



**HAL**  
open science

## Des enigmes de la recherche d'information

Olivier Perrin

► **To cite this version:**

Olivier Perrin. Des enigmes de la recherche d'information. domain\_shs.info.docu. 2008.  
mem\_00349234

**HAL Id: mem\_00349234**

**[https://memsic.ccsd.cnrs.fr/mem\\_00349234](https://memsic.ccsd.cnrs.fr/mem_00349234)**

Submitted on 25 Dec 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Université de Paris 8  
Vincennes-Saint-Denis**

**Sciences de l'Information Documentation**

## **Des énigmes de la recherche d'information**

**Olivier Perrin**

**Sous la direction de M. Claude Baltz**

**Mémoire de recherche**

**Master**

**Culture & communication**

Politiques et industries de la culture,  
de l'information, de la communication et des arts

**Numérique : Enjeux, Technologies  
M2 Recherche**

2007-2008

Toute connaissance est une réponse à une question. S'il n'y a pas eu de question, il ne peut y avoir de connaissance scientifique. Rien ne vient de soi. Rien n'est donné. Tout est construit.

Gaston Bachelard,  
*Le nouvel esprit scientifique*

Tout existe déjà, il faut seulement le retrouver.

Roland Barthes,  
*L'aventure sémiologique*, seuil, 1991, p.125.

## Remerciements

Mes remerciements vont à l'ensemble des professeurs des départements documentation et hypermédia de l'Université, et plus particulièrement à Claude Baltz, qui a accepté de diriger ce travail.

Je tiens à remercier également Jean-Marie Firdion, (Ined) pour l'aide précieuse qu'il m'a apportée dans l'élaboration du questionnaire et Georgia Leguem, de la Cité des sciences de la Villette, pour sa relecture et ses remarques pertinentes.

## Table des matières

<b>Introduction générale</b> .....	10
Une approche cognitive pour tenter d'appréhender un phénomène complexe.....	10
Problématique.....	12
Hypothèse(s).....	15
<b>Première partie</b> .....	18
<b>1. Société de l'information et stratégies de recherche d'information</b> .....	18
1.1. La connaissance comme réponse à la question de l'énigme.....	18
1.1.1. Cognition et information.....	19
1.1.1.1. De l'information psychologique à l'information physique.....	19
1.1.1.2. Architecture du cerveau humain : éléments d'information.....	19
1.1.1.3. La mémoire humaine comme contenant informationnel.....	20
1.1.1.4. Connaissance procédurale* et connaissance déclarative*.....	20
1.1.1.5. La connaissance : connaître et comment connaît-on ?.....	20
1.2. Information : bref aperçu sur un terme devenu un paradigme contemporain.....	21
1.2.1. Qu'est ce que l'information ?.....	21
1.2.2. Sens biologiques.....	21
1.2.3. Sens médiatique.....	22
1.2.4. Qu'est-ce qu'une information en documentation : essai de définition.....	22
1.2.5. La valeur de l'information.....	23
1.2.5.1. Connaissance et information.....	23
1.3. La société de l'information : histoire et concepts.....	24
1.3.1. La théorie mathématique de l'information de Claude Shannon [1948] : la quantité d'information mesurable.....	24
1.3.1.1. Qu'est-ce qu'une information qui répond à un besoin d'information?.....	25
1.3.1.2. L'information : un besoin pour un utilisateur de système d'information.....	25
1.3.1.3. Les sources d'information.....	25
1.4. Le système d'information.....	26
1.4.1. La place de la recherche d'information dans le système d'information.....	26
1.4.1.1. Les spécificités de l'information dans le contexte de la documentation.....	26
1.4.1.2. L'approche orientée usager.....	27
1.4.1.3. L'usabilité de l'information.....	27
1.4.1.4. L'information et sa représentation.....	27
1.4.1.5. Emetteur et récepteur : le message informationnel.....	28
1.5. Le dialogue avec l'utilisateur.....	28
1.5.1. La précision des critères de recherche.....	28
1.5.1.1. Identifier l'utilisateur.....	28
1.5.1.2. Le <i>rechercheur</i> d'information entre « cyber et culture ».....	29
1.5.1.3. L'information et son médiateur dans le processus de recherche d'information.....	29
1.5.1.4. Le <i>rechercheur</i> d'information est toujours un <i>Apprenant</i> .....	30
1.5.1.5. Le <i>rechercheur</i> d'information entre savoir et vérité.....	30

1.5.1.6. L'erreur comme source de progrès.....	31
1.6. Médiation et médiations.....	31
1.6.1. Le <i>rechercheur</i> d'information comme médiateur.....	32
1.6.1.1. L'entretien de référence.....	32
1.6.1.2. Le dialogue enseignant /enseigné.....	32
1.6.1.3. Des qualités nécessaires du médiateur dans le processus de recherche.....	33
1.7. Trouver l'information demandée.....	33
1.7.1. Brève histoire de la recherche d'information.....	33
1.7.1.1. Existe-t-il une science de la recherche d'information ?.....	34
1.7.1.2. Méthodologie(s) de la recherche d'information : un processus cognitif...36	
1.7.1.3. Les <i>technologies</i> de l'information.....	36
1.7.1.3.1. Technologie des stratégies de recherche et résolution de problèmes..36	
1.7.1.3.2. Stratégie de recherche d'information.....	36
1.7.1.4. Exemples de stratégies de recherche d'information simples.....	36
1.7.1.4.1. Les aléas de la recherche d'informations : Stratégies de recherche erronée.....	37
1.7.1.4.2. Stratégies menées sur des supports classiques : le dictionnaire papier.....	37
1.7.1.4.3. Modes de raisonnement dans le procès de recherche de l'information.....	37
1.7.1.4.4. Méthodologie de la recherche avec les opérateurs booléens.....	38
1.8. Outils de recherche.....	38
1.8.1. Le thésaurus : définitions.....	39
1.8.1.1. Les relations dans le thésaurus.....	39
1.8.1.2. Evolution des techniques de recherche d'information.....	40
1.8.1.3. Stratégies de recherche documentaire informatisée.....	40
1.8.1.3.1. Les requêtes : un nouveau paradigme ?.....	41
1.8.1.3.2. Analyse d'une question pour la recherche en langue naturelle.....	41
1.8.1.3.3. La recherche d'information sur le web dans les bibliothèques.....	42
1.8.1.3.4. Problématique de la recherche sur le web : éléments.....	42
1.8.1.3.5. Introduction indicative.....	42
1.8.1.3.6. La recherche des mots-clés pertinents.....	43
1.8.1.3.7. Qu'est ce qu'un moteur de recherche ?.....	44
1.8.1.3.7.1. Les balises Méta.....	44
1.8.1.3.7.2. Google : les clés d'une technologie efficace.....	44
1.8.1.3.7.3. L'algorithme du <i>PageRank</i> .....	45
1.8.1.3.7.4. La recherche par navigation.....	45
1.8.1.3.7.5. Interrogation et navigation.....	45
1.9. Apports de la recherche par navigation.....	46
<b>Conclusion de la première partie.....</b>	<b>46</b>
<b>Deuxième partie.....</b>	<b>48</b>
<b>2. Le couple questions-réponses et son corollaire la pertinence.....</b>	<b>48</b>
2.1. Le couple questions-réponses.....	48
2.1.1. De la question .....	48

2.1.1.1. Répondre aux questions : un usage social induit par la nécessité de communication professionnelle.....	48
2.1.1.2. Les lieux de l'information : des appellatifs génériques pour une même fonction.....	49
2.1.1.3. Question : éléments définitionnels.....	49
2.1.1.4..Qu'est-ce qu'une question ? Question et philosophie.....	49
2.1.1.5. La question en question.....	50
2.1.1.6. Questions et rhétorique.....	50
2.1.1.7. Questions : « Qu'est-ce que vous avez sur... ? ». Le mode interrogatoire.....	51
2.1.1.8. Questionner et Questionnement.....	51
2.1.1.9. Qu'est-ce qu'une question posée à un service de renseignement/information ?.....	51
2.1.2. Quelle est sa nature ?.....	52
2.1.2.1. Le contexte influe-t-il ? Comment ?.....	52
2.1.2. 2. Le couple questions-réponses dans le contexte documentaire.....	53
2.1.2.3. Une question <i>répond</i> à un besoin d'information.....	53
2.1.2. 4. La forme syntaxique des questions.....	54
2.1.2. 5. La fonction des questions.....	54
2.1.2. 6. Une question est une information.....	55
2.1.2.7. La formulation de la question comme un indicateur du besoin d'information.....	55
2.1.2.8. Essai de typologie des questions : le point de vue des professionnels de l'information.....	55
2.1.2.8.1. Des questions tous azimuts.....	56
2.1.2.9. Aspects temporels des problématiques de catégorisation de contenu des questions.....	56
2.1.3. Espace topographique de la question : le lieu de la question.....	56
2.1. 3.1. La question distribuée : une gestion anticipée des réponses.....	57
2.1.3.2. Analyse de la question.....	57
2.1.3.3 Les étapes du processus d'analyse de la question dans le contexte de la recherche d'information.....	57
2.1.3.4. Pourquoi est-il nécessaire de connaître (analyser) les questions dans un centre de documentation ?.....	58
2.1.3.5. Reformulation et dégradation de la question.....	58
2.1.3.6. Dégradation de la question.....	59
2.1.3.7. Bonnes et mauvaises questions.....	59
2.1.3.8. Questions simples et questions complexes.....	59
2.1.3.9. Pragmatique des questions.....	60
<b>2.2. Des réponses.....</b>	<b>60</b>
2.2.1. Qu'est-ce qu'une réponse ?.....	60
2.2. 2. Le stimulus : une réponse conceptuelle.....	61
2.2.3. Théorie de l'information et réponse.....	61
2.2.4. Essai de typologie des réponses : une pragmatique des réponses.....	62
2.2.5. La réponse comme contenu informationnel.....	62
2.2.6. Deux types de réponses sont obtenus au sein des systèmes d'information : Les réponses <i>clôturantes</i> et les réponses ouvrantes.....	62
2.2.7. Répondre et déontologie du <i>chercheur</i> d'information : le <i>droit</i> de réponse.....	63

2.2.8. Invariabilité et <i>prévisibilité</i> de la réponse : le cas du dictionnaire.....	63
2.2.9. Variabilité et imprévisibilité de la réponse comme message informationnel.....	63
2.3. La réponse comme élément du message informationnel.....	64
2.3.1. Le principe théorique de non <i>répétabilité</i> des réponses.....	64
2.3.2. Temporalité contextuelle des réponses.....	64
2.3.3. La réponse différée.....	64
2.3.3.1. Délimitation de la réponse.....	65
2.3.4. Valorisation de la réponse fournie.....	65
2.3.4.1. La réponse incertaine ou approximative.....	65
2.3.4.2. La réponse localisée.....	65
2.3.4.3. La réponse modifiée par les débuts de la recherche.....	66
2.3.4.4. La non réponse ou l'absence de réponse.....	66
2.3.4.5. Type de réponse erronée.....	66
2.3.4.5.1. La réponse experte ou spécialisée.....	66
2.3.4.6. Le niveau de réponse.....	67
<b>2.3.4.7. La construction de la réponse : éléments de réponses définitives</b> .....	67
2.4. La validation de la réponse dans les bibliothèques et centres de documentation.....	68
2.4.1. En bibliothèques publiques.....	68
2.4.1.1. La réponse immédiate sur les livres ou la réponse médiatisée du bibliothécaire.....	68
2.4.1.2. La réponse médiatisée par les livres ou la réponse par le livre (document).....	68
2.4.1.3. Dans les centres de documentation ou d'information spécialisés.....	69
2.5. La réponse <i>par</i> le web 2.0*.....	70
2.5.1. La réponse selon Google : une expertise virtuose ou virtuelle ?.....	70
2.5.2. Y a-t-il un lien entre une question et une réponse ?.....	70
<b>2.6. Les services de questions-réponses en ligne</b> .....	72
2.6.1. Les objectifs des services de questions-réponses dans les bibliothèques.....	73
2.6.1.1. Les services de questions-réponses à distance des bibliothèques.....	73
2.6.1.2. Services de questions-réponses aux publics : les exemples de « <i>Sindbad</i> de la Bibliothèque nationale de France (BnF) et de <a href="#">biblioSés@me</a> de la BPI (Bibliothèque publique d'information du Centre Pompidou), Paris.....	73
2.6.1.2.1. <i>Sindbad</i> de la Bibliothèque nationale de France (BnF).....	73
2.6.1.2.1.1. Règles et pratiques d'uniformisation et éléments d'activité du service : Indication quantitatives sommaires.....	74
2.6.1.2.2. Le service questions-réponses <a href="#">biblioSés@me</a> de la Bpi (Centre Pompidou).....	74
2.6.1.2.3. Le <i>Guichet du savoir</i> de la Bibliothèque municipale de Lyon.....	75
2.6.1.2.4. « <i>Questions ? Réponses !</i> » le service de l'Enssib.....	75
2.6.1.2.5. La participation des internautes : Les <i>Faq</i> ( <i>Frequently Asked Questions</i> ) ou « <i>Foire aux questions</i> ».....	76
<b>2.7. De la question à la réponse : la pertinence comme réponse documentaire évaluée</b> .....	76
2.7.1. Le triptyque pertinence- bruit – silence.....	76
2.7.1.1. Pertinence : éléments définitionnels.....	76
2.7.1.1.2. Le concept de pertinence en documentation et ailleurs.....	77

2.7.1.1.3. Pertinence et documentation.....	78
2.7.1.1.4. Toute petite histoire de la pertinence documentaire.....	79
2.7.1.1.5. L'usage contemporain distingue deux types de pertinence.....	80
2.7.1.1.6. La pertinence : une variable des réponses possibles ?.....	80
2.7.1.2. Efficacité de la recherche des moteurs et pertinence des réponses.....	81
2.7.1.2.1. Le tri par pertinence.....	81
2.7.1.2.2. Le tri par popularité : l'exemple de Google.....	81
2.7.1.2.3. Le mode de calcul de la pertinence selon Google.....	81
2.7.1.2.4. Critères de pertinence ou impertinence des critères ?.....	82
2.7.1.2.5. La pertinence objective : les mesures de la pertinence.....	82
2.7.1.2.6. Les mesures mathématiques simples de la pertinence documentaire..	82
2.7.1.2.6.1. Les mesures (calculs) complexes de la pertinence documentaire....	83
<b>2.8. Le monde du bruit et du silence.....</b>	<b>84</b>
2.8.1. Le bruit documentaire.....	84
2.8.1.1. Le bruit et la recherche plein texte ou texte intégral.....	85
2.8.1.2. Indexation libre et bruit.....	85
2.8.1.3. Recherche d'information par des noms propres et bruit.....	85
2.8.1.4. La <i>lutte</i> contre le bruit et le silence dans la recherche d'information.....	85
2.9. Le silence.....	86
2.9.1. Silence et indexation.....	86
2.9.1.1. Silence et recherche d'information : le cas des suffixes.....	86
2.9.1.2. Le silence et la recherche plein texte ou texte intégral.....	86
<b>Conclusion de la deuxième partie.....</b>	<b>87</b>
<b>Troisième partie.....</b>	<b>89</b>
<b>3. De quelques causes de l'échec documentaire.....</b>	<b>89</b>
3.1. Des causes d'échec documentaire.....	90
3.1.1. Principales causes linguistiques inhérentes à la langue française.....	91
3.1.1.1. Principales difficultés linguistiques et sémantiques.....	91
3.1.1.2. Polysémie des mots, dénotation et connotation.....	92
3.2. Causes linguistiques et langages documentaires.....	92
3.2.1. Causes techniques inhérentes à la langue anglaise.....	92
3.3. Causes liées à l'analyse et à l'indexation des documents.....	93
3.3.1. L'analyse documentaire.....	93
3.3.1.1. L'indexation documentaire : une pratique à haut risque.....	94
3.3.1.2. L'indexation au risque de l'indexeur et de l'idiosyncrasie.....	94
3.3.1.3. Métalangage et ethnométhodologie.....	94
3.3.1.4. L'indexation comme reformulation du document.....	96
3.3.1.5. La représentation du document : l'étude d'Annick Bertrand.....	96
3.3.1.6. La « surindexation » documentaire comme source de bruit.....	97
3.3.1.7. Recherche sur le terme « métadonnée ».....	98
3.3.1.8. <i>Rameau*</i> et <i>Le Métier de bibliothécaire</i> .....	98
3.3.2. Principales causes liées aux modes et usages professionnels de classement des références.....	99
3.3.2.1. Cas des périodiques.....	99
3.3.2.2. L'information cachée.....	99
3.3.2.3. Le traitement automatisé de la langue : indications sommaires.....	100
3.3.2.4. La lemmatisation.....	100

3.3.2.5. L'indexation automatique.....	100
3.4. De quelques causes liées à l'évolution incessante des technologies de l'information.....	101
3.4.1. Surabondance informationnelle.....	101
3.4.2. Erreur de manipulation.....	101
3.5. Principales causes psychologiques et intellectuelles.....	102
3.5.1. Erreur de navigation et surcharge cognitive.....	102
3.5.2. Causes psychologiques et psychosociales.....	102
3.6. Difficultés sémantiques et dialogiques.....	103
3.6.1. Un dialogue imparfait, partiel ou l'incompréhension mutuelle.....	103
3.6.1.1. Les obstacles à la communication interpersonnelle.....	103
3.6.1.2. Le respect de l'autre : l'absence d'écoute, d'empathie et de symétrie...	103
3.6.1.3. Management de la performance : stress et pénibilité au travail.....	104
3.6.2. Questionné/ Questionneur : une relation ambivalente.....	104
3.6.2.1. La proxémie ou la gestion des distances.....	104
3.7. Causes économiques : aperçu rapide.....	104
3.8. De quelques incidences politiques sur la recherche d'information.....	105
3.8.1. La manipulation de l'information.....	105
3.8.2. La censure : un filtre politico-moral de l'information.....	106
3.8.3. Les droits d'auteurs et l'accès aux documents.....	106
3.9. Les problématiques des mises à jour des ressources informationnelles liées à la durée de vie des pages web.....	106
3.9.1. Difficultés liées aux particularités fonctionnelles des moteurs.....	107
3.9.1.1. Google : une gestion des connaissances et de l'information accommodées.....	107
3.9.2. Critères de visibilité des moteurs de recherche sur internet : les problématiques du référencement.....	108
3.9.3. Le web invisible et ou profond.....	108
<b>Conclusion de la troisième partie.....</b>	<b>108</b>
<b>Conclusion générale.....</b>	<b>110</b>
<b>Annexes.....</b>	<b>113</b>
<b>Annexe 1</b> : Réponse envoyée par le service de réponses à distance de la BnF ( <i>Sinbad</i> ) à notre question sur les sources bibliographiques concernant les questions réponses, [juin 2007].....	114
<b>Annexe 2</b> : Document <i>Guichet du savoir</i> Bibliothèque municipale de Lyon.....	115
<b>Annexe 3</b> : Quelques sites de questions-réponses commerciaux et leurs domaines.....	117
<b>Annexe 4</b> : Questionnaire réalisé pour cette étude.....	118
<b>Annexe 5</b> : Glossaire des expressions et termes employés. Ils sont porteurs d'une*.....	126
<b>Annexe 6</b> : Bibliographie et références des sources citées dans ce mémoire.....	130

## Introduction générale

Précisons d'emblée que sous le terme d'énigmes se cache une partie du vaste ensemble des problématiques liées aux difficultés et aux obstacles rencontrés par celui qui cherche une information et se heurte à des problèmes inattendus voire insurmontés. Ces problématiques relèvent entre autres des sciences cognitives, (cognitivisme\*)<sup>1</sup>, [Ganascia, 1996], des Sciences de l'information et de leurs nécessaires interdisciplinarités.

La définition du *Tlf : Trésor de la langue française* [en ligne] est claire : une énigme « est un jeu d'esprit mettant à l'épreuve la sagacité de l'interlocuteur qui doit trouver la réponse à une interrogation dont le sens est caché sous une parabole ou une métaphore ». Pour l'auteur de ce mémoire, il s'agit, non d'un jeu d'esprit, sauf à considérer l'exercice de rédaction d'un travail universitaire comme tel, mais de tenter de trouver réponses à des interrogations suscitées par une partie de l'exercice professionnel de bibliothécaire.

L'acception du terme « énigme » portant sur le « Langage ésotérique (chiffré, figuratif, symbolique, etc.) pour initiés. *Enigme hiéroglyphique ; le mystère sacré, le secret de l'énigme* », ne concerne pas ce mémoire, puisque celui-ci fait appel *a priori* à la réflexion et aux commentaires scientifiques appuyés par des travaux de professionnels, universitaires ou non.

« L'énigme est le ressort le plus intime de la question, quand bien même celle-ci ne fonctionnerait plus en termes d' « être », [Axelos, 2005, p. 67].

### **Une approche cognitive : pour tenter d'appréhender un phénomène complexe**

Cette approche cognitive, qui couvre « le domaine de la communication humaine liée aux activités intellectuelles, -connaissance, compréhension, raisonnement- sera l'un des fondements de notre démarche, puisque l'approche cognitive « met l'accent sur le processus psychologique lié à l'interprétation et à la production de messages dans la communication », [Lazar, 2004, p. 41]. Mais l'approche cognitive concerne également la cognition, « c'est à dire l'activité par laquelle un système (humain ou artificiel) acquiert, stocke, récupère et utilise des connaissances sur l'environnement avec lequel il est en interaction et sur lui même », [Tricot, Rouet, 1998, p. 7].

Néanmoins, nous retenons les apports et les réflexions de Damasio, à propos du dualisme cartésien. Il ne saurait y avoir de séparation tranchée, entre le corps fait de matière et l'esprit immatériel, et nuançons « (...) la métaphore de l'esprit considéré comme un ordinateur », [Damasio, 2008, p. 313].

Il n'échappera à personne que cette présente démarche cognitive et heuristique\* que nous tentons de problématiser procède à son tour des sciences heuristiques, c'est à dire de la découverte. Une énigme n'est-elle pas une question précise à laquelle il faut trouver une réponse? Celle du Sphinx, ou celles de l'oracle de Delphes ?

