 1999. The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science* 50, no 12 : 1119-1120.
- Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Cambridge (Mass.) : MIT Press.
- Talja, Sanna, Kimmo Tuominen et Reijo Savolainen. 2005. "Isms" in information science: constructivism, collectivism and constructionism. *Journal of Documentation* 61, no 1 : 79-101.
- Taylor, Arlene G. et Daniel N. Joudrey. 2009. *The organization of information*. 3<sup>e</sup> éd. Westport (Conn.) : Libraries unlimited.
- Université de Montréal, Département de géographie. *Présentation des expertises de recherche*. <<http://geographie.umontreal.ca/recherche/expertises-de-recherche/>> (consulté le 01 juin 2014).
- Wilson, Patrick. 1977. *Public knowledge, private ignorance: toward a library and information policy*. Westport (Conn.) : Greenwood Press.

## **Annexe A : Glossaire**

**Aboutness** : Indice langagier minimal servant à informer l'utilisateur de l'existence d'un document dont le contenu pourrait lui être utile. Cet indice s'insère le plus souvent dans l'index d'un système documentaire.

**Analyse documentaire** : Nous utilisons la définition donnée par Hudon : « Opération dont le double objectif est de déterminer les éléments d'information pertinents dans un document ou une collection de documents et de les représenter sous une forme concise et non ambiguë pour en faciliter le repérage ultérieur » (Hudon 2013, 264).

**Appariement** : Correspondance établie entre le vocabulaire utilisé par l'utilisateur dans ses recherches documentaires et celui mobilisé par l'analyste documentaire dans son travail.

**Caution bibliographique** : Important principe développé par Hulme qui prescrit l'inclusion d'un terme (ou d'un code) dans un langage documentaire seulement s'il traduit un contenu entériné par la littérature publiée (*published literature*).

**Caution usagère** : Selon la définition de la caution usagère donnée par Lancaster (1986, 26), l'inclusion d'un terme dans un index serait justifiée par l'intérêt démontré par les usagers d'une institution documentaire pour ce même terme.

**Énoncé de recherche** : Problématique(s) associée(s) à un sujet, fréquemment exprimée(s) sous une forme interrogative. Il est important de distinguer ce terme de la « question de recherche », qui sert à interroger directement un système documentaire (en menant une recherche dans le catalogue ou encore en faisant appel à l'expertise du bibliothécaire de référence).

**Entrée sujet spécifique (*specific subject entry*)** : Partons de la définition d'« entrée » proposée par Hudon : « Unité élémentaire d'un index composée d'une clé d'accès suivi d'un indice de localisation ou d'un renvoi » (Hudon 2013, 267) L'entrée sujet spécifique, qui

emprunte cette forme composée, traduit le sujet spécifique d'un document ou d'un ensemble de documents plutôt que la classe à laquelle appartient ce sujet (Cutter 1904, 22).

**Environnement de recherche documentaire** : Dans le contexte du présent mémoire, ce terme décrit l'ensemble des ressources langagières offertes dans un système documentaire pour permettre aux usagers de trouver diverses sources d'information (documents). Cette définition est volontairement circonscrite de manière à souligner les difficultés liées directement à la problématique de l'appariement.

**Épistémè (ou epistêmê)** : Terme tiré du grec qui, au sens premier, signifie « science » (l'épistémologie désignerait par conséquent « l'étude des principes, des hypothèses et des résultats des diverses sciences » ; Lalande 2010, 293). Par extension, l'*épistémè* renvoie aussi à la connaissance de manière générale, ce qui fait du néologisme « épistémographie » l'action de répertorier le vocabulaire servant à l'acquisition, l'évaluation, le partage et le développement des connaissances.

**Folksonomie intégrée** : La *folksonomie* est un langage documentaire qui répertorie les termes choisis par les usagers pour représenter le contenu d'un document, la fonction d'un item (ex. : « pot de confiture »), la/les thématique(s) d'une image, etc. La *folksonomie intégrée* est un cas particulier de *folksonomie* qui intègre une terminologie liée directement aux intérêts de recherche des usagers en milieu universitaire.

**Hyponymie** : L'hyponymie indique la relation d'inclusion d'un terme dont le signifié est hiérarchiquement plus spécifique que celui d'un autre terme (définition inspiré de l'article « Hyponyme » dans *Le trésor de la langue française informatisée*). Il est à noter que l'hyponymie entretient un lien étroit avec la notion de « subsumption » (voir l'entrée correspondante dans le glossaire).