---

<sup>1</sup> **N.b.** : une \* accompagne tout au long de ce mémoire certains termes techniques cités et explicités dans le glossaire de l'annexe 5.

Un des objectifs de cette étude est de montrer que, si tout responsable ou toute personne en charge de système d'information et de communication d'information, se *doit* de fournir des réponses et d'en évaluer la validité et la pertinence, le risque d'erreur, (qualifié ici **d'échec documentaire**), ou d'approximation, existe de manière inévitable comme dans toutes les entreprises humaines. Pour qu'il y ait échec documentaire, ne faut-il pas qu'il y ait au préalable tentative d'apporter une réponse par le moyen de la recherche documentaire, autrement dit, qu'une question soit posée ?

Sont confondues, pour une facilité de l'exposé, les expressions « recherche d'information », « recherche documentaire », voire « recherche bibliographique ». De même, toute personne, du néophyte au professionnel en recherche d'information, est qualifiée par commodité, au cours de ce travail du néologisme « *rechercheur* ». On objectera que les pratiques du néophyte sont sans commune mesure avec celles du documentaliste. C'est oublier que celui qui a un *besoin* d'information peut mettre en œuvre de nombreux dispositifs intuitifs opératoires et développer mille stratégies d'essais et erreurs et parvenir à un résultat probant. Sauf en de rares exceptions, nous nous limitons à l'espace langagier francophone.

Ce travail se situe dans le cadre d'un mémoire de Master 2. Il est nécessairement limité dans le temps et l'espace. Quelques questions et problématiques importantes ne sont qu'évoquées, d'autres pourtant indispensables sont volontairement ignorées. Enfin, devant la prolifération croissante des types de documents et supports numériques sur internet et ailleurs, sont privilégiés pour les nécessités de cette étude, les documents textuels.

Les pratiques de bibliothécaire et de documentaliste évoluent vers des « convergences et spécificités des métiers des archives des bibliothèques et de la documentation », comme le rappelle le sous-titre de l'intitulé de l'article « Journées d'étude interprofessionnelles », paru dans la revue *Documentaliste-Sciences de l'information*, 2005, vol.42, n°1, p .48-57

Un des contributeurs de ce compte rendu d'étude, Michèle Battisti, dressant le parallèle des métiers de bibliothécaires et de documentalistes, notait que « Si le souci de collection physiquement présente dans la bibliothèque a longtemps prévalu, c'est aujourd'hui le développement de services d'informations divers; intégrés et adaptés à une collectivité donnée afin de favoriser la découverte de l'espace insoupçonné de liberté qu'offre le libre accès à l'information et à culture » et, évoquant le travail des documentalistes, M.Battisti notait dans ce même article, que « la recherche documentaire est devenue plus complexe... », [Battisti, 2005, p.50].

En effet, ce paradoxe trouve une partie de son explication dans le fait que ceux qui s'adressent maintenant aux professionnels de l'information, le font après avoir tenté, avec plus ou moins de bonheur, de trouver seuls leur information. La puissance des moteurs de recherche, la connaissance des méta moteurs et leur « convivialité » étant mieux connues du grand public, celui-ci se retourne alors après ses infructueuses recherches vers le documentaliste, ce qui augmente *de facto* le niveau de spécialisation des interrogations et concomitamment la difficulté des recherches.

## Problématique

La recherche et la communication d'informations fiables restent, quelles que puissent être les évolutions technologiques de l'information et celles des métiers des professionnels de l'information, une fonction essentielle, toujours centrale dans l'exercice de ces professions. Pourtant, des difficultés, des aléas, peuvent survenir à n'importe quel moment des étapes de la recherche d'information<sup>2</sup>. Quelles sont les causes qui font qu'on ne trouve pas l'information recherchée, qu'on ne parvient pas à retrouver une information, qualifiée en retour de réponse, et qu'un utilisateur attend ?

Pourquoi, en certaines situations ou occasions, ne parvient-on pas à obtenir le résultat escompté ? Pourquoi, de l'utilisateur lambda au spécialiste en recherche d'information, se produisent des mises en situation d'échec documentaire ? Qu'est-ce qu'une question ? Existe-t-il des questions difficiles ? Des questions simples ? Quelles sont les raisons qui font qu'une *question simple* (pour tout un chacun) ne reçoit pas sa réponse *ad hoc*, ou qu'une requête avancée, menée par un spécialiste se brise sur les murs du silence\* ou du trop plein du bruit\* de l'interrogation ? Un échec ne peut-il pas être qualifié de réponse en sursis ?

S'agit-il des problèmes liés à la nature de la question ? A la nature de la réponse ? A la nature des domaines de connaissance interrogés ? S'agit-il des embarras engendrés par des difficultés liées aux stratégies de recherche<sup>3</sup>, à la méconnaissance des sources et/ou de leur usage (bases de données), des problématiques concernant la formulation du besoin d'information ? (Questions, quiproquo, pré-jugement, sémantique, etc.....). Un entrelacs de possibilités se cache, qu'il faudra essayer de démêler pour espérer, peut-être, trouver... réponse(s) à ces questionnements.

Mais la binarité oui/non des réponses aux questions est-elle inéluctable ? [El-Bèze, 2006, p. 277-297] parlant de questions booléennes. Existe-t-elle réellement ? Et un halo d'imprécisions des réponses est-il inévitable ? En effet, la réponse dépend aussi d'un contexte social et culturel, d'enjeux psychologiques, sociologiques et politiques. On ne saurait négliger ces aspects multiples et connexes qui interfèrent lors des processus d'interrogation puis de restitution (communication) de l'information.

### L'information introuvable, improbable ou *inconnue*

Evidemment, il ne s'agit pas ici de traiter de l'information qui n'a pas été révélée, pour le dire comme Nietzsche, parlant de *révéler des vérités muettes*. On

---

<sup>2</sup> Précision sémantique : concernant la recherche d'information nous ne faisons pas de distinction entre chercher une information (en anglais : *search*) et retrouver une information (en anglais : *retrieval*).

<sup>3</sup>(Nous ne tenons pas compte des concours de stratégies de recherche du type <[http://www.bases-publications.com/revues/bases/e-docs/00/00/03/87/document\\_article.phtml](http://www.bases-publications.com/revues/bases/e-docs/00/00/03/87/document_article.phtml)> réalisées, lors de salon professionnel tel *I-expo*). Consulté le 05/08/2008.

ne saurait, par exemple, trouver une thérapeutique vaccinale efficace du Sida pour la raison, hélas, qu'elle n'existe pas encore. On ne peut chercher que dans le réel de l'existant des informations disponibles.

D'une part, il ne saurait être question de problématiques soulevées par les informations confidentielles et personnelles, ou relevant du renseignement militaire ou policier, voire des informations de veille technologique difficiles à se procurer tout en restant dans la légalité. D'autre part, savoir par exemple, qui de Freud et de ses épigones a raison contre Sartre, qui enterre le concept d'inconscient pour mieux exhumer, en lieu et place, sa « mauvaise foi » prend le documentaliste en otage d'un débat dans lequel il n'a pas, selon nous, qualité pour répondre. Le documentaliste, (ou le médiateur), ne délivrant pas encore de certificat d'authentification.

Néanmoins, sur le plan de l'éthique professionnelle, le *chercheur* a le *droit* de ne pas savoir répondre, auquel cas, il se *doit* de réorienter l'utilisateur qu'il a en face de lui vers une structure appropriée ou une personne source.

L'obligation de résultat ne pouvant s'imposer, ou lui être opposée que dans des contextes où la question posée relève, comme dans notre exemple, de la gageure philosophique. Pourtant, le *chercheur* ne saurait faire l'impasse sur les questions de type tous azimuts, saugrenues ou farfelues au prétexte qu'elles n'entrent pas dans son champ de compétence.

Néanmoins, les interrogations contingentes ou usuelles du type : « Où est la sortie ? », « Où sont les toilettes ? », ne concernent pas ce travail.

Faut-il alors distinguer les problématiques soulevées par la recherche manuelle (ou classique) et celles nouvelles, qui seraient générées par la recherche automatisée ?

De même, la réponse fournie est-elle toujours *vraiment* pertinente ou bien ne contient-elle pas un halo d'imprécision et d'indécision inévitables? En un mot, la pertinence est-elle de ce monde de la documentation ?

Subséquentement, y a-t-il des réponses *vraies* ? Une réponse univoque peut-elle exister hors d'un contexte ?

Et les réponses calculées des logiciels sont elles réellement pertinentes ? En adéquation correcte ou exacte avec les questions ? Celles que soulèvent les problématiques concernant le concept de **pertinence**. Ce critère de pertinence, précisément défini par la norme ISO 11620, (il s'agit du degré de corrélation entre la réponse et la question posée).

La connaissance ou l'ignorance des sources, la connaissance ou la méconnaissance de leur usage induisent elles-mêmes des résultats d'échecs, (non-réponses). Pour répondre à quelques-unes de ces questions, un questionnaire a été réalisé en collaboration avec un sociologue de l'Ined, Jean-Marie Firdion, chercheur associé à l'équipe ERIS, (Equipe de Recherche sur les Inégalités Sociales), au Centre Maurice Halbwachs du CNRS, (Cf. Annexe 4). Le temps nous a manqué pour aller plus avant et nous souhaiterions vivement pouvoir pousser plus loin notre étude. Ce questionnaire serait proposé aux utilisateurs, (en priorité aux professionnels de l'information), via la liste *biblio.fr* ou celle de l'Adbs, afin d'en dresser, non seulement une typologie des utilisateurs

*désappointés* (catégories socioprofessionnelles, sexe, âge, etc., ...), mais surtout afin de connaître d'abord, comment et pourquoi ils n'ont pas su résoudre leurs besoins d'information, puis quelles en sont les conséquences pour eux-mêmes, et de savoir enfin si, avec un peu de recul, ou davantage d'expertise, on aurait pu résoudre positivement leur besoin d'information.

De même, et sans vouloir porter atteinte à l'honneur professionnel des spécialistes de l'information, il serait intéressant d'établir une typologie des interrogations qui n'ont pu recevoir de réponses satisfaisantes et que, pour faire court, on pourrait nommer : « questions en question ». Il existe sûrement des domaines de connaissance plus difficiles et plus délicats à appréhender. (Et tous, nous avons connu des échecs documentaires plus ou moins déroutants). Quels sont-ils et pourquoi y déambuler relève-t-il de la gageure ? A ce propos, l'expérience enseigne que les sujets (questions) en apparence les plus abscons ne sont pas forcément les plus difficiles à résoudre. Il est bien plus facile, dans le cadre d'une recherche bibliographique de trouver des informations bibliographiques pertinentes à l'intention d'un étudiant sur par exemple « le rôle de la bio luminescence chez les êtres vivants », si l'on ne se laisse pas hypnotiser par le seul mot scientifique inconnu, plutôt que d'établir une bibliographie sur « le rire chez Freud », second exemple dans lequel tous les termes de la recherche sont connus *a priori*.

S'interroger sur le pourquoi de l'échec documentaire conduit également à poser le questionnement inverse : pourquoi et comment trouve-t-on ? Les chemins de la logique conduisent-ils tous inévitablement à la *bonne* réponse ? Le hasard n'entre-t-il pour rien dans toute démarche interrogative ?

L'espace des bibliothèques publiques, parce que j'y travaille depuis plusieurs années, est un lieu d'observation privilégié, d'autant plus que les bibliothèques, voire médiathèques, ont su développer et dédier à leur public, des salles de lecture ou de documentation, les centres de documentation, spécialisés ou non, s'intégrant parfaitement aux problématiques soulevées par ce travail d'étude.

Nous examinons plus loin, quelques-uns des éléments du *paramétrage* qui constituent et gouvernent le corpus intelligible de l'interrogation et donc son processus en recherche documentaire. L'enjeu est de tenter de montrer en quoi la connaissance de ces éléments ou leur méconnaissance participe à l'efficacité ou non d'une démarche heuristique\*. Nous donnons quelques exemples de recherches qui n'aboutissent pas faute d'une maîtrise de la situation, (expérience), faute, pour le *rechercheur*, de connaissance dans l'organisation du savoir et de la structuration de l'information, faute enfin de l'absence de traitement efficace de l'information par des professionnels, (Exemple *infra* : recherches avec le terme « métadonnées » sur le site de la BnF). L'information étant révélée par des questions et les réponses *recouvrant* l'information.

Remarquons également que les explications très générales présentées dans cette étude peuvent paraître à certains égards schématiques, alors que les actions et les données produites s'inscrivent dans le contexte systémique des systèmes d'information dont nous souhaitons nous inspirer.

## Hypothèses

Une de nos hypothèses postule que tout *rechercheur* d'information, le néophyte comme le spécialiste, face à une question soumise par un utilisateur, est en **situation d'apprentissage (s) permanent(s) et de découvreur** et que de là peuvent découler des situations d'échec.

Ainsi, une question interroge aussi le sujet, (ses connaissances procédurales\* et déclaratives\*), à qui elle est soumise, (Double effet). Dans le domaine des Sciences de l'information, une question posée est déjà en soi une information. On peut aller au-delà, en écrivant qu'une question énoncée est une somme d'informations, en ce sens qu'elle contient des éléments intrinsèques à la question et des éléments extrinsèques, qui participent de son énonciation. S'y ajoutent le contexte de l'énonciation et celui « du langage comme médiation », [Corroy, Gonnet, 2008, p. 207-208].

**Une question est toujours un fait nouveau. Un message perturbateur.** C'est là notre seconde hypothèse fondamentale. Les linguistes diraient un hapax. Un événement linguistique et social, porteur de sens variables en fonction de l'énonciateur de la question, de son contenu informationnel, du contexte, contexte lui-même soumis à des reconstructions incessantes, [Lévy, 1993, p.26]. Comme tel, ce **fait nouveau** surgit semblablement à l'éclair et *bruit*, simultanément dans l'espace informationnel et à l'esprit, (cerveau) du récepteur de la question, dont il dérange ou perturbe les dispositions, mettant provisoirement fin à une situation homéostatique interne. Perturbation de l'ordre intérieur établi ; l'éclair figure de l'éblouissement, de l'étonnement qui aveugle l'esprit. Etonné, c'est être frappé. Sidération, etc..... Cette perturbation intellectuelle génératrice de bruit parasite le processus de la recherche d'information. Elle est d'ordre psychosociologique. Pour radicale qu'elle soit, cette posture épistémologique adoptée dans le cadre de ce mémoire permet d'en problématiser la question. Elle paraît d'autant plus pertinente, que : « Le mot « information » renvoie toujours à quelque chose de nouveau », [Baltz, cours de Cyberculture, 2008].

Reste une hypothèse sous-jacente, celle des conséquences liées à la société de l'information: le nécessaire apprentissage continu, que l'accélération ininterrompue du renouvellement des connaissances et des pratiques et usages rend indispensable, vient provoquer des perturbations cognitives et métacognitives.

Cet apprentissage n'est pas seulement la formation continue, organisée par les textes des lois successifs concernant la vie professionnelle, mais la formation continuée indispensable et continuant tout au long de la vie, aussi bien au travail comme hors de l'espace professionnel, pour rester *dans la course...* Il s'agit encore d'apprendre à apprendre.

Double mouvement piagétien d'assimilation et d'accommodation.

Mouvements d'accommodation (adaptation), devant la nouveauté technologique et d'assimilation à celle-ci, à des usages, à des procès nouveaux et continus : les nouvelles technologies de l'information et de la communication conjuguées à l'explosion vertigineuse des savoirs nécessaires à la connaissance.

Une autre composante de ces phénomènes perturbateurs, (*bruits*, parasitages divers) est due aux dispositifs techniques et aux problématiques de compétitivité (critères d'excellence).

Effets déstabilisants dus à la vitesse des innovations technologiques : (Voir le développement d'une nouvelle science : l'ergonomie)

L'évolution permanente et incessante des technologies de l'information renforçant ce concept d'apprentissage : procédures d'adaptation récurrentes à des dispositifs et interfaces logicielles. L'apprenant remet son savoir chaque jour sur le (s) métier(s). Il est, en situation professionnelle, ce néophyte que nous sommes tous chaque jour devant la vie. « Apprendre à s'informer c'est apprendre à apprendre et à tenir à jour ses compétences, dans un environnement en continuelle évolution, [Le Coadic, 1997, p. 111].

Cette notion d'apprentissage permanent est confirmée sur le terrain, si l'on peut *dire*, par Francis Agostini, conservateur qui fut longtemps chargé du service *Références* à la médiathèque de la Cité des Sciences de la Villette :

« Toute question est pertinente et légitime. La vulgarisation n'est pas une « verticale descendante » ; le lecteur traite d'égal à égal avec nous : « **je n'en sais pas plus que vous nous allons chercher ensemble** »<sup>4</sup>, [Souligné par nous].

Ainsi se pose la question du comment apprend-on son métier en bibliothèque et autres lieux de savoir ? C'est l'expérience et la formation sur le *tas*, qui enseignent comme l'enseignement scolaire réalise ses apprentissages, (lecture, écriture,...). On y apprend le plus souvent par mimétisme, (culture d'entreprise culturelle).

L'apprentissage est alors dit vicariant, [*vicarious Learning*] apprentissage par observation du comportement d'autrui.

L'*Apprentissage*, notion définie dans le glossaire de *Les hypermédias : conception et réalisation* [Saleh (dir.), 2005 p. 309], comme « processus conscient que suit un individu pour acquérir des connaissances, (selon la terminologie de S.D. Krashen qui oppose « apprentissage » et « acquisition ». L'apprentissage a à voir avec l'*usabilité*, (Cf. *Infra*). Formation et information : se former, c'est s'informer.

Il ya un rapport notable, évident, (origine étymologique *forme*) entre formation et information : la première notion désigne un dispositif qui permet des apprentissages de nouveaux savoirs, dans la seconde notion, l'information, du savoir est mis à disposition de l'apprenant.

Trois grandes parties divisent notre travail. La première partie de ce mémoire, après avoir interrogé la matière informationnelle, c'est-à-dire, comment l'esprit connaît et accède à la connaissance, situe la recherche d'information dans un contexte général d'usages professionnels en explicitant les principales étapes, particulièrement celles des médiations et des interfaces technologiques. En présentant quelques-uns des outils et des technologies auxquels le *rechercheur* d'information est amené à se confronter de manière incessante, à cause de l'évolutivité et du renouvellement de ces technologies, qu'accompagne l'explosion des savoirs qui interfèrent dans le mode d'interrogation, nous soutenons que le *rechercheur* est toujours en situations d'apprentissages, créatrices de *bruit* et de *silence* documentaires.

---

<sup>4</sup> *Bibliothèques(s)*, oct. 2007, n°34-35, p.10.

La seconde partie interroge le couple questions-réponses et la pertinence, qui forment le fondement de ce triptyque documentaire. Nous tenons que la question et la réponse sont des objets de connaissance et que c'est la question qui mène à la connaissance, (réponse). Plus précisément, des séries de questions et de réponses mettent au jour l'information. En postulant que la question, est le message informationnel initial, nous souhaitons montrer que le questionnement est le processus médiateur (série de questions-réponses) des interactions qui lient le cerveau du *rechercheur et ses représentations*, celui de l'utilisateur, avec la connaissance et l'ensemble des systèmes d'information. Puis, au travers des problématiques de pertinence, nous posons la question de la validation des informations distribuées aux utilisateurs par les bibliothécaires. En outre, nous tentons de savoir s'il existe des liens de correspondance, autres que la pertinence, entre une question et une réponse. Enfin, nous dressons un panorama des services de questions-réponses en ligne mis en place récemment en France dans différentes structures.

La troisième partie de ce travail évoque quelques unes seulement des causes possibles de l'échec documentaire (non réponses). Il s'agit de savoir pourquoi l'information reste parfois introuvable. L'impossibilité où nous sommes de trouver nous-mêmes des études qui montreraient des professionnels en situation d'échec documentaire (non réponse) - ce qui ouvre de nouvelles perspectives de recherche - nous a conduit à présenter sommairement quelques unes de ces causes d'échec qui concernent les professionnels. Faute de temps, nous n'avons pu effectuer un balayage exhaustif des éventuelles études qui existeraient, en langue française, sur ce sujet et qui affecteraient les professionnels de l'information. Nous pensons, sauf erreur de notre part, et sources qui nous auraient échappées, que ces études n'existent pas, ou sont difficiles à trouver. En ce cas, elles entreraient, par effet réflexif, doublement dans notre sujet, puisque cette absence, voire cette difficulté, constituerait une cause d'échec documentaire....pour l'auteur de ce mémoire.

C'est pourquoi notre attention s'est portée, dans un premier temps, sur les problématiques du langage desquelles sourdent la plupart de causes d'échecs documentaires. Nous avons tenté de dresser, dans un second temps, un panorama, non exhaustif et donc partiel, des causes psychologiques et sociales qui interfèrent dans le processus d'échec.

# Première partie

## 1. Société de l'information et stratégies de recherche d'information

Nous traitons de la question et des réponses documentaires comme des « objets de connaissance », des objets qui servent aussi à connaître, à interroger et à répondre dans le réel. A produire de la connaissance. Connaissons qu'il y a d'abord la question au singulier et que le pluriel des réponses entraîne d'autres interrogations. Questions et réponses servant ainsi tels des outils conceptuels, d'abord à piocher dans le réel, (questionner) puis à construire, c'est à dire à représenter les connaissances humaines, (répondre).

### 1.1. La connaissance comme réponse à la question de l'énigme

« Toute connaissance est une réponse à une question. Et s'il n'y a pas eu de question, il ne peut pas y avoir de connaissance scientifique » a écrit Bachelard.

Cette citation sera le fil rouge ou épistémologique de notre tentative pour tenter de résoudre l'énigme, ou les énigmes sous-jacentes de notre questionnement, que nous voulons éclaircir. Mais il nous faut d'abord établir des rapports et des filiations entre la connaissance, l'acte de connaître et l'information, vérifier la longueur et l'épaisseur des liens qui unissent ces termes, en faire une brève généalogie exploratoire parce qu'il s'agit encore de parler de la connaissance de la connaissance, [Besnier,1996], de savoir comment l'homme connaît et de montrer que c'est au fur et à mesure de ses échecs multipliés avec entêtement de chercheur qu'il peut espérer parvenir à ses fins : connaître et savoir. Car rappelle le philosophe « Rien n'est donné ».

Avant de parler de l'objet de cette étude, des outils et des systèmes d'information et de leurs interactions, il paraît indispensable de définir la notion d'information et d'en montrer l'importance au sein de la société où l'information projette ses rhizomes et tisse ses toiles réticulaires d'un grand nombre de réseaux, informationnels, neuronaux, génétiques, psychologiques, cognitifs, politiques, etc..

Le terme d'information est fondamental en Sciences de l'information où il a suscité et suscite encore un grand nombre de tentatives de définitions, de débats et de controverses épistémologiques. Il recouvre en effet des significations protéiformes très diverses, tantôt objectives, tantôt subjectives. Même si l'on s'écarte de la définition que Claude Shannon (1916-2001) lui a donnée, comme de toutes celles qui relèvent de la mesure d'un degré d'organisation, les sens possibles restent nombreux. Il serait bien incertain de prétendre en faire maintenant la recension exhaustive, c'est pourquoi nous nous bornerons à examiner *infra* quelques-unes des définitions proposées ci-après et à recenser à partir d'elles des notions reliées à l'information.

C'est la raison pour laquelle notre travail, en premier lieu, portera un intérêt particulier vers des problématiques connexes, liés à la connaissance mais toujours en rapport étroit avec celles de l'information.

Car celui qui interroge l'interrogation, qui pose la question de la question, de sa résolution ou non dans l'espace documentaire ne peut faire l'impasse sur la question de la connaissance, du « comment on connaît ? » et de ce que c'est que de connaître.

### 1.1.1. Cognition et information

Le terme « cognition » vient du latin *cognitum*, « connu », forme nominale de *cognoscere*, « connaître ». Utilisé en français dès le XIV<sup>e</sup> siècle, il prend au XVI<sup>e</sup> siècle, le sens de « ce qui concerne la connaissance ». Son utilisation au XX<sup>e</sup> siècle est renforcée par l'influence du terme anglo-saxon *cognitive*, relatif à la cognition, il est devenu l'étude de la connaissance. Il couvre l'ensemble des sciences qui ont partie liée à la connaissance et au processus d'apprentissage des connaissances. Et sont dites Sciences cognitives.

« La cognition est la faculté d'enregistrer, de trier des informations sur son environnement afin que cela devienne une connaissance. La psychologie, la linguistique, la cybernétique et l'informatique s'y intéressent», [Corroy, Gonnet, 2008, p. 36]. Comme elles intéressent le *rechercheur* d'information puisqu'il est confronté à un certain nombre de problématiques et pratiques qui ressortissent des sciences cognitives mais aussi des Sciences de l'information et de la communication (Sic).

*« Le terme d'information est lié, en psychologie, au courant théorique qui fait du psychisme humain (individuel) un système de traitement de l'information c'est à dire un système qui calcule à partir de symboles et de règles. L'information est sa matière première, codée sous forme propositionnelle. Si ce courant s'est caractérisé par son intérêt pour les modélisations informatiques, il a contribué à placer au centre des débats théoriques la notion de représentations mentales. C'est cette dernière et non l'information qui semble la notion centrale, même au sein de ce courant théorique. Les difficultés rencontrées pour modéliser diverses activités humaines ont entraîné une prise de distance avec la référence informatique et le traitement de l'information au profit des informations; celles-ci sont « le résultat des traitements des systèmes sensoriels. Ces informations sont de deux ordres: des informations de nature spatio-temporelle concernant les objets et les événements et des informations de nature symbolique (langagière ou iconique) qui véhiculent des significations et sont à interpréter à l'intérieur de systèmes de signes et du contexte de la situation. » (RICHARD, 1990, p. 12) [Nasse-Kolmayer, 1997, p. 169].*

#### 1.1.1.1. De l'information psychologique à l'information physique

Le cerveau humain est le lieu de naissance de toute connaissance humaine, le lieu conservatoire et provisoire de la connaissance, sa bibliothèque hypertextuelle mémorielle et neurobiologique. C'est par le truchement de ses sens et l'exercice de ses facultés et capacités intellectuelles que l'homme est porté à connaître, à s'informer et à se former. « L'Homme est plongé dans un monde qu'il perçoit par l'intermédiaire de ses organes des sens. L'information qu'il reçoit est coordonnée par son cerveau et son système nerveux (...), [Norbert Wiener, In Bougnoux, 1993, p, 446].

#### 1.1.1.2. Architecture du cerveau humain : éléments d'information

« Nous vivons à l'ère d'une neuroscience, qui vise à étudier *l'organisation du cerveau en tant qu'objet physique*. [Changeux, 2004, p.17].

On sait que le cerveau humain est structuré en trois parties qui correspondent aux étapes historiques de son développement : le cerveau reptilien, chargé du contrôle de la survie, des besoins fondamentaux et des automatismes ; le cerveau limbique, centre de la vie émotionnelle ; enfin le cerveau supérieur ou néocortex, centre physiologique de la vie rationnelle. L'essentiel de ses informations proviennent par le truchement de nos cinq sens. Ces trois cerveaux

sont interconnectés et fonctionnent avec les deux hémisphères. L'hémisphère gauche est analytique et l'hémisphère droit analogique. Tous nos apprentissages dépendent du bon fonctionnement et des interactions entre toutes ces entités cérébrales. Le tissu nerveux est formé de neurones au nombre de plusieurs milliards organisés en circuits locaux, « lesquels constituent à leur tour des régions corticales, (...), ou des noyaux. Les régions corticales et les noyaux sont interconnectés pour former des systèmes et des systèmes de systèmes à des niveaux de complexité de plus en plus élevé », [Damasio, 2008, p. 50]. Le cerveau est ainsi un véritable système d'information. L'information chimique et *électrique*, (influx nerveux) y circule au moyen des neurotransmetteurs.

Certaines informations se prêtent davantage à un traitement analytique, cas d'un message verbal, et d'autres à un traitement synthétique, cas de la reconnaissance d'un visage.

Le cerveau est le siège de la mémoire. La définition de la mémoire est influencée par la métaphore informatique qui est décrite comme « un processus de mémorisation en trois étapes (encodage/stockage/récupération) admis dans la plupart des théories (...), [Dennecker, Kolmayer, 2006, p. 45], ce qui induit que le cerveau est aussi un système d'information.

#### **1.1.1.3. La mémoire humaine comme contenant informationnel**

Les psychologues cognitivistes distinguent la mémoire à court terme, dont la capacité est limitée, qui permet le stockage d'information, et la mémoire à long terme, aux capacités *illimités*. « Souvenirs et connaissances y sont remémorés selon un processus complexe qui s'appuie sur la catégorisation et l'association ». [Dennecker, Kolmayer, 2006, p. 55].

Sans la mémoire il n'y a pas de connaissance possible et « (...) l'homme possède cette propriété unique de placer sa mémoire en dehors de lui-même, dans l'organisation sociale », [Leroi-gourhan, cité par Debray, 2000, p. 53].

#### **1.1.1.4. Connaissance procédurale\* et connaissance déclarative\***

Les sciences cognitives, anticipant les distinctions du *Knowledge management* entre savoirs et savoir faire, appellent connaissances procédurales, les connaissances liées à l'action, et connaissances déclaratives, les connaissances sur le monde. (Faute de place, nous ne ferons que de brefs rappels aux connaissances procédurales).

#### **1.1.1.5. La connaissance : connaître et comment connaît-on ?**

Connaître, revient sur le plan théorique à distinguer le sujet qui connaît et l'objet à connaître puis à se demander par quels moyens la réalité est connue par le sujet connaissant ? Le sujet a essentiellement affaire avec ses représentations, (concepts) qui sont autant de *signes* pour interpréter le monde réel. Comment ce qui est *en-soi* demandera Hegel, par exemple, peut-il devenir *pour-moi* ? Sortir de soi, s'extraire de son monde, pratiquer un *Hors soi*, [Baltz], est un autre plausible moyen pour tenter de connaître, de répondre à la question de la connaissance. Toutefois, sur un plan pratique, le raisonnement est un des moyens pour parvenir à la connaissance. Mais la psychologie cognitive postule que l'homme ne raisonne pas toujours selon les canons de la logique.

Dans l'acte de connaître, il y a toujours une part d'incertitude. «Le raisonnement humain est une activité en contexte.», [Dennecker, Kolmayer, 2006, p. 155]. On connaît de plusieurs manières par essai et erreur (tâtonnements), par déduction logique, par hasard, (sérendipité), par intuition, par acquisition de connaissances, etc. ....Ce que l'on cherche à connaître est encore et toujours du domaine de l'information. Connaissance et information ont indissolublement partie liée. « L'information est une connaissance inscrite (enregistrée) sous forme écrite (imprimée ou numérisée) orale ou audiovisuelle ». [Le Coadic, 1994, p.6].

L'information structure la société selon certains auteurs. Ce qui est acquis, c'est que l'information fait société, en ce sens qu'elle est créatrice de lien social, et qu'elle accompagne les technologies de l'information dans leur *brusque* mouvement d'accélération des connaissances.

## **1.2. Information : bref aperçu sur un terme devenu un paradigme contemporain**

La révolution informationnelle a bouleversé l'ordre historique et immuable de nos sociétés et est devenue, même s'il est contesté, un nouveau paradigme. Trois cycles longs peuvent être repérés: celui de « la matière qui nourrit l'homme (révolution agricole), [celui de] l'énergie et de la production, (révolution industrielle), et [celui] de l'information, [Ghernaouti-Hélie, Arnaud, 2001, p.94].

### **1.2.1. Qu'est ce que l'information ?**

« L'information est un concept dont la portée n'a d'égale que la difficulté qu'il y a à l'évoquer en termes généraux. », [Dion, 1997, p.9].

Ce vocable est particulièrement difficile à cerner. Il est surabondant dans notre espace contemporain, « autoroute de l'information », « Société de l'information », qui semble en user et en abuser puisqu'il connaît des dérivés nombreux : « sur- information », « désinformation », « contre-information ».

C'est par le concept de « société de l'information » que le terme continue sa proliférante diffusion polysémique au cœur de nos sociétés qui organisent, depuis quelques années, des « sommets mondiaux de l'information », (Genève 2003, Tunis, 2005). L'ONU a même décrété que le 17 mai de chaque année verrait la « Journée mondiale de l'information », la première ayant eu lieu le 17 mai 2007.

Issue de l'ancien français *enformer*, au sens étymologique de donner une forme, l'information est la mise en forme d'une connaissance en vue de la communiquer. C'est de cette mise en forme cognitive et informationnelle (et redondante) qu'il est question ici. L'information a encore des acceptions biologiques et médiatiques.

### **1.2. 2. Sens biologiques**

« Le schéma de Shannon a servi à mieux comprendre un certain nombre de phénomènes biologiques. La notion de code biologique lui doit beaucoup », [Breton, Proulx, 2006, p.131].

Dans les systèmes vivants, l'information génétique est contenue dans les structures de l'ADN et de l'ARN (messager). L'information cérébrale est communiquée à la fois par des signaux électriques et des substances chimiques, les neurotransmetteurs. « Le cerveau se comporte naturellement comme un système autonome qui projette en permanence de l'information en direction du monde extérieur (...). [Changeux, 2004, p.37].

L'analogie et le rapprochement entre information numérique et information neurobiologique n'ont rien de fantaisiste. La capacité des micros processeurs double, selon la loi de Moore, tous les 18 mois. Moore a prédit que cette capacité arrivera à saturation physique en 2017, sauf... « Si l'on parvient à un changement technologique radical, par exemple les mémoires informatiques « chimiques » (les informations binaires sont représentées par des molécules), voire biologiques (les informations tiennent dans les cellules) », [Jacquesson, Rivier, 2005, p. 33].

Dès lors qu'elle n'est plus n'est uniquement et toujours objectivement mesurable, l'information devient subjective et psychologique.

### 1.2.3. Sens médiatique

Le vocable information est généralement compris dans un sens journalistique et médiatique, comme celui d'une nouvelle communiquée par un média, (radio, blog, presse, télévision).

En ce sens, le terme information ne caractérise pas un ensemble de faits objectifs, car l'information est *fabriquée* avant d'être diffusée et médiatisée par un canal et communiquée à son destinataire final public.

Les journalistes reçoivent, sélectionnent et traitent leurs informations. Les événements retenus sont souvent agencés et mis en scène, les interviews et les images découpées, montées, commentées, censurées. Souvent, cette information fabriquée est proche des pouvoirs et participe du *consentement*, de la fabrication de l'opinion, (Chomsky), (Halimi). Elle s'apparente alors à la propagande. Pour certains chercheurs, nous l'avons mentionné, la prégnance de l'information est telle qu'elle structure le social. (A cet égard, l'information a à voir avec la communication). L'information est omniprésente dans la société. Cette surabondance informationnelle fait dire à Jean-Louis Ermine : « (...) que ce désordre informationnel concrétise la notion d'entropie dans la théorie de Shannon », [Ermine, 2000, p.43]. On se souvient que l'entropie désigne en thermodynamique le degré de désordre d'un système physique, mais devient en « information, le degré d'incertitude, mesuré en *bits*<sup>5</sup> ou unités binaires, qui s'attache à une source émettant des messages », [Bougnoux, 1993, p.407]. Norbert Wiener ira jusqu'à écrire : « (...) plus le message est probable, moins il fournit d'information. Les clichés ou les lieux communs, par exemple, éclairent moins que les grands poèmes, [Norbert Wiener, cité par Bougnoux, 1993, p. 448].

### 1.2. 4. Qu'est-ce qu'une information en documentation : essai de définition

Donnée et Information. Il faut distinguer nettement information et données. Les données (*data*), ne sont pas *stricto sensu* de l'information. L'information étant le résultat d'une transformation ou d'un traitement imposé à des données afin de fournir des réponses spécifiques à des questions spécifiques. Ou, pour le dire

---

<sup>5</sup> Bit ou (*Binary digit*) : Unité élémentaire d'information représentée par 0 ou 1.

autrement, l'information est une donnée numérique qui *dort*, (n'a pas encore été extraite) dans une base de données, par exemple. Information, données et connaissance : ces termes sont fortement polysémiques. Parfois synonymes, ils sont opposés par les spécialistes de disciplines différentes...et les dictionnaires usuels et spécialisés. Une recension et une comparaison de tous ces mots et de leurs acceptions dépassent ce travail. Pour résumer, l'information a le sens courant de nouvelles (*média*), le sens spécialisé de données électroniques, (*data*) et le sens de savoir, (*Knowledge*) et/ou, celui de connaissance.

Enfin, pour en terminer provisoirement avec « le chantier de la définition de l'information », [Baltz, 2003, p. 363], dans le cadre de travaux réflexifs et théoriques de l'enseignement d'une cyberculture à l'Université Paris 8, l'information « se définit comme in-formation, co-relation observée de variations entre un HT<sup>6</sup> émetteur et un HT récepteur », [Baltz, 2003, p. 363].

L'information, on vient de le voir, même brièvement, est partout et protéiforme. Il n'est pas exagéré non plus de dire et de penser que tout est information. La capacité circulatoire de l'information, sa ductilité infinie, multipliable qu'elle est en une infinité de variations hypertextuelles, à la fois vecteurs d'information et informations, font que l'information ne cesse pas d'être... *physiquement* dans le monde et que celle-ci a aussi parfois un prix.

### 1.2.5. La valeur de l'information

L'information ne possède pas *a priori* de valeur. Elle commence à prendre de la valeur lorsqu'elle devient indispensable à quelqu'un. Mais qui détient l'information, dit-on, détient le pouvoir, que ce soit par rétention d'information, ou bien lors de la circulation de l'information.

Une caractéristique insolite fait de l'information le seul bien que son détenteur puisse échanger sans en être dépourvu à son tour. Elle ne s'épuise pas lorsqu'on la consomme. (On doit, selon Didier Frochot<sup>7</sup>, cette remarque à Thomas Jefferson, (1743-1826).

L'information est porteuse de sens, elle est le vecteur de la connaissance : c'est peut-être là sa seule valeur.

#### 1.2.5.1. Connaissance et information

Selon la Bible, la connaissance naît d'une faute, d'une insoumission, [Genèse, II, 17], et parallèlement, pour les Grecs, la connaissance commence par une action délictueuse: « Alors Prométhée, ne sachant qu'imaginer pour donner à l'homme le moyen de se conserver vole à Héphaïstos et à Athéna la connaissance des arts avec le feu ; car, sans le feu, la connaissance des arts était impossible et inutile ; et il en fait présent à l'homme » [Platon, *Protagoras*, 321c]. La connaissance débute par un vol, un rapt lumineux. Rappelons en outre, le propos de l'épistémologiste : rien n'est donné. C'est qu'il faut s'attacher à dérober au réel ce qu'il cèle au regard des hommes pour connaître et dissiper ses territoires d'ombre. Le feu semble être l'une des premières technologies et l'un des premiers symboles éclairants de la

<sup>6</sup> HT signifiant hypertexte.

<sup>7</sup> Frochot, Didier [en ligne] [réf. du 10/07/08 sur Infostratégès]. Disponible sur le world wide web <<http://www.les-infostrateges.com/article/031275/aspect-multiforme-de-l-information>>.

connaissance acquis par l'homme. Pour Héraclite, (vers 500 av J.C.- 480 ? A.J.C.), le feu sera un principe explicatif fondamental.

### **1. 3. La société de l'information : histoire et concepts**

« Le terme « société de l'information » (*information society*) est un néologisme assez vague. Il désigne en fait la préoccupation des dirigeants politiques d'offrir des conditions favorables aux nouvelles technologies de l'information, considérées comme un enjeu capital pour leurs pays, et accessoirement pour le développement dans le monde entier ». [Rivier, 2007, p. 12].

Pour Miège, l'expression de « société de l'information » « est toujours aussi inconcevable », [Miège, 2004, p.56]. C'est également pour l'ancien professeur des Sciences de l'information de l'université de Grenoble, une butée définitionnelle. Car ce syntagme polysémique ne fait pas consensus en Sciences de l'information et attendra peut-être encore longtemps sa définition précise, d'autant plus que le débat oppose également les tenants de l'information et ceux de la communication. (Cf. D. Bougnoux, *La communication contre l'information*, Hachette, 1995).

L'explosion de la communication a lieu en pleine Seconde guerre mondiale avec les cybernéticiens, N. Wiener (1894-1964), Claude Shannon. « L'emploi du terme communication par la cybernétique en fait un quasi équivalent de la notion d'information (...) », [Breton, Proulx, 2006, p.125]. Cette polysémie sémantique ne manquera pas d'ajouter de la confusion dans la perception et le traitement de l'information.

Il est cependant possible de distinguer ce qui ressortit à l'information mesurable et ce qui ressortit à l'information psychologique.

#### **1.3.1. La théorie mathématique de l'information de Claude Shannon [1948] : la quantité d'information mesurable**

Ingénieur aux laboratoires Bell, (Etats-Unis) Shannon construit sa théorie mathématique de l'information de 1943 à 1945, qu'il publie en une reprise de trois articles antérieurs réunis sous le titre de « *Théorie mathématique de l'information* » en 1948. Ses travaux portent sur la capacité d'un canal à transmettre des informations, (signal électrique) et au codage des messages. Intervient une notion statistique, probabiliste : la quantité d'information. Ce dernier terme n'a aucun rapport avec l'usage courant d'information. En outre, cette information, et ce point est capital, ne fait pas intervenir le sens dans le message.

« La quantité d'information est une mesure de la liberté de choix entre plusieurs messages. Elle est mesurée par le logarithme d'un nombre de choix possibles. Elle possède une unité, le bit » (*binary digit*), [Leclerc-Reynaud, 2006, p. 28].

Dévolue au domaine des télécommunications, la théorie mathématique de la communication, (autre appellation de la théorie de Shannon), « s'applique au problème technique de la fidélité de la transmission de types variés de signaux d'un émetteur à un récepteur ». [Weaver, cité par Leclerc-Reynaud, 2006, p. 29].

Ainsi, la communication n'est rien d'autre que la science du transport des messages, (les données) de machine à machine.

Nous venons de le rappeler, l'information dans le modèle shannonien est dénuée de sens et de signification, mais qu'est-ce qu'une information qui répond à un besoin d'information d'un utilisateur?

##### **1.3.1.1. Qu'est-ce qu'une information qui répond à un besoin d'information?**

Une information peut s'entendre ici comme une donnée, une référence qui fait sens avec et pour son utilisateur. L'information n'existe pas pour elle-même. Elle est cet objectif à atteindre et en même temps cette possible finalité. Des moyens mis en œuvre par un médiateur pour atteindre cette donnée informative découle la réussite ou non des stratégies de résultats. On le verra par la suite, l'information, qui est au cœur des sociétés humaines, ne prend sens que dans les actes et les procès de communication. L'information est liée au besoin de savoir et à la connaissance : c'est de cette information-là dont nous parlons.

### **1.3.1.2. L'information : un besoin pour un utilisateur de système d'information**

De quelle information s'agit-il ? Car ce concept a beaucoup fluctué ces dernières années. Jadis, (avant les années 70), reconnu sous l'appellation « d'information scientifique et technique », et naguère, (jusqu'au début des années 90), connu par l'expression d'«information spécialisée\*», il semble que le terme d'information, sous la pression peut-être, de l'expression très connotée, de « société de l'information », connaisse un regain d'intérêt. En effet, l'information est devenue omniprésente dans notre univers contemporain, c'est une valeur fondatrice puisqu'elle nous projette dans un au-delà de la société post industrielle. Les professionnels de l'information savent que les flux informationnels dont ils ont besoin gisent dans des sources d'information.