**Langage documentaire** : Dans sa forme classique, le langage documentaire est le résultat d'un processus de normalisation dont le but est d'éliminer les ambiguïtés inhérentes aux langues naturelles (synonymie, homonymie, etc.). Empruntant le plus souvent la forme d'un thésaurus, d'un répertoire de vedettes-matières ou encore d'un schéma de classification, le

langage documentaire est utilisé pour indexer ou classer la documentation de manière à optimiser son repérage éventuel dans un système documentaire.

**Pragmatisme** : Philosophie définie par l'attention portée aux effets significatifs, aux gains pratiques qu'une idée (ou un concept) peut avoir sur notre expérience du réel (Lalande 2010, 803-807). Dans le contexte de l'analyse documentaire, l'accent est mis plus particulièrement sur l'expérience de recherche documentaire.

**Prétention au savoir (*knowledge claim*)** : Dans une perspective relativiste (voir l'entrée correspondante dans le glossaire), la prétention au savoir indique l'état incertain d'une connaissance (état momentané ou durable), telle qu'elle est formulée et située, par un individu ou un groupe de personnes, à l'intérieur d'un bassin de connaissances préalables. Dans le contexte de la recherche documentaire, cette notion peut suggérer une forme d'incertitude (faillibilité) concernant les termes qui composent une question de recherche (interrogation d'un système documentaire), une fois ceux-ci confrontés à un langage documentaire préexistant.

**Réduction sémantique inter-idiomatique** : Transition du contenu d'un document à son expression indexicale (voir la section 1.1 du mémoire) dans un système documentaire (le plus souvent à l'aide d'un langage documentaire). Cette expression indexicale est spécifique à l'utilisation du système documentaire (par l'entremise d'index), ce qui la place dans un idiome (un moyen d'expression) différent du contenu de la documentation.

**Relativisme** : Le relativisme est une « doctrine qui admet que toute connaissance (ou toute connaissance humaine) est relative » (Lalande 2010, 914).

**Scientisme** : Le scientisme exprime « l'idée que l'esprit et les méthodes scientifiques doivent être étendues [*sic*] à tous les domaines de la vie intellectuelle et morale sans exception » (Lalande 2010, 960-961).

**Subsomption** : Dans le contexte de l'analyse documentaire, la subsomption décrit l'action de placer un contenu particulier, tiré d'un document (ce qui constitue une distinction par rapport à

l'« hyponymie » ; voir l'entrée correspondante dans le glossaire), sous une classe ou une catégorie générique.

**Système documentaire indexical** : Système documentaire dont les outils langagiers de représentation pointent vers des valeurs à vocation neutre et stable (indices) consignées dans des index prévisibles (*i.e.* dont la terminologie appelle le plus souvent une signification consensuelle).

## Annexe B : Exemple d'énoncés de recherche

Nous présentons ici quelques exemples d'énoncés de recherche potentiellement significatifs pour orienter la recherche documentaire. Ces énoncés sont tirés d'un article de Hufford sur la contribution de la bibliothèque à la culture académique (« culture for learning ») qui prévaut dans un milieu universitaire. La forme interrogative met en évidence une forme de *déficit informationnel* auquel pourrait palier au moins en partie une recherche documentaire subséquente. Bien sûr, il s'agit d'un article déjà publié, et non de notes concernant une recherche éventuelle, mais la correspondance sur le plan formel nous semble significative et l'auteur parle lui-même de « research questions » supplémentaires :

- How does the library affect student behavior and attitudes?
- How do student reports of what they expect to find challenging on campus, including challenges posed by the library and its services, compare with what they actually experience?
- How has the college or university documented the connection between student well-being measures and learning? Does the library play a role in this connection? (Hufford 2013, 296)

Certains termes pourraient être extraits de ces questions pour donner une idée du point de vue privilégié pour aborder un domaine de la connaissance : *student, learning, behavior, challenges, library services, college, university, well-being*. Une démarche analytico-synthétique subséquente (voir la section 3.2) pourrait permettre, par exemple, de repérer le noyau (*learning*) du sujet, aussi bien que l'agent (*student*), le milieu (*college* ou *university*) et le moyen (*library services*) associés à la problématique soulevée. Ainsi, il serait possible de favoriser une certaine organisation de l'information **avant** que la recherche documentaire ne soit entreprise. La terminologie tirée d'un vocabulaire contrôlé pourrait être mise à contribution pour établir certaines correspondances avec les termes que souhaite utiliser l'utilisateur (les termes acceptés et les termes rejetés, par exemple, mais aussi les termes associés et subordonnés pour suggérer le niveau de précision sémantique le plus adéquat pour la recherche de documents pertinents).