### **1.3.1.3. Les sources d'information**

La connaissance et la gestion (mise au jour et mise à jour) des sources d'information est une préoccupation majeure de la profession. Connaître les sources, les atteindre et savoir y déambuler, sans s'y noyer, est indispensable. Nous ne pourrions pas traiter, hélas faute de place ici, de l'ensemble des problématiques d'échec documentaire induites par les sources d'information, (connaissance et repérage des sources, accès navigationnel, structure et organisation, etc., ...). Nous nous limitons à en effleurer les domaines. A titre d'exemple, ce sont, pour les principales d'entre elles : les sources bibliographiques : *Livres hebdo*, *Electre*, *Books in print*, etc.,...Les catalogues en ligne des grandes bibliothèques, *Bn opale plus*, le catalogue collectif de France ([www.Ccfr.bnf.fr](http://www.Ccfr.bnf.fr)), le Sudoc, (<http://sudoc.abes.fr>), le Kvk : [Http://ubka.uni-karlsruhe.de/hylib/virtueller-katalog.html](http://ubka.uni-karlsruhe.de/hylib/virtueller-katalog.html) , etc.,...

Les organismes spécialisés, les banques d'images, les brevets, les agrégateurs de presse et guichets de l'information tels : *Pressed* (presse francophone, *Factiva*, (presse internationale), l'Inist, (pour l'information scientifique et technique), etc..... (Pour une solide base d'information, on consultera avec profit : [PrévotEAU, Utard, 2005] et pour les sources d'information, [Lawson-Mathieu, Barsanti, 2006], et *Oriente express* de la Bpi en ligne sur : <http://www.bpi.fr>.

Mais l'expression « des besoins en information est à la base de toutes les études et recherche sur la conception des systèmes d'information », [Mathelot, 1998, p. 53].

## **1.4. Le système d'information**

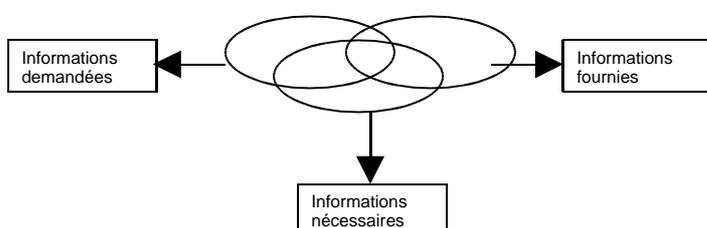
Issue de l'approche systémique, reprise par la suite par les cognitivistes, la notion de « système d'information inclut la fonction documentaire, sous système de l'entreprise, (...), [Accart, Réthy, 2003, p. 82]. Un système d'information recueille, traite et diffuse l'information. On a vu plus haut que le cerveau pouvait être considéré comme un système d'information. On désignera également sous cette appellation le cœur informationnel et stratégique du centre de documentation, les bibliothèques, les bases de données, les outils d'interrogation et de conception, et les relations qu'ils entretiennent nécessairement entre eux et avec les personnels. Nous ne retiendrons pas l'appellation de « Système de recherche d'information », (SRI), [Mkadmi, Saleh, 2008, p. 18] à cause de son sens essentiellement technique et spécialisé, (dénoté) des chercheurs en sciences de l'information.

#### **1.4.1. La place de la recherche d'information dans le système d'information**

L'information et sa recherche sont au cœur des dispositifs des systèmes d'information. Il s'agit de retrouver, par différents procédés et tactiques exploratoires, (essais et erreurs) des éléments d'information disponibles ou bien à demeure, (documentation interne) ou bien d'aller repérer et découvrir ces éléments informationnels distants du lieu d'exercice professionnel, (documentation externe). On le devine, il sera souvent plus difficile d'agir et de connaître leur contenu informationnel, par exemple, sur des documents repérés extérieurement, contrairement à la possibilité, plus aisément maîtrisable, d'agir sur des documents effectivement possédés par le système d'information. Une recherche dans les rayons d'une bibliothèque donne, par exemple, la possibilité d'une réponse immédiate.

##### **1.4.1.1. Les spécificités de l'information dans le contexte de la documentation**

Point capital, dans le cadre de la recherche d'information, il n'est évidemment pas possible d'être complètement informé sur tous les sujets et de tout connaître.



La reconnaissance des informations [Mathelot, 1998, p 52]

En suivant le schéma proposé ci-dessus, « La reconnaissance des informations » et les explications données par [Mathelot, 1998, p. 52-53], on appréhende mieux le hiatus triadique qui opère au sein des systèmes d'information, lorsque l'on étudie les étapes de la restitution d'informations à l'utilisateur. Face à la résolution d'un problème d'information,- et c'est déjà anticiper sur un des entrelacs d'une problématique de ce mémoire,- informations demandées, informations fournies et informations nécessaires, ne se recoupent pas.

#### **1.4.1.2. L'approche orientée usager**

Tirant l'expression orientée usager de l'informatique, (modèle orienté objet), Le Coadic définit ce modèle comme un nouveau paradigme qui replace l'utilisateur des services d'information et son besoin d'information au centre du dispositif des systèmes d'information, [Le Coadic, 1997, p.13]. Nous nous inspirons de ce modèle, dans lequel l'utilisateur devient le partenaire principal, non pas tant à cause de son caractère éventuellement novateur, mais à cause de son évidence principielle, et quand bien même « la figure difficilement identifiable et toujours discutable « de l'utilisateur » [idéal serait à nouveau] remise en cause » en science de l'information, [Favier, Martin-Juchat, 2002, p. 255]. Quoi qu'il en soit de ce débat, à l'utilisateur, reviendra toujours le dernier mot, (*pertinence*) quand bien même il se tromperait.

#### **1.4.1.3. L'usabilité de l'information**

L'approche orientée usager définit un nouveau concept, « l'usabilité ». L'usabilité, qui précède toujours l'utilité, mesure jusqu'à quel point un produit d'information, un système d'information, un service d'information, une information est prête à l'usage. C'est un concept clé à côté des concepts d'utilité et d'efficacité. [Le Coadic, 1997, p. 52].

Produits et services d'information, à cause de la complexité des fonctionnalités des nouveaux matériels et des versions successives des logiciels, des interactions des techniques qu'ils recèlent, rendent leur apprentissage difficile et sont à l'origine de nombreuses erreurs. (Mémorisation des modes d'emploi délicats à saisir, ergonomie des systèmes peu adaptée..).

#### **1.4.1.4. L'information et sa représentation**

La solution d'un problème d'information est une opération complexe, conditionnée par les représentations psychologiques et cognitives qu'en ont les utilisateurs. Selon Claude Morizio, « Dans le cadre de la recherche documentaire, la difficulté est de redéfinir le problème en fonction des représentations », [Morizio, 2002, p. 70]. L'étymologie nous indique que ce terme vient du latin *repraesentare*, qui signifie « rendre présent, reproduire ». La représentation s'entend également comme au théâtre, et dans le commerce, le représentant est celui qui montre... « Dans le domaine des sciences de l'information et de la communication « représentation » est le plus souvent utilisé en rapport avec l'étymologie latine : présenter à nouveau, rendre présent ». [Corroy, Gonnet, 2008, p. 291-292].

La représentation d'un énoncé intellectuel, (ici, une question) est déjà une stratégie cognitive, en réponse à un problème. En effet, en sciences cognitives, l'acte de connaissance passe par « une représentation de la situation » et une série de questions « qui aident à créer une image plus ou moins complète de la situation dans laquelle l'apprenant se trouve, du problème qu'il veut régler » [Morizio, 2002, p.74]. Autrement

dit, une représentation est un acte par lequel un objet *pensé* devient présent à l'esprit, et prend en même temps la place de l'objet perçu dans l'intellect du sujet regardant et se substitue à l'objet perçu. Pour nous, une représentation est également une réponse, dans la mesure où cette réponse est une construction du sujet connaissant, (percevant).

#### 1.4.1.5. Emetteur et récepteur : le message informationnel

L'objectif pour les professionnels de l'information est de permettre la médiation entre des ressources informationnelles et un demandeur, voire un utilisateur d'information, qu'on appelle usager.

Dans l'acte de communication, le message est ce dont il est question. « (...) l'objectif du message est formulé par un émetteur, (E), (...), et transmis à travers un canal vers un récepteur-destinataire (R.). [Lazar, 2004, p.147]. Ce récepteur est en l'occurrence tour à tour le *chercheur* d'information et l'utilisateur des services du système d'information. Sous ce terme d'utilisateur, se cachent aussi bien l'utilisateur, le visiteur de bibliothèque que son lecteur et évidemment, le client, dernier avatar de la société mercantile.

L'acte de communication qui unit plus particulièrement les acteurs de ce processus interactif et cognitif est le dialogue.

### 1.5. Le dialogue avec l'utilisateur

« On peut comprendre le mot « dialogue » dans un sens élargi, c'est-à-dire non seulement comme l'échange à haute voix et impliquant des individus placés face à face, mais tout échange verbal, de quelque type que ce soit. Le livre, c'est à dire l'acte de parole imprimée, constitue également un élément de l'échange verbal » [Bakhtine, 1977, p. 136]

Généralement, une recherche d'information menée par un professionnel, (qui est aussi un usager) suppose un échange verbal, un dialogue avec un utilisateur en vue d'obtenir une réponse satisfaisante. Il leur est indispensable de préciser sous la forme de questionnement, les besoins informationnels, (l'utilisateur ne les ayant pas toujours explicitement formulés). De définir les objectifs et l'utilisation de la recherche. D'établir ce que chacun sait de la recherche, de poser des limites, le cas échéant à la demande, enfin de préciser les critères de recherche.

#### 1.5.1. La précision des critères de recherche

Ils visent à définir la forme du produit : bibliographie, dossier documentaire, dossier produit, fichier numérique etc.....

Les types de supports souhaités : papier (monographies, articles, rapports,..), ou numérique, la période couverte par la recherche, les langues du ou des documents, etc..... Le moment où les résultats doivent être communiqués...

##### 1.5.1.1. Identifier l'utilisateur

Pour répondre à une question, donc à un besoin d'information, le professionnel de l'information doit savoir **identifier son interlocuteur**.

Le questionnement, (anamnèse) de l'utilisateur est un des préalables de la recherche d'information. Identifier un utilisateur permet de mieux préciser son contexte informationnel, et d'approcher la *cible* informationnelle à atteindre.

« La manière d'interroger n'est pas neutre. L'efficacité d'une communication est subordonnée à une évaluation adéquate des connaissances de l'interlocuteur. » [Sutter, 2002, p.199-200].

« L'informateur [le professionnel de l'information], doit savoir situer son interlocuteur afin de pouvoir déterminer le niveau de connaissance qu'il faut considérer comme acquis pour émettre sa réponse, (...) », [Sutter, 2002, p.200].

Cette quasi *anamnèse*, pour être complète, doit rechercher, le non-dit, les motivations implicites, sous-jacentes, masquées par la requête initiale souvent formulée de façon hâtive. En outre le professionnel doit **pouvoir définir son public** : Les utilisateurs appartiennent-ils à la structure documentaire ou en sont-ils extérieur ? Sont-ils des spécialistes, des étudiants, des universitaires, des professionnels ou font-ils partie du « grand public » etc., ?

### 1.5.1.2. Le *rechercheur d'information* entre « cyber et culture »

Cyberculture entendue comme déambulation *navigationnelle*, navigation *éclairée* dans l'espace des savoirs et des connaissances structurées ou non. La boussole ou le sextant ou encore la lampe torche à la main, en l'occurrence machines de vision éclairantes, au sens d'objets de médiation de la *visio-connaissance*<sup>8</sup>, pour se repérer, se conduire, (cyber, du grec *kubernetes*, pilote, gouvernail), dans le corpus épars informationnel, qui va du médium, dans le plus simple appareil, (l'écriture, le signe), à l'hypermédia.

Le professionnel est entouré d'ouvrages dits de références et d'outils d'*assurance*, (dictionnaires spécialisés, encyclopédies, annuaires, traités, manuels bibliographies, bases de données, catalogues, signets, Opacs\*, etc.), qui lui permettent d'opérer en première instance.

Peut-être parfois à la manière étrange et provocatrice du bibliothécaire de *L'homme sans qualités* de Musil, pour qui « Le secret de tout bon bibliothécaire est de ne jamais lire, de toute la littérature qui lui est confiée, que les titres et la table des matières. « Celui qui met le nez dans le contenu est perdu pour la bibliothèque ! m'apprit-il. (...) Ainsi vous ne lisez jamais un seul de ces livres ?- Jamais. A l'exception des catalogues, [Musil, 1995, p. 581].

### 1.5.1.3. L'information et son médiateur dans le processus de recherche d'information

Le médiateur est *rendu* nécessaire dans les bibliothèques et centres de documentation ouverts sur le public, parce que c'est lui qui collecte, traite et communique l'information. En effet, les savoir-faire bibliothéconomiques visent également à maintenir en permanence une distance et une médiation professionnelles avec le visiteur en quête de besoin. Distance que les moteurs de recherche et d'indexation semblent abolir, comme ils abolissent *de facto* tous les discours sur la prétendue autonomie du lecteur et la démocratisation du savoir, tenus par la profession. Ainsi, le bibliothécaire explique rarement comment il procède pour trouver l'information, se réservant l'usage exclusif des technologies intellectuelles, qu'il s'approprie à son propre profit. La plupart des technologies utilisées en bibliothèque, avant l'internet, ont été conçues par des techniciens et des professionnels du livre et de l'information pour des professionnels du livre et de l'information. Le médiateur crée par là des liens de dépendance d'usage informationnel.

---

<sup>8</sup> Visio-connaissance : on est tenté d'écrire sur le modèle de visioconférence : Visio-connaissance.

Le médiateur peut prendre plusieurs formes, ce peut être une personne, un ouvrage ou une technologie, (interface). Le chercheur d'information, (documentaliste, bibliothécaire, néophyte etc.....) que l'on qualifiera ainsi définitivement maintenant, de *chercheur*, pour bien montrer qu'il opère en recherche (de) cette information, est également un, (son) médiateur.

Cette quête d'information, à tout le moins son processus, est une des difficultés inévitables auxquelles doit se confronter tout chercheur d'information qui en devient, simultanément et à son corps défendant, son médiateur. C'est lui, le médiateur, cet intermédiaire *nécessaire*, qui est au cœur du processus d'interrogation. De lui, dépend en grande partie, le succès ou non de sa (re)-quête d'information. Cette quête que commande une stratégie plus ou moins consciente, plus ou moins élaborée.

#### **1.5.1.4. Le *chercheur* d'information est toujours un *Apprenant***

Pour rechercher, ne faut-il pas déjà connaître ? Comment en effet, celui qui ne saurait rien pourrait-il chercher ce qu'il ignore ? Comment chercher ce que l'on ignore ? Ne faut-il pas avoir au préalable une certaine idée de ce que l'on (re)-cherche ? Platon développe ces paradoxes dans *Ménon* et plus précisément par la doctrine de la *Réminiscence*, [Platon, 1991, p 153 et sqq.]. On se souvient que la solution socratique de la *Réminiscence* considère que l'âme est immortelle et a acquis antérieurement la connaissance de toutes choses. La connaissance est ici associée au souvenir qu'il s'agit de redécouvrir en soi pour recouvrer cette connaissance *perdue*. Le *chercheur* est comparable à celui qui recherche une information pour un utilisateur, il ne sait pas ce qu'il va trouver, ni s'il va trouver : il sait seulement qu'il *doit* trouver (se souvenir), de quelque chose...Il fait en quelque sorte l'apprentissage toujours recommencé de son ignorance.

Saint Augustin<sup>9</sup> : « Qu'est-ce donc que le temps ? Si personne ne me le demande, je le sais, mais si on me le demande et que je veuille l'expliquer, je ne le sais plus ».

Les savoir-faire du *knowledge management* ne sont pas une garantie de réussite. Savoir comment chercher et où se trouve l'information demandée par un utilisateur n'entraîne pas de résultats positifs automatiques. Dans un monde en pleine évolution technologique accélérée, et c'est une banalité que nous réitérons, le professionnel de l'information est condamné à apprendre sans cesse, à se former continuellement faute de quoi, il risque au mieux de se répéter et au pis de périr.

#### **1.5.1.5. Le *chercheur* d'information entre savoir et vérité**

L'objectif de ce paragraphe est de montrer que le *chercheur* d'information est confronté à de nouvelles problématiques concernant les concepts de vérité et de pertinence.

En effet, Pierre Lévy l'avait noté à propos du savoir informatique et de l'information : nos sociétés de l'efficacité ne visent plus la vérité comme fondation des savoirs, mais privilégient la « pertinence et la vitesse de l'exécution ». [Lévy, 1993, p.

---

<sup>9</sup> *Confessions*, Livre XI

135]. Nous consacrons un développement important au concept de pertinence dans la deuxième partie de ce travail. « La vérité n'est jamais immanente mais sociale et disputationnelle », [Bougnoux, 1993, p. 571]. La vérité est presque toujours une hypothèse qui ne *demande* qu'à être invalidée.

Popper explique qu'une vérité n'est jamais définitivement assurée. La **multiplication des expériences ne constituant pas une preuve**. L'exemple poppérien suivant est relativement connu. Ce n'est pas parce que je viens de rencontrer un millième cygne blanc que je démontre que tous les cygnes sont blancs. En fait, j'aurais plus de certitudes en rencontrant un cygne noir. Alors, je pourrais démontrer qu'ils ne sont pas tous blancs. J'aurais réfuté sur un seul cas une conjecture fautive. Finalement, je peux prouver une erreur tandis que je ne peux pas établir définitivement une vérité.

Pour Gilles Deleuze, il n'y a vérité que créée, [Zourabichvili, 2004, 59]. La vérité, points de vue pluriels : plusieurs personnes décident ensemble, vrai et faux en même temps. La vérité est affaire de représentations du sujet observant (connaissant) et il n'y pas de connaissance sans le truchement de signes pour interpréter et connaître le réel. Ce sont là, brièvement abordées des difficultés épistémiques soulevées lors de la restitution d'information, (réponses).

**La communication authentique renvoie à la médiation. Espace et vérité. [C. Baltz en cours de Cyberculture].**

#### 1.5.1.6. L'erreur comme source de progrès

On peut considérer que l'erreur, donc l'échec documentaire, est une **réponse** non conforme au résultat attendu et présentée comme *vraie* ou pertinente, (la meilleure plausible : de *bonne foi*). La représentation de l'erreur relève d'une appréciation de son adéquation avec la vérité. Dans son acception positive, (non moralisante), l'erreur correspond plutôt au chemin incertain que l'esprit emprunte pour trouver des connaissances. Cette conception cognitiviste et positive de l'erreur est gage des stratégies de réussite. L'erreur est un fondement pédagogique positif parce que « les erreurs ne sont connaissables qu'après coup, une fois que les obstacles ont pu être franchis » [Astolfi, 2008, p.43].

Une pédagogie de l'erreur ? L'erreur comme source de progrès : Je me trompe, (je n'obtiens pas *ma* réponse) : je sais que je me suis trompé, donc je sais qu'il y a erreur. C'est ainsi, par *jeux* d'essais et d'erreurs que je parviens à obtenir une réponse, qu'il me restera à valider auprès du demandeur par action de médiation.

### 1.6. Médiation et médiations

Médiation vient du mot latin *mediatio* (de *mediare*), qui signifie « s'interposer » ou « diviser », ou de *medius* – « milieu »-, la médiation se définit en effet dans son sens moderne comme l'action de mettre en relation deux termes ou deux êtres. Il s'agit là de l'hypothèse de toutes les théories de la connaissance, si l'on excepte peut-être la phénoménologie, selon lesquelles le réel n'est pas connaissable *immédiatement* « Toute connaissance passe donc nécessairement par des médiations, dotées d'un plus ou moins grand pouvoir de *représentation*, c'est à dire de restitution ou de reconstruction du réel, [Briant, Palau, 1999, p. 41].

#### 1.6.1. Le *rechercheur* d'information comme médiateur

En situation de médiation : deux *postures*. Une posture cognitive face à la problématique de la question, (son contenu), l'autre interrogative, face à l'utilisateur. Le but : faciliter la compréhension et la résolution de la question.

#### 1.6.1.1. L'entretien de référence

Dans un article du *Bulletin des bibliothèques de France* qui a fait date, Jean-Philippe Lamy, [Lamy, 1998, p. 52-58], après avoir constaté le retard pris par la bibliothéconomie française dans les cursus universitaires, en matière de renseignement de l'utilisateur, proposait d'intégrer des exercices pratiques, qualifiés de « mise en situation », l'enseignement devant permettre aux futurs professionnels de mieux appréhender les réalités lors des entretiens (face à face), avec les usagers. Il insistait sur les points suivants : « – la nature de l'information recherchée ; par exemple, une question sur le cheval porte-t-elle sur les aspects pratique, historique, vétérinaire ? – la quantité : un fait, un ouvrage un article, une référence ? – la destination finale : un exposé, un discours, le début d'une recherche ? – le niveau de spécialisation, qui découle des deux paramètres précédents ; – le temps dont dispose l'utilisateur et, éventuellement, l'investissement financier qu'il est prêt à consentir ; – directement liée à ce qui précède, la date limite de cette recherche. »<sup>10</sup>

#### 1.6.1.2. Le dialogue renseignant /renseigné

Calqués sur les syntagmes enseignants et enseignés, nous utilisons ici les termes de *renseignants*, (médiateurs) et de *renseignés*, (utilisateurs ou usagers), pour désigner ceux qui *subissent* alternativement les effets des interactions dialogiques dans les espaces informationnels. Effets de domination et de position : effets de bricolage de postures professionnelles, (rôles sociaux et *mise en scène de soi*).

*Rien n'est donné*, il faut donc apprendre à parler, qui est un *agir communicationnel*, en situation de renseignant, ce dernier considéré par le tout venant comme celui qui sait. Qui ne sait pas tout, mais qui doit savoir au moins, où tout est cédé. « La connaissance est de deux sortes : soit on connaît son sujet, soit on sait où trouver des informations dessus », [Johnson, cité in Line, 2004, p.41]. Absent, on vient de le voir des enseignements de la bibliothéconomie, l'apprentissage de la fonction de renseignement se fait empiriquement, par observations mimétiques des pratiques des collègues expérimentés. Le dialogue avec l'utilisateur revêt d'autant plus d'importance que les processus mis en œuvre peuvent induire des échecs. Il arrive que ce soit le demandeur qui ...donne finalement la réponse !

Le renseigné ne l'est, (renseigné) qu'au produit des informations fournies par la subjectivité et les savoir-faire professionnels du renseignant. Le renseigné est ainsi soumis à son corps défendant à une forme de hasard, (ou bruit ou silence) : les compétences du renseignant. Le dialogue est le canal métaphorique et interrelationnel par où transite le message.

Le dialogue, « conformément à son étymologie grecque désigne une forme de conversation entre deux ou plusieurs personnes. Le dialogue repose donc sur une alternance de tours de parole produit par des locuteurs différents », [Garric, Calas, 2007, p. 108].

---

<sup>10</sup> Lamy, Jean-Philippe, « L'entretien de référence : Une notion à introduire dans un enseignement rénové », *BBF bulletin des bibliothèques de France*, 1998, n° 6, p. 52-58 [en ligne] <<http://bbf.enssib.fr>> Consulté le 14 avril 2008.

### 1.6.1.3. Des qualités nécessaires du médiateur dans le processus de recherche

Eric Sutter, dans son ouvrage : *Documentation, information, connaissances : la gestion de la qualité* ADBS, 2002, p.72, citant Guy Le Boterf, énumère avec lui, quelques unes des compétences requises et indispensables aux professionnels de l'info-doc : « Connaissances générales (concepts, savoirs disciplinaires, etc....), elles servent à comprendre un phénomène, une situation, un problème, un procédé. Connaissances spécifiques à l'environnement professionnel : contexte de travail, règles de gestion, culture organisationnelle, codes sociaux, organisation de l'entreprise, etc. Elles permettent d'agir sur mesure. Connaissances procédurales : le « comment il faut faire » ou « comment s'y prendre pour ». Elles décrivent des méthodes, des modes opératoires. Savoir-faire opérationnels : démarches, méthodes, instruments dont la personne maîtrise l'application pratique. Elles (*sic*) permettent de savoir opérer. Connaissances et savoir-faire expérientiels : ils sont issus de l'expérience, de l'action (« tour demain », « astuces », etc.).

**Savoir faire relationnels** : capacités qui permettent de coopérer efficacement avec autrui ; capacité d'écoute, de négociation de travail en équipe, de travail en réseau, etc.

**Savoir-faire cognitifs** : opérations intellectuelles nécessaires à l'analyse et à la résolution de problèmes, à la conception et à la réalisation de projets, à la prise de décision, à l'invention. Elles permettent de créer des informations nouvelles à partir d'informations existantes. » (...).

A ce tableau, qui décrit en les listant, les qualités d'un professionnel de l'information quasi idéal, faut-il modestement ajouter : être capable de se mettre au niveau du demandeur, être persévérant et susceptible de se remettre en question. (Les échecs sont formateurs, si l'on sait en tirer profit et analyser pourquoi l'on s'est trompé). Enfin, savoir faire appel à d'autres professionnels en cas de difficulté, (carnet d'adresses) et orienter à bon escient l'utilisateur.

## 1.7. Trouver l'information demandée

Trouver ou ne pas trouver l'information, là où est la réponse, là aussi est **la question**, telle est la question. Quelle stratégie mettre en œuvre face à l'énigme de la recherche, qui est une tentative d'élucidation (essai-erreur) ?

Pour retrouver une information, c'est évident, il faut savoir chercher.

Ou bien ranger ! Ce sont là toutes les problématiques liées aux classements et aux classifications documentaires et aux aménagements des espaces documentaires (signalétique), que nous ne pourrons pas développer faute de place.

Il est indispensable de connaître les stratégies et les méthodes de recherche : comment chercher ? C'est-à-dire connaître les stratégies à employer, où chercher ? C'est-à-dire connaître les sources d'information, avoir une idée de ce que l'on cherche. C'est-à-dire connaître le domaine couvert. La sérendipité ou la rencontre fortuite, elle favorise la découverte aléatoire de ce qu'on ne cherchait pas. Mais chercher expose au risque de s'égarer.

### 1.7.1. Brève histoire de la recherche d'information

Ce bref aperçu, extrêmement rapide et succinct, de l'histoire de la recherche d'information n'a aucune prétention. Il rappelle simplement au lecteur que les obstacles à la recherche d'information s'insèrent dans l'histoire concomitante du développement des connaissances et des moyens d'y accéder. Histoire du livre et des bibliothèques, des catalogues et de la bibliographie, du classement et des classifications et des choix qui présidèrent à la création de ces outils.

Les technologies de recherche d'information ont évolué en plusieurs étapes. Des méthodes classiques de la recherche *papier*, puis les systèmes booléens et les systèmes statistiques pour arriver au traitement automatisé du langage (Tal), jusqu'à l'internet et à la recherche en texte intégral.

Le mot bibliographie vient du grec « *biblion* », livre et de « *graphein* », écrire. Ce terme est postérieur à l'objet qu'il désigne : le répertoire de titres imprimé qui date de 1494. Son origine manuscrite remonte au II<sup>e</sup> siècle, avec le *De Libris propriis* du médecin grec Gallien, dont le travail « donne la première expression de l'idée bibliographique qui équivaut à « liste d'œuvres » ». [Malclès, Lhéritier, 1989, p. 7]. Ces listes constituèrent les premiers répertoires bibliographiques bien qu'ils soient désignés sous les termes de « *Bibliotheca* », « *catalogus* », « *repertorium* », « *index* », etc. ... Grâce à leur élaboration devenue plus rigoureuse à partir des 18<sup>e</sup> et 19<sup>e</sup> siècles, ces répertoires bibliographiques furent parmi les premiers instruments fiables de recherche. Ils permirent la recherche bibliographique.

Longtemps, la logique du titre de l'œuvre et du nom, (du prénom, jusqu'au moyen âge) de l'auteur aura été privilégiée dans la constitution des catalogues de bibliothèques, Cf. [Mouren, Peignet, 2007, p. 17-36]. Il était donc nécessaire et indispensable de savoir **qui** avait écrit, et surtout de connaître le nom et le titre sous lequel le livre avait été publié, pour retrouver le document correspondant. Il semble (les auteurs divergent à ce propos), que c'est dans les années 1930, aux Etats-Unis, que les bibliothécaires ont élaboré des fichiers (catalogues) analytiques de matière, à l'intérieur desquels des fiches, classées alphabétiquement, indiquaient le contenu (**quoi**), de l'ouvrage. On était ainsi passé d'une logique du « qui ? » à une logique du « quoi ? » qui permettait de savoir ce que la bibliothèque possédait sur tel ou tel sujet.

Ce qui augmentait, en même temps les réponses mais aussi le bruit\*, puisque ces fichiers analytiques étaient construits sans langages contrôlés.

Le développement de l'informatique des années 1950 à aujourd'hui a permis la recherche plein texte, (*full text* en anglais), c'est-à-dire que le texte est considéré comme un seul champ,\* dans lequel tous les *mots*, (en fait des chaînes de caractères), font l'objet d'une entrée dans l'index. On verra plus loin que cette dernière possibilité n'est pas sans inconvénients

#### 1.7.1.1. Existe-t-il une science de la recherche d'information<sup>11</sup> ?

---

<sup>11</sup> « La méthodologie de la recherche scientifique s'appuie sur les écrits de C. Bernard et sur le schéma dit OHERIC (O= observation, H=hypothèse, E=expérience, R=résultats, I=interprétation, C= conclusion).

Le concept de science : « La science » nous dit Jean-Lévy-Leblond, « c'est l'art de transformer les questions jusqu'à ce qu'elles trouvent une réponse ». [Lévy-Leblond, 2003 p.108], et plus encore « Trouver l'information est un art, pas une science », [Lardy, 2001, p.3].

Il n'existe pas en France de science de la recherche d'information au sens strict. Intégrée au cours des développements successifs des évolutions de l'information de ces dernières années, (1950), dans les cursus des écoles de bibliothécaires : Ecole des Chartes, (1821) ; Institut national des techniques documentaires, Intd, (1950) ; Ecole nationale des sciences et des bibliothèques, (1963) Ensb, devenue en 1992, Enssib, Ecole nationale supérieure des sciences de l'information et des bibliothèques, et les modules des parcours universitaires des métiers du livre et de la documentation, (Dut, licence, maîtrise, Master), la formation à la recherche d'information repose essentiellement sur un savoir-faire professionnel. Mais cette affirmation doit être explicitée.

En effet, si un empirisme professionnel semble dominer, la formation à l'information s'est lentement développée en France dans les universités, qui ont mis en place dès 1982, des Unités Régionales de Formation à l'Information Scientifique et Technique, (URFIST), destinées à l'origine, aux étudiants de 3<sup>e</sup> cycle. Depuis 1997, la France s'est dotée d'un portail d'information, FORMIST, (Formation à l'Information Scientifique et Technique), dont les missions vont de la veille documentaire à la production et à la diffusion de documents pédagogiques et à la formation de formateurs.

Des chercheurs en Sciences de l'information, ont élaboré au début des années 90 des modèles cognitifs de la recherche d'information et l'on ne saurait ignorer les noms d'Eisenberg et Berkowitz, pour la résolution de problèmes, Guthrie et son schéma de spécification de repérage de l'information, enfin et sans vouloir être exhaustif, il nous faut mentionner, Rouet et Tricot et leur modèle EST (Evaluation Sélection Traitement), [Rouet, Tricot, 1998, p. 57-74].

Coexistent et s'entremêlent ainsi des savoirs normés, (Afnor<sup>12</sup>, ISO<sup>13</sup>, etc., ...), et des techniques, (scientifiques ?) plus ou moins élaborés dans l'exercice quotidien des professionnels de l'information et des enseignements spécialisés, dispensés au sein des universités et de quelques écoles dédiées. L'expérience des professionnels, principale boussole pour se repérer dans le flux informationnel, repose à la fois sur des connaissances universitaires et sur un savoir plus ou moins empiriquement acquis, et parfois *appris*. Appris, en effet des divers publics qui n'ont pas peu contribué, par leurs conseils et leurs besoins d'informations, à former en *feed-back* les professionnels.

On le sait, des échanges de savoirs et de savoir-faire ont lieu concomitamment et simultanément parmi les professionnels de l'information (documentalistes, bibliothécaires, etc....).

---

Les travaux de M.D. Grmek à propos du raisonnement expérimental et des recherches toxicologiques chez C. Bernard ont montré que ce raccourci méthodologique n'était qu'une reconstruction intellectuelle *a posteriori* et non la méthode universelle empruntée par l'auteur », [Astolfi, Develay, 1989, p. 23].

<sup>12</sup> Association française de normalisation.

<sup>13</sup> International Organisation for Standardization.

Deux associations professionnels historiques, l'ABF<sup>14</sup>, pour les bibliothécaires et l'ADBS<sup>15</sup>, pour les documentalistes militent activement pour une « défense et illustration » de leur métiers réciproques souvent complémentaires, quand bien même les premiers travailleraient à partir de l'offre d'information et les seconds sur la demande d'information.

### **1.7.1.2. Méthodologie(s) de la recherche d'information : un processus cognitif**

Puisqu'il n'est pas de science de la recherche d'information, il s'est constitué un savoir-faire professionnel empirique, regroupé sous l'appellation générique de méthodologie de la recherche d'information.

La stratégie de recherche d'information à déployer dépend de la question posée et à résoudre, des sources et des supports à utiliser.

Dans leur *Petit guide d'accès à l'information juridique française*, ADBS [S.d.], Stéphane Cottin et Sophie Moyret tentent de définir, page 7, non sans humour, des axiomes de base pour une « meilleure recherche documentaire juridique » :

Axiomes « optimistes

1 - On trouvera toujours ce que l'on se donne la peine de chercher. Au pire, on découvrira que ce que l'on cherchait n'existe pas.

2 - Tout juriste est avant tout un documentaliste juridique. On ne s'improvise pas documentaliste ».

### **1.7.1.3. Les technologies de l'information**

Technologie : définition et précision du terme.

Notre époque est caractérisée par un formidable développement des technologies de l'information. Nous entendons ici par technologie, tout dispositif intellectuel mis en œuvre afin de réaliser des opérations de la pensée au moyen de dispositifs technologiques ou non.

#### **1.7.1.3.1. Technologie des stratégies de recherche et résolution de problèmes**

« Le langage est une carte et non un calque », [Deleuze, Guattari, 1980, p.98].

#### **1.7.1.3.2. Stratégie de recherche d'information**

Explicitation d'un concept :

Le mot stratégie n'est pas à prendre évidemment au sens militaire, (art de faire évoluer une armée), mais au sens figuré, « d'ensemble d'actions coordonnées, de manœuvres en vue d'une victoire », (*Le Robert 1*), de but à atteindre. Ici une réponse documentaire et documentée.

#### **1.7.1.4. Exemples de stratégies de recherche d'information simples**

---

<sup>14</sup> Association des bibliothécaires de France : < <http://www.abf.asso.fr/>>

<sup>15</sup> Association des professionnels de l'information et de la documentation : < <http://www.adbs.fr/>>

Il s'agit de définir ou de proposer des stratégies, des techniques de recherches bibliographiques classiques *a priori* simples ou faciles et de mettre en évidence certains processus cognitifs d'échec quant au résultat. Faciles, cela veut dire quelles supposent un minimum de pré-requis scolaires et la compréhension par un adolescent ou jeune adulte de ce qui lui est demandé.

#### **1.7.1.4.1. Les aléas de la recherche d'informations : Stratégies de recherche erronée**

Ne sont pas prises en compte les stratégies de recherche manifestement erronées liées par exemple à une erreur d'orthographe. Que celui qui cherchant la liste des ouvrages de l'auteur des *Rougon-Macquart* nous pardonne, s'il cherche encore dans le catalogue d'une bibliothèque à... Zolla. Nous omettons volontairement les conséquences induites par des comportements d'excès de confiance ou de distraction dans la recherche d'information. Nous savons par expérience que les erreurs orthographiques et autres *fautes* de frappe sur les claviers des ordinateurs sont sources d'innombrables échecs documentaires.

#### **1.7.1.4.2. Stratégies menées sur des supports classiques : le dictionnaire papier**

« Le bon sens est la chose la mieux partagée du monde » ou « Le bon sens est la chose du monde la mieux partagée » (citation exacte) ? Descartes, certes, mais comment vérifier l'exactitude de la citation sans recourir au livre... ou à un tiers, étudiant ou professeur de philosophie, et face à un utilisateur pressé ?

La réponse se trouve à l'entrée « Monde » du dictionnaire *Le Robert* et non... à l'entrée « sens », ce qui démontre au chercheur qu'il fallait un peu de chance et de la persévérance en l'occurrence et qu'il n'y a là aucune espèce de logique dans l'ordre et la préférence des termes retenus de la citation. C'est le hasard qui conduit à la bonne réponse. Bien sûr, un dictionnaire des citations philosophiques eût eu raison de ces incertitudes linguistiques. Mais cet exemple illustre bien les difficultés rencontrées. A savoir que la logique ne préside pas toujours à l'interrogation et constitue un des aléas de cette recherche. (*Le Robert* est réputé pour être également un dictionnaire de citations).

#### **1.7.1.4.3. Modes de raisonnement dans le procès de recherche de l'information**

On le sait, le raisonnement - lorsqu'il y en a un - en recherche d'information repose théoriquement sur un mode hypothético-déductif, (du général au particulier), par opposition au raisonnement déductif, (ou mode inductif), c'est-à-dire du particulier en remontant au général. Pourtant, ces deux modes cohabitent et sont employés en fonction de la question. Lorsque la question porte sur la durée de vie d'une crevette, par exemple, et qu'il n'y a pas de document sur la crevette, la recherche se fera sur les crustacés.

Les problématiques soulevées par les questions répondent aux besoins inhérents d'information de l'être humain. Ce besoin d'information est motivé par

deux objectifs: « un besoin d'information en vue de la connaissance et le besoin d'information en vue de l'action ». [Le Coadic, 1994, p. 47].

Pour résoudre ou espérer résoudre son besoin d'information, une personne peut choisir de se tourner vers un lieu *ad hoc*, le plus souvent une bibliothèque ou un centre de documentation. A elle, de pousser la porte et de venir auprès d'un bureau d'information et ou d'orientation, où elle pourra exprimer son besoin cognitif ou poser (déposer ?) sa question. Pour obtenir la réponse qu'elle convoite ou attend, cet utilisateur *devra* se fier au savoir faire d'un professionnel censé maîtriser un ensemble de technologies et parmi elles, les techniques de recherche d'information.

#### 1.7.1.4.4. Méthodologie de la recherche avec les opérateurs booléens :

Leur nom provient de celui de George Boole, (1815-1864), mathématicien anglais et auteur de la théorie des ensembles ou algèbre de Boole.

Ce sont le « ET », le « OU » et le « SAUF ». Ils sont utilisés dans le cadre de la recherche documentaire informatisée. Il est indispensable de maîtriser ces opérateurs booléens logiques si l'on veut éviter le bruit (obtention de réponses hors sujet) et le silence (ensemble des documents pertinents qui n'ont pas été retrouvés). Ils sont combinés soit avec des termes soit avec des ensembles de résultats de réponses, (historique des réponses). Pourtant, ils sont à manier avec circonspection.

Ainsi, sur le site *Repère* de l'Enssib, l'auteur de « *Comment interroger* », Annie Léon<sup>16</sup>, recommande « de traiter d'abord les termes équivalents séparés par « ou », ensuite d'opérer les croisements de paniers (résultats intermédiaires de recherche), avec des « et ». Les longues équations de recherche comprenant à la fois des « ou » et des « et » sont à éviter (...) car elles aboutissent souvent à des résultats nuls dont on ne peut pas comprendre la raison ». L'opérateur « ET » (parfois noté AND ou encore +), est recommandé pour réduire le nombre de réponses à une requête. Il donne l'intersection de deux ou plusieurs ensembles. L'opérateur « OU » (parfois noté OR), est utilisé pour augmenter le nombre de réponses à une question. Il est le résultat logique de la somme de deux ou plusieurs ensembles appartenant aux mêmes ensembles. L'opérateur « SAUF » (parfois noté NOT ou encore -), est employé pour exclure des réponses ».

Avec le « et », il faut toujours penser à mettre en tête de l'équation de recherche l'élément le moins probable (le plus *petit*). Avec le « ou », l'élément le plus probable. C'est l'algorithme de programmation qui impose ce *rangement* des informations<sup>17</sup>.

Il est important de savoir que les systèmes d'information automatisés utilisent parfois d'autres opérateurs, tels les opérateurs de proximité que sont (ADJ), (NEAR), par exemple, qui permettent de rechercher la position d'un terme lors de recherches en texte libre. La troncature autorise la recherche à tous les mots dont la racine est commune.

Mal utilisés ces opérateurs peuvent produire des résultats inattendus.

## 1.8. Outils de recherche

---

<sup>16</sup> Léon, Annie. Conservateur des bibliothèques. URFIST de Paris. [anleon@ccr.jussieu.fr](mailto:anleon@ccr.jussieu.fr) <<http://reperere.enssib.frontOffice/afficheArticle.asp?idThem=11> > Consulté le 26/08/2007.

<sup>17</sup> Précisions données par G. Delamotte, dans son cours de Master 1, (2007), sur les fichiers à Paris 8.

En approchant rapidement quelques-uns des outils à disposition du *rechercheur* d'information, nous souhaitons montrer au lecteur que leur méconnaissance ou mésusage entraîne des difficultés lors de la résolution de problèmes d'information. En portant notre attention sur le thésaurus\*, nous rappelons d'un part que l'utilisation pertinente de cet outil vise avant tout à éviter le bruit\*et le silence\* documentaires « grâce au contrôle des homographes et au rapprochement des synonymes », [Calcaly, 2008, p. 262], lors des recherches et que d'autre part, le thésaurus\* est également « une liste structurée de concepts destinés à représenter de manière univoque le contenu des documents et des questions dans un système documentaire déterminé, et à assister l'utilisateur dans l'indexation des documents et des questions. », [Mkadmi, Saleh, 2008, p. 23].

### **1.8.1. Le thésaurus : définitions**

Le thésaurus\* établit des relations entre les descripteurs d'un domaine et les hiérarchise entre termes génériques, spécifiques et associés.

Un thésaurus\* est une liste organisée de termes descripteurs sélectionnés pour indexer une documentation relative à un domaine de la connaissance, ou à un secteur d'activité, [Richter, 1990, p.165]. Selon l'AFNOR, les descripteurs sont « des termes, des mots ou des expressions retenus parmi un ensemble de synonymes, de quasi synonymes et de termes apparentés, pour représenter de manière univoque un concept (...) ».

Claude Morizio, augmente cette définition, en précisant que dans cette liste de termes normalisés «on aura éliminé les problèmes de polysémie et de synonymie, qui représentent des concepts pour un domaine de connaissance voire tous » [Morizio, 2002, p. 40]. Le nombre de descripteurs est variable et dépend du thésaurus\*. Il revient à l'utilisateur de les combiner entre eux dans une équation de recherche. Les thésaurus\* sont dits langages contrôlés\* par opposition aux langages libres de mots clés.

L'ensemble des termes qui constituent le thésaurus\* est réparti en champs sémantiques\*, à l'intérieur desquels les descripteurs sont reliés par des relations de significations qui sont de 3 types : la relation d'équivalence, la relation hiérarchique, la relation d'association.

#### **1.8.1.1. Les relations dans le thésaurus**

La relation d'équivalence : les termes équivalents, appelés aussi non-descripteurs, sont les synonymes, les quasi synonymes et les termes voisins ou associés. La relation d'équivalence est notée EM (employer) ou EP (employé pour) dans le thésaurus\*. Cette relation qui élimine les risques de synonymie et de polysémie, lève les ambiguïtés du langage, recherche le terme univoque, distingue les descripteurs des non-descripteurs.

La relation hiérarchique : selon l'AFNOR, il s'agit d'une « relation entre deux descripteurs dont l'un est subordonné à l'autre ». L'intérêt des relations hiérarchiques est de répartir les descripteurs aux différents niveaux de spécialisation. Il existe deux types de relations hiérarchiques : une relation entre termes génériques [notée TG dans le thésaurus\*] et spécifiques [notée TS dans le thésaurus]. Un terme B est dit spécifique d'un terme A si l'on peut répondre

affirmativement à la question « B est-il A ? ». Par exemple : le moineau est un oiseau.

Une relation de partitivité. Le terme B est une partie ou une composante du terme A. Par exemple : le tiroir est une partie de l'armoire. (B est un élément de A).

On aperçoit ainsi les possibilités d'introduire dans les questions des opérateurs logiques (et, ou, sauf).

La relation d'association. L'Afnor précise qu'il s'agit de « relations qui indiquent des analogies ou liens de signification entre les descripteurs, autres que la relation hiérarchique ou d'équivalence ».

Dans la relation d'association, [Notée TA où VA (voir aussi) dans le thésaurus\*] les deux concepts s'évoquent mutuellement. « L'exemple le plus simple est la relation entre une pathologie et ses symptômes : grippe et fièvre ou toux ; guérison et soin, [Prax, Larcher, 2004 p.125-133].

La relation d'association peut s'établir avec des descripteurs d'un autre champ sémantique Dans la pratique, le concept de relation d'association est un concept « vague ». Les thésaurus sont souvent complétés par une liste dite « permutée » : chaque groupe de mots est représenté autant de fois qu'il comprend de termes descripteurs.

Une mauvaise utilisation des descripteurs, un thésaurus\* mal construit, perturbent le processus d'extraction de l'information. Des thésaurus en ligne sont disponibles sur le réseau des réseaux.

### **1.8.1.2. Evolution des techniques de recherche d'information**

Les modalités de la recherche d'information sont en pleines transformations et expansion. Le développement de l'outil informatique constitue-t-il une rupture épistémologique et définitive avec la recherche papier d'autrefois, c'est-à-dire, ce que Bruno Menon qualifie plaisamment avec nous de « modèle de la boule de billard », par opposition à l'actuel modèle navigationnel de lien en lien de l'hypertexte, (ou hypermédia) ? En effet, la recherche pratiquée naguère dans les catalogues bibliographiques sur support papier conviait pratiquement le *rechercheur* d'information à aller d'une référence à l'autre, (le plus souvent un nom, ou un mot sujet) dans un court et limité espace informationnel, (le plus souvent dans les imprimés) puis d'une page à un numéro de notice, (ou inversement).

Par exemple, pour qui devait effectuer des recherches dans « *Biological Abstracts* », une bibliographie américaine imprimée spécialisée en biologie, il fallait repérer l'ouvrage d'abord, déterminer ensuite la date, puis le volume idoine, puis le fascicule, puis utiliser « *Major concept* » lequel donnait la page où l'on trouvait la première notice, ou utiliser « *Subject guide* » qui donnait la liste des sujets et le numéro des notices, etc..... [Such, Pérol, 1987, p.227 et *sqq.*]. Pour autant, l'interrogation obéit toujours à une stratégie de recherche qui emprunte ses principaux modes actuels aux anciens modèles.

Les stratégies de recherche informatisée ne sont rien d'autre que des stratégies assistées par l'ordinateur qui démultiplie les vitesses d'exécution. Des cyberstratégies.

Des cyberstratégies qui augmentent de façon exponentielle les domaines d'informations repérées, fouillées, traitées, communiquées.

### **1.8.1.3. Stratégies de recherche documentaire informatisée**

Si la recherche d'information nécessite une stratégie, c'est que précisément l'information demandée n'est pas disponible immédiatement et qu'il faut ou la retrouver ou la chercher. Plusieurs modes de recherches sont possibles :

- 1 – la recherche par arborescence ou menu ;
- 2 – la recherche par champs ;
- 3 – la recherche par liens hypertextes.

On peut dater la recherche informatisée des débuts des années 1960. Elle est concomitante du développement des technologies de l'information et de la communication qui préfigurent les outils actuels sur l'internet.

Elle s'est développée en deux directions, la première selon les besoins des usagers et la seconde, selon le besoin de l'industrie des grands serveurs de bases de données, (Dialog, etc.....). La télématique française (Minitel) étant un cas spécifique.

#### **1.8.1.3.1. Les requêtes : un nouveau paradigme ?**

Les langages d'interrogation de bases de données. Il existe trois types de langages d'interrogation :

- 1 - les langages colonnés ;
- 2 - les langages pseudo naturels (SQL)<sup>18</sup>
- 3 - les langages en langues naturelles.

Selon G. Delamotte, (Cf. son cours de M1 sur le fichier), les plus faciles sont les langages colonnés qui utilisent le « et » et le « ou inclusif » et les langages pseudo naturels car SQL permet d'interroger n fichiers simultanément *inter-relationnés* entre eux.

#### **1.8.1.3.2. Analyse d'une question pour la recherche en langue naturelle**

Lors d'une recherche en langage naturel, l'utilisateur formule sa question sans se soucier des détails techniques.

Un traitement spécifique, une analyse linguistique de la question posée, est appliquée par le système. Ce traitement se décompose en plusieurs phases d'analyses, qui permettront de passer progressivement de la question en tant que chaîne de caractères, à une représentation linguistique. Cette représentation linguistique permettra la recherche dans les corpus de documents. Elle devra être la moins ambiguë possible et suffisamment structurée pour permettre à la fois la dégradation de la question et la génération de la requête étendue. Plusieurs étapes caractérisent ce processus : L'analyse morpho syntaxique, la segmentation (processus de découpage de chaînes de caractères en formes partout où il ya présence de séparateurs, tirets, ponctuation, etc.), et la lemmatisation. [Ce passage est inspiré de la page 111 [Lallich-Boidin, Maret, 2005, p.111].

---

<sup>18</sup> *Structured Query Language* : langage standard structuré de requêtes.

Une « désambiguïisation » sémantique (utilisation des mots sémantiquement proches de chacun des mots de la question pour élargir la recherche) [Lallich-Boidin, Maret, 2005, p. 198].

Cette désambiguïisation dépend d'abord du contenu des documents, texte, son, images, puis en grande partie de leurs formes et de leur description, car les professionnels de l'information décrivent et caractérisent le contenu des documents avec des outils visant à normaliser la terminologie et à structurer les relations entre les termes utilisés. Le but étant d'éviter la polysémie.

L'indexation analytique à l'aide de descripteurs ou mots clés est le premier mode de représentation d'un document. Les termes sont choisis pour décrire les principaux concepts. (A noter que l'on peut également effectuer des recherches à partir des indices systématiques Dewey<sup>19</sup> et CDU<sup>20</sup>).

#### **1.8.1.3.3. La recherche d'information sur le web dans les bibliothèques**

D'après les plus récents comptes rendus publiés des bibliothèques publiques répondant au questionnaire annuel de la Direction du Livre et de la Lecture, 14% des bibliothèques disposaient d'un accès à internet en 2003 et 9% seulement des usagers avaient connaissance d'un site concernant leur bibliothèque. [Maresca, 2007, p. 119]. De plus, selon l'enquête du Credoc (Centre de recherche pour l'étude et l'observation des conditions de vie), réalisée en 2005 pour la Direction du livre et de la lecture, 55% des usagers inscrits dans une bibliothèque disposaient d'un accès personnel à internet. Si 26% des français utilisent internet pour leur recherche d'information, ce pourcentage augmente jusqu'à 30 % pour les usagers des bibliothèques. Pour 17 % des personnes interrogées « la bibliothèque est un lieu utile aux enfants pour la recherche documentaire » [Maresca, 2007 p. 171]. Malheureusement cette étude ne quantifie pas les usages et les pratiques de recherche documentaire en bibliothèques publiques. (Les études récentes manquent).

#### **1.8.1.3.4. Problématique de la recherche sur le web : éléments**

La recherche d'information sur le web a introduit de nouvelles problématiques au regard des ressources informationnelles devenues classiques voire obsolètes pour certains professionnels, papier et bases de données des années... d'avant le web ! (1995). Notons ce fait technologique, sur le web, nul ne sait où sont précisément stockées, (dans quels *tuyaux*), les données.

Hétérogénéité et variété de contenus, (commercial, universitaire, grand public,...). Absence ou peu de contrôle, mais proliférations des normes et des standards de toutes natures, (*HTML*<sup>21</sup>, *XML*<sup>22</sup>, *RDF\**, *Dublin Core\**) et des tags\*, (indexation libre).

Croissance exponentielle des informations et des formats. Impression de facilité de disponibilité de l'information. Evolution technologique incessante. Existence d'un « web invisible », c'est-à-dire de pages de bases de données générées

<sup>19</sup> Classification décimale de Melvil Dewey (1851-1931), bibliothécaire américain, répartissant les connaissances en dix grandes classes, elles-mêmes subdivisées en dix sous-classes...Très utilisée en bibliothèques publiques et par la Bnf.

<sup>20</sup> Classification décimale universelle. Davantage présente en bibliothèques universitaires.

<sup>21</sup> *Hyper Text Markup Language* : langage d'encodage ou de balisage du web, [Rivier, 2007, p.29].

<sup>22</sup> Extensible Markup Language : langage destiné à remplacer Html.

dynamiquement, « à partir d'une interrogation *via* une zone de recherche, et les pages cachées, qui ne sont reliées à aucune autre, [Rivier, 2007, p.97] qui échappent aux moteurs de recherche, et qui ne sont donc pas indexées.

#### 1.8.1.3.5. Introduction indicative

Loin d'être une technique rigoureuse, la recherche d'information sur internet, en fait les pages Web, requiert l'usage d'outils spécifiques tels les moteurs de recherche, les métamoteurs et les annuaires. En effet, l'information est abondante sur le Web et ces outils généralistes que sont les moteurs de recherche et d'indexation, (Google, Yahoo, MSN, totalisent à eux seuls 80% du trafic mondial), sont devenus incontournables. Nous présentons quelques techniques sans aucune prétention à l'exhaustivité.

L'évolution des techniques de recherche d'information dépend en premier lieu de la formulation des requêtes (questions) et de leur traitement et, en second lieu de la présentation des résultats (réponses). D'après [Jacquesson, Rivier, 2005, p. 55-56], (dont nous donnons un résumé *infra*), dans le cas des requêtes plusieurs possibilités techniques existent : recherche d'une chaîne de caractères séquentiels, recherche indexée (construction d'un index et recherche sur des termes avec leur adresse), recherche booléenne, recherche par opérateurs spécifiques, (adjacence ou distance, recherche sur les termes dans la même phrase etc.,...), recherche qualifiée : la recherche porte sur un type d'information précis, logique floue : exemple *Londres* et *London*.

Présentation des résultats (réponses) : présentation brute dans l'ordre séquentiel, présentation triée par ordre croissant, décroissant, alphabétique ou numérique, présentation par ordre de pertinence ou pondération, affichage du texte.

Disons rapidement un mot des techniques et technologies *pull* et *push*. Pull : l'information est tirée par l'utilisateur, en ce sens que c'est lui qui va chercher l'information. Push : l'information est poussée (*To push*), via l'internet et des technologies automatisées, agents d'alerte, *newsletters*, flux RSS\*, etc., vers la boîte mail de l'utilisateur, qui aura au préalable déterminé un domaine ou un sujet de recherche, (question). A l'usage, ces technologies peuvent être sources de bruit\* informationnel, (trop de flux d'informations continus tuent l'information).

#### 1.8.1.3.6. La recherche des mots-clés pertinents

Une pratique nouvelle chez les professionnels de la recherche d'information sur le web : la recherche de mots-clés *pertinents* avant l'élaboration des requêtes. Il s'agit également de savoir avec quels mots ou expressions les internautes interrogent le web. Ces mots-clés sont des questions qui appartiennent des réponses. (Des réserves sont apportées à cette pratique).<sup>23</sup>

Différents outils peuvent aider à retrouver des mots-clés proposés par les internautes lors de leurs recherches, par exemple : *outiref* : <http://www.outiref.com/>, *wordtracker*, <http://www.wordtracker.com>, (en anglais et payant), on peut également consulter le générateur de mots clés de Google : <https://adwords.google.com/select/KeywordToolExternal>, ou encore utiliser, après

<sup>23</sup> Voir : Arrington, Michael [en ligne] <http://fr.techcrunch.com/2008/04/26/la-recherche-par-mots-clefs-atteint-elle-ses-limites/> consulté le 07 /07/2008.

interrogation, les mots-clés proposés par les métamoteurs *Kartoo* : [www.kartoo.com](http://www.kartoo.com) et *Exalead* : [www.exalead.com](http://www.exalead.com), etc. L'intérêt principal de ces générateurs de mots-clés réside dans le fait qu'ils donnent des indications sur le trafic des occurrences demandées.

Cependant, certaines technologies (Google), dominant aujourd'hui cet univers de la recherche d'information et les enjeux d'une telle domination ne sont pas sans conséquences sur nos sociétés. C'est la raison pour laquelle le succès de Google doit être interrogé aussi bien au plan de sa technologie, que sur le plan des effets culturels d'une telle domination. Les réponses fournies par Google ne sont pas neutres et obéissent à des traitements, technologiques culturels (américains), algorithmiques, commerciaux, politiques, (Google obtempère aux demandes de censure de certains gouvernements), etc.

### 1.8.1.3.7. Qu'est ce qu'un moteur de recherche et d'indexation?

Selon Véronique Thomas et Armelle Mesguisch, « un moteur de recherche est un outil automatique destiné à collecter et indexer un grand nombre de pages Web (contrairement aux annuaires qui recensent des sites). » [Thomas, Mesguisch, 2006 p. 21].

Trois fonctions principales sont réalisées par les moteurs de recherche et d'indexation.

1) Exploration par un robot, nommé spider ou crawler, qui collecte le contenu de milliards de pages Web structurées dans une base de données en champs.

2) Indexation automatique : construction d'un index de la base de données constituée par explorations successives

3) Interrogation de l'index : l'utilisateur lance un ou des mots-clés qui interrogent l'index.

#### 1.8.1.3.7.1. Les balises Méta

Les métadonnées (données sur les données), sont des aides à l'indexation automatique et au repérage de l'information. Dans le cas du langage à balises HTML, en usage sur internet, les balises META sont des solutions standardisées qui décrivent le contenu de chaque document. Ainsi les balises :

<META NAME= « Author » CONTENT= « »>

< META NAME= « Keywords » CONTENT= « »>

< META NAME= « Description » CONTENT= « »> permettent des recherches sur ces champs.

Mais, les moteurs de recherche ne tiennent plus tellement compte des ces balises et « dans la pratique, seuls 34% des sites renseignent les balises Méta », [Lefèvre, 2000, p. 199]. La recherche se fait le plus souvent en texte intégral, c'est-à-dire sur l'ensemble des chaînes de caractères d'un document, (mots).

Les professionnels, (et les autres) utilisent beaucoup de ces « trucs et astuces » de recherche sur le web. Sur son blog, *Life with web 2.0*, Emmanuel de Taillac<sup>24</sup> en a listé une quinzaine, qui vont de « *define* : » pour chercher une définition, « à *related* : » pour la recherche de sites similaires... [Lenormand, 2007], propose aux professionnels une méthodologie et des exercices pratiques qui procèdent de la connaissance de ces tournemains astucieux. Cependant, la

<sup>24</sup> < <http://emmanueldetaillac.fr/index.php?tag=inurl>> Consulté Le 14/08/2007.

source et la technologie d'information la plus prisée des internautes reste Google en raison de son efficacité.

#### **1.8.1.3.7.2. Google : les clés d'une technologie efficace**

Fin 1997, l'invention de Google par Larry Page et Sergey Brin, deux jeunes doctorants de l'Université de Standford (E.U.) va accélérer la popularité et l'accès aux services Web.

#### **1.8.1.3.7.3. L'algorithme du *PageRank***

Un moteur de recherche fournit des réponses à des requêtes générées par des utilisateurs et la pertinence de celles-ci (*relevancy*). 80 % des requêtes s'arrêtent à la première page.

Le moteur de recherche et d'indexation de Google est basé sur un algorithme appelé *PageRank* (parce qu'il donne un rang aux pages Web). Néanmoins, même ses plus zélés détracteurs oublient de signaler que cet algorithme est propriété de l'Université de Standford.

« L'idée de Page et de Brin, pour classer la pertinence des réponses et les hiérarchiser, c'est de prendre en compte non seulement les mots clés qui déterminent un site, mais les liens qui pointent vers lui. » [Cassin, 2007. p. 97].

#### **1.8.1.3.7.4. La recherche par navigation**

La recherche et l'obtention d'informations numériques ont pris une place prépondérante dans la vie quotidienne et professionnelle de chacun d'entre nous. C'est par métaphore que l'on parle de recherche par navigation et parce que l'outil logiciel qui permet cette navigation est appelé navigateur, (*browser* en anglais). *Explorer* et *Mozilla*, sont des navigateurs, qui rendent possible l'affichage des pages web.

La navigation sur internet peut être associée à une promenade intellectuelle et ludique. Cette déambulation non-linéaire aiguise la curiosité et le plaisir de découverte informationnelle, (le succès de l'internet est là). La navigation consiste à aller de lien en lien en cliquant sur ceux que l'ont souhaite.

En 1963, ces liens furent qualifiés de liens hypertextes\* et hypermédia par Ted Nelson, mais c'est Vannevar Bush qui, dans un article célèbre, « *As We May Think* » en 1945, eut le premier l'idée de mettre en relation le savoir humain par le truchement de liens (traces) entre les documents sur le modèle du fonctionnement du cerveau et par association d'idées.

Un lien hypertexte\* ou hypermédia est constitué d'une ancre, (petit programme informatique), qui peut être activée de différentes manières : bouton, soulignement...Chaque site web et chaque document sont identifiés par une adresse Url.<sup>25</sup> Naviguer, c'est cliquer de lien en lien d'une page web, et de page web en page web ou de site web en site web.

#### **1.8.1.3.7.5. Interrogation et navigation**

---

<sup>25</sup> Url : *Uniform Ressource Locator*. « l'Url est le moyen le plus répandu de localiser de façon univoque une ressource Internet », [Rivier, 2007, p. 30].

Les navigateurs autorisent la recherche dans le cyberspace de façon non linéaire. Le chemin de lien en lien est réversible et l'on peut en conserver les traces, (historique de navigation). La recherche (avec les moteurs) est possible sur l'auteur, sur le titre ou mot du titre, par troncature, par sujet, (mots clés ou termes approchant le langage naturel), tag\*, Url, sites web, etc..... Les catalogues de bibliothèques (celles qui appliquent la norme Z 39 50\*), sont accessibles en ligne et consultables sur internet en temps réel. Nota : nous ne traitons pas de la question et des réponses au sens informatique du protocole Z 39 50\* qui fonctionne d'une part, selon des *clients* Z 39 50\*, qui permet d'interroger un bouquet de catalogues : **envoi des questions, réception des réponses** et d'autre part, selon des *serveurs* Z 39 50\* qui rendent chaque catalogue accessible à une interrogation simultanée, (**réception des questions, envoi des réponses**).

### 1.9. Apports de la recherche par navigation

La recherche d'information sur le web a permis de passer très souvent de l'information secondaire (références bibliographiques) au texte intégral (information primaire), apparemment plus facile à se procurer.

## Conclusion de la première partie

Le cerveau humain, doué de capacités fonctionnelles complexes, cognitives et intellectuelles, permet à l'homme d'appréhender son environnement en l'informant à la fois sur lui même et sur le monde. L'aventure de l'humanité et de la connaissance est consubstantielle à celle des technologies intellectuelles. De là, les surgissements ininterrompus et historiques des connaissances et l'explosion de l'information. Dans cette première partie, nous nous sommes étendus non sans raison<sup>26</sup> sur la notion d'information, fortement polysémique pour ne pas dire polémique, parce que l'information est le **corpus**, [Corroy, Gonnet, 2008, p.73] générique dans lequel fouille le *chercheur* d'information. Des réponses à la question des énigmes ou échecs documentaires qui surgissent, imprévus le plus souvent, dans le processus de recherche d'information se trouvent dans ce *corpus* générique, épars et informationnel de l'information, plus précisément d'informations. En interrogeant et en revisitant le concept d'information et des stratégies mises en œuvre pour chercher et trouver l'information pertinente demandée, ainsi : informations demandées, informations nécessaires et informations fournies ne se recoupent pas. Les stratégies adoptées doivent être adaptées, telles des flèches, (questions) lancées vers un but (réponse) et dont la direction doit être corrigée, ( évaluée : pertinence) à mesure des résultats.

Après avoir rappelé *supra* les différentes acceptions du terme information, nous reprenons ici la typologie opératoire d'un ouvrage récent, (juillet 2008), en la résumant succinctement : « L'information est **action** d'informer, un **état**, une **connaissance**, un **contenu** [d'un message], et enfin un **contenant**. », [Mkadmi, Saleh, 2008, p.19]. Nous avons essayé de montrer que la théorie de l'information reste un modèle plausible, malgré les nombreuses réserves et notamment, celles

---

<sup>26</sup> Nous souhaitons montrer ainsi notre vif remerciement pour le cours d'épistémologie de l'information du M1 et le cours de cyberculture du M2 de C. Baltz.

de Pierre Lévy, [Lévy, 1993, p. 26] qui considère que les « modèles systémiques et cybernétiques de la communication » ne sont plus pertinents parce que ces modèles « réduisent l'information à une donnée inerte et décrivent la communication comme un processus unidimensionnel de transport et de décodage », (p. 26). Enfin, l'apprentissage continué et la connaissance des outils et systèmes d'information sans cesse renouvelée restent indispensables pour la résolution pertinente de problèmes, quand même ceux-ci induiraient malgré tout des échecs documentaires.

## **Deuxième partie**

## **2. Le couple questions-réponses et son corollaire la pertinence**

Cette seconde partie interroge le couple questions-réponses et la pertinence qui forment le fondement de ce triptyque documentaire

### **2.1. Le couple questions-réponses**

Pour tenter de répondre à notre question, (*Des énigmes de la recherche d'information*) nous allons paradoxalement désunir le couple questions-réponses et interroger d'abord la question et ensuite la réponse. L'appariement question/réponse étant une des modalités du processus automatisé de la recherche documentaire, (réponse). La question verbale, énoncée oralement, est au centre de notre questionnement. Quel est cet objet de connaissance, d'interrogation et de communication, qu'est une question?

#### **2.1.1. De la question**

##### **2.1.1.1. Répondre aux questions : un usage social induit par la nécessité de communication professionnelle**

De tout temps, les bibliothécaires ont eu à répondre aux questions que leur posaient leurs lecteurs et les simples visiteurs : « Répondre avec douceur aux diverses questions oiseuses, savantes ou de simples curiosité, en donnant avec complaisance les livres qui lui seront tour à tour demandés » [Achard, 1787, cité par Blanc-Montmayeur, Danset, 1984, p. VII].

Aujourd'hui, les professionnels de l'information font face aux mêmes sollicitations que celles rapportées par M. Achard au 18<sup>em</sup> siècle. « Les questions oiseuses, savantes ou de simples curiosité » n'ont pas cessé de résonner (raisonner ?) à l'intérieur des murs des bibliothèques et autres services d'information où les livres et donc la connaissance qu'ils cèlent s'alignent toujours sagement sur des étagères, même si les documents numériques, les hypermédia, y ont fait une apparition remarquée, et davantage encore si la bibliothèque *hybride*, [Eboli, 2007, p.9] d'abord, et virtuelle et numérique dans un futur rapproché ensuite, s'annonce comme un nouveau paradigme informationnel.

Quoiqu'il en soit, les questions des lecteurs et de tous ceux qui pénètrent dans les bibliothèques et les centres de documentation restent, nous l'avons dit, encore à l'ordre du jour.

##### **2.1.1.2. Les lieux de l'information : des appellatifs génériques pour une même fonction**

Appelés bureau d'accueil ou d'information voire poste de renseignement, salle d'information ou d'orientation, etc.,... dans les bibliothèques on parle également de banque de prêt, ces services de référence occupent une position de centralité à haute visibilité dans l'organisation et la fonction documentaire.

(En Angleterre, ces services sont appelés « *reference services* »).

C'est le lieu *précisé, signalé* (au sens de remarquable), du **point de départ** de la question, du message informationnel. Le lieu par excellence, où entre autre se noue l'avenir incertain du **couple question et réponse**. C'est à partir de ce lieu d'information qu'il s'agira pour le *chercheur* d'information de replacer la question *dans* un arbre de connaissance afin d'essayer de cueillir réponse(s).

### 2.1.1. 3. Question : éléments définitionnels

Etymologiquement, le mot « question » prend sa source dans le latin « *quaestio* » recherche, lequel procède à son tour de « *quaerere* », chercher. Les termes voisins ou synonymes sont **énigmes**, interrogation, problème. L'étymologie voisine avec la sémantique.

D'après *le Petit Robert 1*, « une question est une demande qu'on adresse à quelqu'un en vue d'apprendre quelque chose ». « Interrogation » et « questionner » sont donnés par ce dictionnaire comme étant des synonymes, ou des syntagmes en relation de synonymie.

Autrefois, le mot signifiait « Torture [légale] infligée aux accusés et aux condamnés pour leur arracher des aveux. [*Le Petit Robert 1*].

Le terme « requête » lui, est d'usage plus récent. « Demande instante, verbale ou écrite », signale *Le Robert*. Mais ce syntagme vient, selon *Wikipedia*, dans son acception contemporaine, du mot anglais « *Request* », qui correspond à une demande de traitement. Ce néologisme est employé dans l'univers des bases de données informatiques.

Une question est encore une « entrée » (*input*) dans le système d'information et le système de connaissance du *chercheur* d'information. Cette métaphore qui procède de l'informatique ....

« Question » a donné le verbe « questionner » qui signifie poser des questions à quelqu'un, « questionneur » avec le sens de « personne qui aime poser des question », « questionnaire », qui est une série de questions auxquelles on doit répondre et enfin « questionnement », pour le fait de poser un ensemble de questions.

### 2.1.1.4. Qu'est-ce qu'une question ? Question et philosophie

D'après le dictionnaire de philosophie de référence qu'est devenu le « *Lalande* », le terme « question » a connu un sens primitif « d'enquête et d'investigation » qui a disparu, [Lalande, 1991, p. 872]. Le sens premier équivaut pour ce dictionnaire spécialisé à « Ce qui est mis en discussion ; sujet dont on traite. « Ordre des questions de physique », titre de la 5<sup>e</sup> partie du *Discours de la méthode*. Le second sens se comprend dans une acception linguistique : « acte linguistique consistant à énoncer soit une fonction propositionnelle, soit une *lexis*, en marquant (par l'intonation, par la forme grammaticale, ou par un signe de ponctuation) que l'on demande à quelqu'un soit de la compléter dans le premier cas, soit de l'affirmer ou de la nier dans le second ».

C'est de ce second sens « d'acte linguistique » (et d'objet de connaissance), que nous tirerons partie. Car une question posée dans un service d'information est un acte de langage. Notons que ce dictionnaire ne dispose d'aucune entrée au mot ... « réponse », ce qui ne laisse pas de surprendre.

Aristote<sup>27</sup> distingue quatre sortes principales de questions : « le fait, le pourquoi, si la chose existe, ce qu'elle est ».

La philosophie grecque a inauguré et consacré la question comme méthode de connaissance de l'universel et la maïeutique socratique, (accoucher, faire accoucher), n'est rien d'autre qu'une *technologie intellectuelle* écrivait Pierre Lévy au service de l'acte de connaître. Au sein des Universités, le moyen âge scolastique en fera un mode particulier, le commentaire questionné, et un genre littéraire *la questio*, qui avec la *disputatio* « est à la fois un exercice d'apprentissage et une méthode d'enseignement (...). Questions disputées et questions extraordinaires sont discutées par le maître, cette dernière pour le seul prestige. [Libera, 1992, p. 30]. D'autres penseurs (nous n'en citerons que quelques-uns), ont utilisé la question comme outil de connaissance capable de *piocher* dans le réel et nos représentations. Montaigne (1533-1592), et son « Que sais-je ? » dubitatif ne cessera de s'interroger et d'interroger la connaissance. A son tour Kant (1724 - 1804) demandera : « Que puis-je savoir ? Que dois-je faire ? Que m'est-il permis d'espérer ? ».

#### 2.1.1.5. La question en question

Le mot question signifie également, selon la définition proposée par un dictionnaire de philosophie, « demande, interrogation **attendant une réponse** », [souligné par nous], [Baraquin, 2005, p.283].

Poser la question de la question équivaut sur un plan philosophique à poser la question du sens.

Ainsi, une question est en *attente*, de quelque chose qui est une réponse. Mais, précisément, sera-ce *la* réponse, (contenu informationnel). La résolution d'une question est un processus complexe.

Dans le contexte documentaire, l'expérience (processus cognitif), confirme qu'il est nécessaire de comprendre la question si l'on veut obtenir une réponse pertinente. Pour cela, il est indispensable *d'habiter* la question et de se donner le temps de la *réflexion* de la question.

Aussi bien pour les questions orales que pour les questions écrites. La non prise en compte de ces impératifs pratiques, pour ne pas dire catégoriques, perturbera le processus de la recherche d'information. .

#### 2.1.1.6. Questions et rhétorique

Bernard Dupriez dans son *Gradus*, propose la définition suivante du terme question : « Assertion dont le prédicat demande à être complétée ou confirmée par l'interlocuteur », [Dupriez, 1984, p.370]. Il donne le mot « requestion » comme tactique « afin de se dérober à une question par une autre question ». Ce qui peut aussi se comprendre comme une demande d'information complémentaire. Dupriez note encore que « La fonction injonctive devient prépondérante dans l'interrogatoire ». Le questionnement documentaire est proche, au moins dans l'état initial, de l'interrogatoire y compris dans ce que ce dernier a d'inquisitorial.

---

<sup>27</sup> Aristote : *Organon Seconds analytiques*, Vrin, livre II, Chap. 1, 86b, 25-35, p.161-163, [Aristote, 1970, p.161-163].

### 2.1.1.7. Questions : « Qu'est-ce que vous avez sur... ? ». Le mode interrogatoire

Si, en langue française écrite, et d'après les règles de ponctuation, un point d'interrogation [?] termine toujours, sauf accident, une question écrite. « Etymologie : du latin « *punctum* » (« point ») et « *interrogatio* » (interrogation), [Drillon, 1991, p. 341], c'est l'intonation orale de la demande qui, le plus souvent, indique au professionnel qu'on s'adresse à lui, en vue d'obtenir *quelque chose*.

La *forme* interrogative constitue une information. Il faut remarquer que si les formules de politesse sortent le bibliothécaire de sa *réserve*, toutes les questions ne sont pas assorties d'intonations interrogatives. Il en est qui sont énoncées sur le mode affirmatif : « Je voudrais des informations sur ... », « Je cherche la date de ... », « j'ai besoin de ... », etc. Mais le marqueur le plus usuel de l'interrogation reste la forme mélodique « Est-ce que ?... », ou l'inversion verbe sujet : « Savez-vous ?... », ou l'inversion assortie d'un adverbe : « comment puis-je trouver ?... »

La forme interro-négative des questions se fait également entendre sur le mode : « Vous n'avez rien sur ?... » qui semble placer le locuteur en position d'un doute dont il ne demanderait qu'à se libérer. Les tournures linguistiques des questions, la manière dont elles sont énoncées donnent des indications précieuses qui aident à cerner, à *profiler* le locuteur, (origine sociale, niveau de langue, etc.). Se multiplient, sous l'influence anglo-saxonne, les tours « *c'est quoi ? C'est qui qu'a écrit ?...* » etc., en même temps que les tics et les codes de langages populaires, notamment parmi les jeunes.

Toutes ces interpellations sont énonciatrices et annonciatrices de besoin devenu, par prescription, le plus souvent, quasi irréductible d'information. Elles commandent le démarrage des interrelations, la négociation entre l'utilisateur et le professionnel de l'information.

### 2.1.1.8. Questionner et Questionnement

Une question est **toujours un fait nouveau**. Qui surgit dans l'espace informationnel, avec une exigence d'être satisfaite. Oui mais, il s'agit de répondre à un besoin d'information. Quand bien même, nous l'avons vu avec la doctrine de la *Réminiscence*, connaître, c'est reconnaître. Questionner, c'est tenter ou vouloir re connaître *quelque chose*. (En entrant dans l'intellect du questionné, la question génère un processus perturbateur). Nous avons dit de la question, qu'elle est le message.

### 2.1.1.9. Qu'est-ce qu'une question posée à un service de renseignement/ information ?

La nature des questions posées à un médiateur soulève un paradoxe. En effet, « La question n'est pas une expression comme les autres car elle est signe d'ignorance (c'est une demande d'information que l'utilisateur ne possède pas) en même temps qu'elle est révélatrice du savoir ». [Leclerc-Reynaud, 2006, p. 100]. Questionner, c'est mettre à découvert une part de soi et partant révéler qu'on ne sait pas quelque chose. Pour Le Coadic, cette ignorance est une « anomalie de connaissance », mais le paradoxe se rencontre ici dans le fait que « nous en savons assez pour savoir que nous avons un besoin d'information, mais nous n'en savons pas assez pour poser les bonnes questions (...), [Le Coadic, 2007, p. 19].

Posée, la question s'insère en même temps dans l'intellect du destinataire qui tente d'en saisir les contours et d'ébaucher les premiers linéaments de la recherche, (contenu, sources). Si l'on osait, nous dirions que la question lui pose question, ainsi que sa résolution, (réponse). Ainsi, une question peut être vue sous l'angle d'un **défi**, d'un *jeu*, lancé à l'adresse du récepteur (*rechercheur*) : l'énigme ne met-elle pas en jeu des dispositifs cognitifs et intellectuels ? La résolution de l'énigme ne passe-t-elle pas par le désir de faire « preuve » de son savoir ? N'est-ce pas là, l'un des enjeux de la motivation du *rechercheur* d'information ?

### 2.1.2. Quelle est sa nature ?

On peut distinguer la question liminaire, (ou de départ), posée en manière d'introduction, qui fonctionne comme une prise de contacts interpersonnels et avec l'environnement, (réception). C'est aussi pour le questionneur un premier point d'appui, lequel permettra les réajustements futurs en fonction de la première réponse. Noëlle Balley<sup>28</sup>, dans un article savoureux qui force le trait, dont nous extrayons cet exemple, a montré que du côté bibliothécaire, le jargon professionnel, « *le bibliothécais* », était aussi cause, sinon d'échec documentaire, au moins de perplexité pour les utilisateurs.

Question d'un usager : « Je voudrais lire du Dickens en anglais. Pas David Copperfield, je le connais déjà. »

Réponse du bibliothécaire : « Mots des auteurs : commence par : Dickens ET langue = anglais SAUF mots du titre : contient : Copperfield ». [Balley, 2007, p.78-81].

On peut également envisager les questions élucidantes, qui visent à soulever l'implicite de la question et posent d'autres questions (en évitant les rafales de questions...). Ces questions sont délicates à formuler car il s'agit de demandes de précisions.

#### 2.1.2.1. Le contexte influe-t-il ? Comment ?

Certes, une question est posée par un émetteur à l'intention d'un récepteur. La voix humaine est le canal, le médium. Pourtant, ce n'est plus toujours le cas aujourd'hui puisque de nouveaux dispositifs techniques, (*cf. Infra*) les **systèmes** de questions réponses, (*Sqr*), apparaissent et notamment ceux qui relèvent de l'intelligence artificielle (I. A.). L'intelligence artificielle, née des progrès de la cybernétique, désigne « le traitement formel et automatique des informations par une machine. (...) Le but est de décrire, d'expliquer et le cas échéant de simuler les principales dispositions et capacités de l'esprit humain ». [Corroy, Gonnet, p. 36]. Nous n'interagissons pas de la même manière face à un humain et face à une machine.

#### 2.1.2. 2. Le couple questions-réponses dans le contexte documentaire

---

<sup>28</sup> Balley, Noëlle, « Le bibliothécais sans peine : Petit manuel de conversation courante à l'usage des visiteurs pressés », *BBF Bulletin des bibliothèques de France*, 2007, n°3, p. 78-81 [en ligne] <<http://bbf.enssib.fr>> Consulté le 19 août 2008.

Le « besoin d'information », [Le Coadic, 2007, p. 13]. C'est par cette approche orientée-usager quasi conceptuelle, que Le Coadic qualifie ainsi ce qui relève de l'usage individuel de la fonction information. Après avoir relevé les différentes acceptions du terme besoin, Le Coadic, les explicite en notant que, paradoxalement si le besoin physiologique est vital, comme le sont, le boire et le manger, etc., le besoin ou le désir d'information n'est pas inhérent à la vie, mais qu'il est une construction sociale, et ajouterons-nous, qu'il est une prescription sociale.

### **2.1.2.3. Une question *répond* à un besoin d'information**

Cette partie est fortement inspirée des travaux de synthèse de Yves Le Coadic, et plus spécialement du chapitre 3 de son livre : « *Le besoin d'information : formulation, négociation, diagnostic*. Paris : ADBS, 2007. 204 p. (Collection Sciences et techniques de l'information).

Le Coadic, dans une note de bas de page, a dit sa dette aux travaux des deux chercheurs américains, Belkin et Vickery, dont il s'est lui-même largement inspiré pour la rédaction de ce chapitre, [Le Coadic, 2007, p. 50]. Ajoutons encore, que l'auteur a légèrement remanié cette partie qui est une mise à jour de l'édition précédente, qui date de 1998.

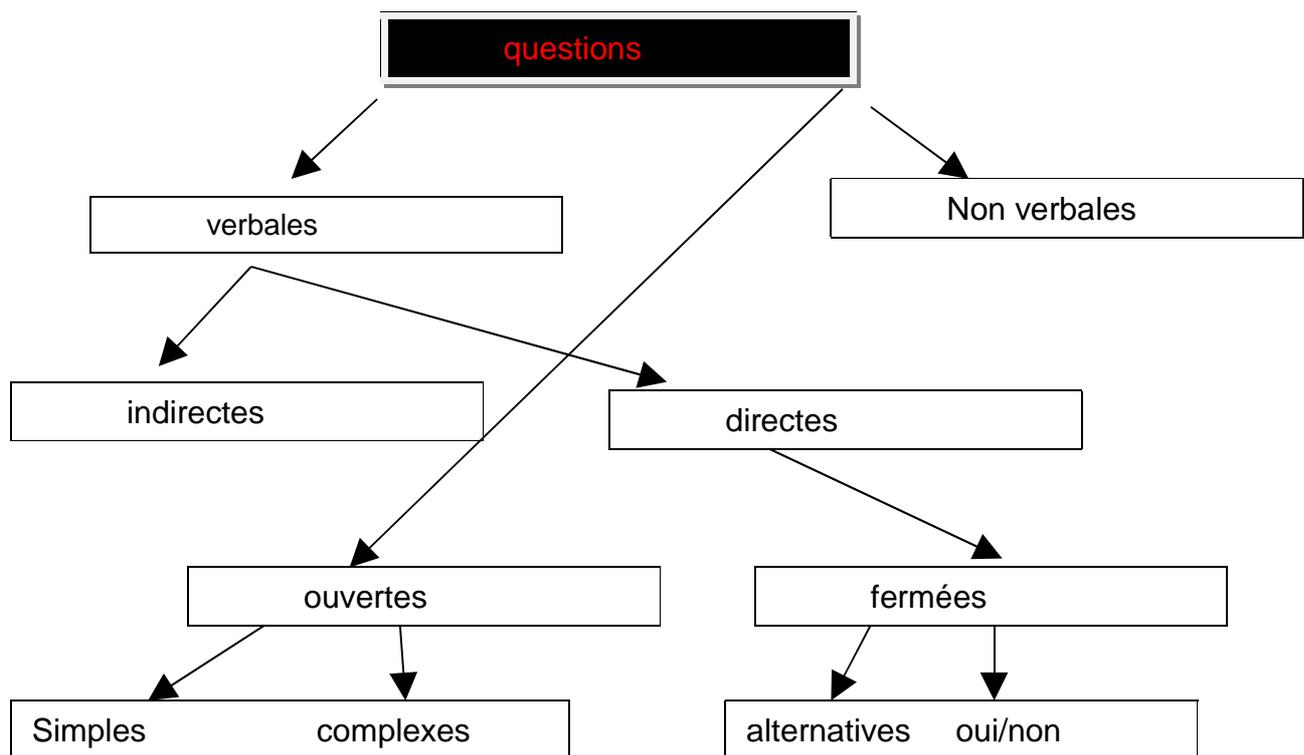
A l'exception notable en France des recherches de Le Coadic, professeur en Sciences de l'information au Conservatoire national des arts et métiers (Cnam), les travaux de recherche sur les problématiques des questions et des réponses n'ont pas, semble-t-il, mobilisé l'attention des chercheurs. La bibliographie en français sur ces sujets est assez pauvre et relève des linguistes.<sup>29</sup>

### **2.1.2. 4. La forme syntaxique des questions**

La forme des questions est principalement syntaxique. Prédominant dans les systèmes d'information, les questions orales directes et indirectes, ouvertes ou fermées.

---

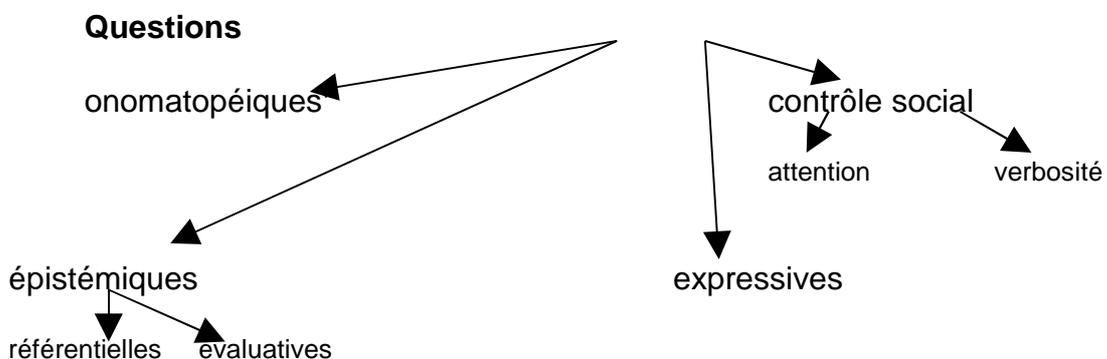
<sup>29</sup> KERBRAT-ORECCHIONI, Catherine (dir.). *La question*. Presses universitaires de Lyon, 1991. 377 p. (Linguistique et sémiologie) ISBN 2-7297-0392-



« D'après des études menées dans quatre bibliothèques publiques américaines concernant le dialogue entre l'utilisateur et le bibliothécaire, ce sont **les questions verbales directes fermées qui sont majoritaires. 90% des questions posées** étaient **des questions fermées** et 8% des questions ouvertes. Les questions ouvertes encouragent le bibliothécaire à répondre plus longuement que les questions fermées ». [Le Coadic, 2007, p.50].

### 2.1.2. 5. La fonction des questions

« Elle est fondée sur des critères sémantiques Chaque question a comme fonction première d'obtenir de la personne questionnée une réponse verbale. Mais alors qu'une question a une forme unique, elle peut avoir plusieurs fonctions ». [Le Coadic, 2007, p.51].



onomatopéiques\* : questions qui demandent que la réponse soit répétée pour confirmer qu'elle a été bien comprise [Le Coadic, 2007, p.52].

### 2.1.2. 6. Une question est une information

Un grand nombre de questions possèdent une fonction d'élucidation. On peut les nommer questions éclaircissantes : ce sont celles qui serviront à démarrer la recherche : « Pourquoi avez-vous besoin de cette information ? Quel est votre niveau d'étude ? Quel est votre but informationnel ?... ». Ainsi, si une question est une information, c'est aussi un réservoir de questions-réponses potentiels.

Une question est encore une problématique définie comme « un art, science de poser les problèmes. Un ensemble de problèmes dont les éléments sont liés », [Petit Robert 1]. Une problématique donc, qui est un ensemble de questions, un thème ou un sujet de recherche, devra apporter un ensemble de réponses documentées.

#### **2.1.2. 7. La formulation de la question comme un indicateur du besoin d'information**

Pour Le Coadic, une question répond à un besoin d'information. Pourtant, dans l'approche bibliothéconomique traditionnelle, « (...) la question n'est pas analysée en tant que telle, mais en tant qu'elle renvoie à un type de source où l'on peut espérer trouver la réponse (démarche orientée émetteur). On cherche à voir ce qu'on peut faire avec le stock. Cette façon d'ordonner les questions montre que l'accent est mis sur la bibliographie et les ressources disponibles dans la collection et qu'une faible attention est portée au développement de la question elle-même ». [Le Coadic, 2007, p. 49].

L'enjeu est de taille ici, car ce qui se joue dans cette démarche orientée émetteur, c'est « que l'hypothèse implicite est qu'il existe pour chaque question une réponse correcte, appropriée et une seule », [c'est nous qui soulignons], [Le Coadic, 2007, p. 49]. Ainsi, à chaque question correspondrait une réponse univoque qui viendrait, tel un bouton-pression, boutonner la réponse. C'est cette réponse idéale qui devrait être la quête, le Graal du *chercheur* d'information, jamais assuré d'atteindre son but.

En fait, dans le cadre des services classiques des bibliothèques, et les modèles organisationnels enseignés par la traditionnelle bibliothéconomie, l'organisation des collections est plutôt « orientée bibliothécaire ». L'utilisateur ayant [aura toujours?] des difficultés à se repérer dans les méandres informationnels se devant de recourir aux services du bibliothécaire rendu en la circonstance omniscient et surtout, on l'a dit, médiateur providentiel et indispensable.

#### **2.1.2.8. Essai de typologie des questions : le point de vue des professionnels de l'information**

Une typologie tentera toujours de cerner imparfaitement des réalités complexes.

##### **2.1.2.8.1. Des questions tous azimuts**

Les professionnels de l'information (documentalistes, bibliothécaires ...), ont coutume de ranger sous deux grandes rubriques génériques, la recherche d'information générale et ponctuelle, et la recherche d'information thématique et/ou spécialisée.

[Beaudiquez, 1989, p. 31] écrit que l'information demandée peut-être rangée en deux grandes catégories : « l'information directe ou immédiate et l'information indirecte, médiante ou différée ». On voit que sous cette typologie fonctionnelle se dessine en outre une notion de temps. Le temps informationnel devient catégorisant, lorsqu'il s'agit d'appréhender et de se saisir du concept de *question*.

Claire Guinchat et Yolande Skoury proposent, dans le volume 2 de leur ouvrage pédagogique que nous citons ici, une typologie empirique, descriptive et précise parce qu'opératoire :

### **2.1.2.9. Aspects temporels des problématiques de catégorisation de contenu des questions**

« Question

**Ponctuelle, limitée, définie.** Exemple, l'eau gèle à..., la loi sur l'énergie solaire est parue le, ...

**Précise, mais étendue dans le temps, « rétrospective »**, le sujet peut être très vaste et mal défini. Les aspects sont multiples et variés. Exemple : état de la législation en France et au Sénégal sur la pêche côtière.

Evolution de la délinquance au Liban sous l'angle social et pénitencier. (Sic).

**« Courante », répétitive avec constante mise à jour.** Exemple : les statistiques du cinéma, chaque semaine ?

Je veux régulièrement les articles parus sur Léopold Sédar Senghor.

**Synthétique, faisant le point d'un sujet.** Exemple : où en est-on sur le séchage solaire au Mali ? Le point sur les télétransmissions en France et en Angleterre ».

Les auteurs insistent sur une nette délimitation entre recherche courante (moins de un an) et qui répond à un type de questions « Qu'avez-vous de nouveau sur le sujet ? » et la recherche rétrospective, qui demande : que sait-on sur le sujet ? », [Guinchat, Skoury, 1996, p. 189].

### **2.1.3. Espace topographique de la question : le lieu de la question**

On **pose** (dépose) une question : idée de lieu.

Une question est portée par un locuteur. Elle s'insère dans un processus de communication interpersonnel, (émetteur /récepteur), lui même inséré dans un cadre spatial, un système d'information, le plus souvent un centre de documentation, spécialisé ou non, ou une bibliothèque, voire une médiathèque. (On parle encore de « réception » dans les hôtels). Ces lieux informationnels revêtent souvent une charge symbolique prestigieuse importante pour celui ou celle qui y pénètre. En effet, on y traite, diffuse et communique de l'information c'est à dire de la connaissance et pour le dire avec les mots d'aujourd'hui, de l'ingénierie des connaissances.

Dans les systèmes informatisés, l'interface rectangulaire insérée sur l'écran de l'ordinateur devient le lieu de la question. Le formulaire de requête, c'est ainsi qu'il se nomme, court espace vierge où s'énoncent les mots de l'interrogation avant que ceux-ci ne s'élancent virtuellement à la rencontre des chaînes de caractères, pour le grand appariement de la plus algorithmique des pertinences.

### **2.1. 3.1. La question distribuée : une gestion anticipée des réponses**

Dans les grands organismes, les questions sont distribuées, c'est-à-dire réparties *a priori* en fonction des domaines couverts par le champ ou la discipline dans laquelle s'insèrera la réponse, et en fonction des compétences des agents répondants. Il s'agit de questions spécialisées, économiques, philosophiques techniques, bibliographiques, etc., ou de question dont la résolution demande une attention particulière et soutenue, « Y a-t-il tel mot à telle page ? telle illustration ? quelle est la date d'entrée d'un manuscrit dans telle collection ?...., (les bibliothécaires, les documentalistes, des universitaires, des chercheurs et autres savants et érudits, constituent l'épine dorsale des ces questions complexes ou pointilleuses).

#### **2.1.3.2. Analyse de la question**

C'est le préalable à toute démarche de recherche.

Selon Marcelle Beaudiquez, la recherche bibliographique doit se comprendre comme « la possibilité de résoudre, sinon une énigme, au moins un problème. » Il est nécessaire, poursuit-elle, « d'établir ce qu'on peut appeler une équation de recherche, que ce soit pour la recherche traditionnelle ou la recherche automatisée ». [Beaudiquez, 1989, p.39]. Cette analyse de la question est menée par le professionnel de l'information qui doit s'assurer de la pertinence des interactions (émetteur/récepteur) liées au processus communicationnel. Cette analyse débouche en particulier sur la traduction de la question en mots clés, (termes significatifs) indices, pistes et bases de connaissances....

#### **2.1.3.3 Les étapes du processus d'analyse de la question dans le contexte de la recherche d'information**

Plusieurs étapes conduisent ou conditionnent le processus d'analyse de la question.

Une question formulée dans un espace informationnel par un utilisateur en quête de satisfaction d'information est un processus qui se déroule en plusieurs phases et interagit constamment entre questionné et questionneur. Questionné et questionneur tour à tour en positions interchangeables.

La reformulation de la demande (il peut s'agir également *d'ouvrir*, par un jeu de questions nouvelles, les questions fermées). La détermination exacte du sujet (dates, contenu, types de documents à identifier, descripteurs, etc.). Le choix des instruments (Sources et outils). Avantages/inconvénients. Les supports (numériques ou non) des documents (texte, son, image). Les besoins de l'utilisateur (quelles sont ses fins en matière d'information). Le niveau de recherche. Les problèmes de mises à jour des informations. Le temps disponible (et de plus en plus les coûts de l'interrogation et de l'information). La forme de la réponse (recherche ponctuelle et factuelle ou recherche thématique, courante ou rétrospective). Les problèmes juridiques, (Droit d'auteur, droit à l'image).

Le feed-back de l'utilisateur (qui a place égale dans le processus d'analyse de la question avec le processus d'analyse de la réponse).

#### 2.1.3.4. Pourquoi est-il nécessaire de connaître (analyser) les questions dans un centre de documentation ?

La connaissance des questions et leur traitement permettent de mesurer et de connaître en profondeur la demande réelle et l'évolution de cette demande. C'est un indicateur d'activité utile pour évaluer l'activité du service d'information. L'analyse des questions, (et de leurs réponses) selon des critères de gestion, permet de cerner les besoins : domaines de la demande, types et nombre de questions, sujets récurrents ou non, origine de la demande, (sexe), période des demandes, etc..... Cette connaissance pourra, par exemple, orienter achats et recherche d'information en fonction des besoins exprimés. Elle permet en outre, de mesurer le taux de satisfaction de la demande (réponse à l'utilisateur).

Voici à titre d'exemple une étude américaine, citée par [Chaumier, 2007, p.100].

« Analyse des questions posées aux bibliothèques de l'organisation de l'énergie atomique américaine.

Analyse sur 4 000 questions

A Types de questions techniques (4/5 de l'ensemble)

Description d'un procédé ou d'une méthode	35,5 %
Propriété d'une substance	24,6%
Description d'un appareil ou d'un équipement	16,8%
Constantes physiques ou chimiques	16,4%
Effets biologiques d'une substance	5,8%
Effets de radiations	2,9%
Autres	8%

En outre, dans le cadre de la politique documentaire au sein d'un système d'information, l'analyse des questions traitées permet d'intégrer ces résultats dans un processus de communication interne, (publicité) au système, (entreprises, grands établissements publics).

#### 2.1.3.5. Reformulation et dégradation de la question

Reformuler une question invite à faire comprendre à l'interlocuteur que l'on a bien saisi les termes de sa demande. Il s'agit bien sûr de dire, (redire) mais autrement. La reformulation fonctionne comme un signal destiné à montrer que l'on a compris le sens et le but de la question. On le sait, Le Coadic les nomme « onomatopéiques ».

Reformuler la question pour « ré – atteindre » le locuteur afin de vérifier si les deux parties, alternativement *émettrices* et *réceptrices* se sont bien accordées sur les buts à atteindre. Car il existe des questions mal formulées<sup>30</sup>, (syntaxe défailtantes, énoncé inaudible, termes imprécis, etc.,...).

#### 2.1.3.6. Dégradation de la question

---

<sup>30</sup> THOMAS, Armelle, en Août 2008, lance sur son blog : un « Avis à la population : cherche exemples de recherches mal formulées pour *Net recherche 3* », ouvrage dont les deux éditions antérieures ont été écrites avec Véronique Mesguich, que nous citons. Disponible sur :< <http://inforizon.blog.com/>>. Consulté le 23/08/2008.

Proche de la reformulation de la question, la dégradation de la question consiste à proposer une suite de reformulations possibles plus générales que la question de départ. « Ce mécanisme est assez courant lorsque nous recherchons une information ou lorsque nous souhaitons nous procurer quelque chose, (...), nous recherchons une solution de rattrapage en relâchant certains critères de départ (...) moins essentiels ». [Lallich-Boidin, Maret, 2005, p.203].

### **2.1.3.7. Bonnes et mauvaises questions**

Existe-t-il de bonnes et de mauvaises questions ? Le débat est d'envergure et nous n'oserions trancher. Pourtant, Le « *Lalande* » parle de « questions mal posées » [Lalande, 1991, p.872], et donne l'exemple suivant tiré d'Aristote : « Pour prendre une question dont l'absurdité saute aux yeux, est-ce la terre qui est la mer, ou est-ce le ciel ? Mais parfois ce vice est difficile à démêler, comme lorsqu'on demande si tel genre de choses est bon ou mauvais, alors qu'il y en a dans le nombre de bonnes et de mauvaises ». [Lalande, 1991, p.873].

On peut raisonnablement postuler qu'une *bonne question* est une question dont on pressent qu'elle va recevoir facilement et *rapidement* sa réponse, soit parce qu'elle est clairement formulée, soit parce sa résolution sera aisée. Une bonne question, (formule journalistique) est une question provocatrice, attendue ou inattendue, « mais je vous remercie de me l'avoir posée !... », pas très loin de la question piège.

### **2.1.3.8. Questions simples et questions complexes**

Existe-t-il des questions difficiles : un débat qui est à peine ébauché chez les professionnels de l'information, (Cf. *supra* le schéma de Le Coadic que nous avons reproduit). Certains, (Marie Montano, *cf. supra*) lui récusent même tout fondement. Les difficultés viendraient de la culture générale, des processus d'incompréhension des énoncés, de l'ambiguïté de certaine formulation des questions, ou l'usager reformulant de nouvelles demandes, (questions dites à tiroir).

On trouve dans la littérature professionnelle les expressions « Questions de bas et de haut niveau » [Denecker, Kolmayer, 2006, p. 243]. Ces deux auteurs citent les travaux de [Rouet, 2002] et donnent les exemples de deux questions du chercheur adressées à des étudiants en note de bas de page (243) : « Quels type de particules Rutherford a-t-il utilisé dans son expérience ? et seconde question « En quoi le modèle atomique de Rutherford différait-il de celui de Thomson ? ». [Denecker, Kolmayer, 2006, p. 243]. Le niveau d'une question est défini par ces auteurs « par la quantité d'information nécessaire pour répondre et/ou pour son caractère littéral ou inférentiel ». [Denecker, Kolmayer, 2006, p.243]. Il semble que la notion de question complexe existe également en référence aux systèmes de questions-réponses automatisés, et aux difficultés rencontrées par leurs concepteurs dans l'extraction des réponses

Existe-il des questions plus difficiles à traiter ? Celles qui demande une grande expertise sont elles pour autant ou peuvent elle être qualifiées comme telles ?

Une question difficile est sûrement une question qui n'a pas (encore ?) reçu de (sa ?) réponse.

### **2.1.3.9. Pragmatique des questions**

La pragmatique s'intéresse au sens que les énoncés prennent dans un contexte donné.

Là où les questions énoncées vont déclencher des processus d'actions : des processus cognitifs et physiques. Processus d'action de médiation : processus de recherche d'information tournés vers la résolution de problèmes : la recherche de réponses.

## 2.2. Des réponses

Les réponses ou la réponse est le but de la recherche d'information. Elle clôture une série d'étapes de recherche et fournit un résultat à un destinataire final et pourra prendre plusieurs formes documentaires. Exception faite des cas où la réponse est jugée *a priori* connue par le *chercheur* d'information, elle sera alors délivrée d'emblée et pourra éventuellement dispenser d'une recherche.

« Proposition : 65 - Une réponse qui ne peut être exprimée suppose une question qui, elle non plus ne peut-être exprimée. L'énigme n'existe pas Si une question se peut absolument poser, elle peut aussi trouver sa réponse », [Wittgenstein, 2001, p.111].

Si dans la vie courante on **pose** une question, en général, on **donne** une réponse. A l'idée d'un lieu *utopique* répond la *gratuité* de l'information remarquée par T. Jefferson. Si, comme nous l'avons affirmé, une question est une information, *a fortiori*, une réponse va de l'information brève, à un ensemble d'informations d'initiation, de vulgarisation et de spécialisation.

### 2.2.1. Qu'est-ce qu'une réponse ?

Qu'est-ce qu'une réponse ? Il faut paradoxalement et inévitablement poser la question pour obtenir un ensemble définitoire. (Ne dit-on pas que la réponse est déjà dans la question ?). Le mot réponse est défini par le dictionnaire *Le Petit Robert 1* comme : «ce qu'on dit en retour à celui qui vous a posé une question, fait une demande » (...). Au sens strict : interrogation sur un sujet donné ».

Dans le même mouvement de balancement dialectique, la réponse n'existe pas sans la question. Telle est sa condition.

« La réponse, loin de se faire attendre, est déjà impliquée dans la question et ne cesse pas elle-même, quand elle est incisive, de demeurer énigmatique. En même temps, la réponse constitue le nerf de la question. L'énigme du surgissement de la question et l'énigme de la réponse ouvrante relèvent d'un même centre vibrant », [Axelos, 2005, p. 67].

Le *Trésor de la langue française* [en ligne] définit la réponse ainsi : « Paroles, écrits, gestes symboliques ou mimiques adressées en retour à ce qui est écrit, dit ou demandé ».

On remarque le pluriel des éléments définitionnels. Une œuvre d'art, (tableau, poèmes, une œuvre musicale, etc.), est aussi une réponse. La sublimation créatrice, comme *forme* de réponse à une frustration, à un besoin impérieux.

Ainsi, et pour nous résumer, une réponse sera entendue comme *tout* ce que l'on cherchera et trouvera qui fera sens et s'accordera avec les besoins d'un utilisateur.

## 2.2. 2. Le stimulus : une réponse conceptuelle

Il est singulier de noter que le cognitivisme, né dans les années 60, est ...une réponse au behaviorisme, école et courant de pensée psychologique antérieurs, qui considère que la psychologie scientifique doit se concentrer sur des réalités observables; les stimulations et ...les réponses de l'organisme.

Une réponse est une « sortie » (*out put*), du système d'information et du (cerveau du *rechercheur* d'information).

## 2.2.3. Théorie de l'information et réponse

En raison de son caractère novateur et scientifique, la théorie de l'information a connu des succès inattendus et un grand nombre de chercheurs de toutes disciplines s'en sont emparés. On trouve dans l'ouvrage d'E. Dion, un exemple probant d'application en science cognitive, en l'occurrence, une réponse pratique dans l'élucidation des capacités mnémoniques humaines.

Emmanuel Dion rappelle dans son *Invitation à la théorie de l'information*, que la théorie de l'information [1948], de C. Shannon, a été utilisée en psychologie expérimentale dès les années 1950, [Dion, 1997 p.129 et sqq.]. Il s'agissait alors de **répondre** à la question suivante : « Quelle quantité d'information un cerveau humain est – il capable de traiter ? ». Des protocoles expérimentaux avaient été élaborés, et après resserrement du champ d'expérimentation, les expériences portèrent vers les capacités sensorielles des individus à traiter l'information. « L'expérience consistait le plus souvent à faire varier à la fois le nombre des stimuli possibles, ainsi que le nombre des réponses possibles, de manière à observer à partir de quel moment les sujets commençaient à les confondre, mettant ainsi en évidence leurs limites à traiter l'information reçue. Par exemple, on utilisait quatre tons de gris différents, puis huit, puis douze, puis vingt, etc., en donnant à chaque fois le nombre de réponses possibles correspondant ». [Dion, 1997, p. 131]. Toujours selon E. Dion, la théorie de Shannon a ainsi permit de répondre à la question initiale que se sont posés les psychologues : « (...) le cerveau humain ne semble guère capable de les [les catégories] classer en plus de six ou sept rubriques différentes, sauf à risquer de les confondre avec les autres ». [Dion, 1997, p. 131].

Ces chercheurs utilisèrent la quantité d'information calculée « (...) dans les termes de Shannon » et les résultats montrèrent que « la capacité limite de traitement humain était (...) comprise entre 2 et 3 bits, correspondant à peu près à la capacité à distinguer six ou sept catégories (...). Ce constat ayant conduit à la formulation de la règle selon laquelle le recodage est à la base de l'amélioration des capacités mnémonique de tout individu, il est resté d'une grande importance dans l'histoire de la psychologie cognitive ». [Dion, 1997, p.131].

## 2.2.4. Essai de typologie des réponses : une pragmatique des réponses

Si les professionnels ont catégorisé les questions, il n'en est pas de même pour les réponses. Une pragmatique des réponses signifie que la typologie adoptée en l'occurrence, faute d'en trouver une meilleure et plus conceptuelle, trouve son cadre de référence dans le *lieu* d'un espace d'information et l'intellect (*siège* dans le cerveau), de celui qui y œuvre.

**Neutralité des réponses :** La neutralité de point de vue : privilégier l'information sur les pré-jugements en multipliant les sources et en indiquant - si possible – les interprétations plausibles. C'est une aporie.

**Sémantique des réponses :** Une réponse fait sens quand elle rencontre l'approbation de son destinataire : l'oracle de Delphes disait à celui qui venait le consulter : « ce n'est pas ce que je dis qui importe, mais comment interprète le message sibyllin celui qui le reçoit .... ». Une réponse, pour faire sens doit rencontrer le... sens de la question. C'est une contrainte.

### **2.2.5. La réponse comme contenu informationnel**

Répondre, équivaut à résoudre un problème d'information. Il s'agit pour le *rechercheur* d'assister le demandeur dans sa quête de besoin informationnel. Une réponse est un contenu informationnel qui permet de « compléter, de confirmer ou d'infirmer une question », [Dupriez, 1984, 385]. En additionnant contenu informationnel, et divers processus cognitifs, résultats de modalités de recherche, stratégie, méthodologie et savoir-faire. Selon nous, l'objet réponse n'est pas une affirmation, c'est une *proposition*.

**2.2.6. Deux types de réponses sont obtenus au sein des systèmes d'information : Les réponses *clôturantes* et les réponses *ouvrantes*.**

#### **Les réponses *clôturantes***

Notre catégorisation est issue de celles des questions. Ce sont bien entendu, toutes les réponses définitives, en termes de résolution. Elles closent l'interrogation, ferment la question, quand bien même cette solution ne serait que provisoire. Le plus souvent, elle *répondent* par « oui » ou par « non », par « vrai » ou par « faux », « on a », ou « on n'a pas » etc.... Ce sont toujours des affirmations positives, au sens ou celui qui dit « non » affirme la vérité de son énoncé, (peut-on parler de réponses booléennes ?).

Sans réplique, elles fonderaient une part de l'autorité intellectuelle du professionnel.

#### **Les réponses *ouvrantes***

Ouverture sur d'autres problématiques liées à la question initiale, ouvertures vers d'autres azimuts de la connaissance, (contenu informationnel) ces réponses enchaînent d'autres questions qui à leur tour entraînent de nouvelles interrogations....

Puisque, nous l'avons dit, la question est action, la réponse, elle, est réaction à la question. D'où le feed- back.

Le feed-back, rétroaction voire boucle de rétroaction. La cybernétique, (Norbert Wiener) étudia, entre autres, les processus de commande et d'autorégulation des systèmes vivants et des machines. La notion de rétroaction, action d'un élément sur un autre, est une réponse en retour.

**2.2.7. Répondre et déontologie du *rechercheur* d'information : le *droit* de réponse**

« Répondre- engage une certaine responsabilité- en répondre » [Axelos, 2005, p. 69].

L'obligation d'information s'impose au fonctionnaire dans l'exercice de sa mission de service public et en l'occurrence au bibliothécaire, (Cf. Statut du fonctionnaire).<sup>31</sup>

Déontologie, du grec *deon,-ontos*, « ce qu'il faut faire, ce qui convient » et « logos » science, discours. (On remarque ce sens apparenté avec pertinence : ce qui convient également). « Il n'est en usage en français qu'à partir du XIX<sup>e</sup> siècle, étant traduit de l'anglais *deontology* présent dans le livre de Jeremy Bentham, *Chrestomatia*, (...). La déontologie est un cadre normatif précisant les règles à respecter par l'ensemble des membres d'une profession et la conduite à tenir entre professionnels mais aussi avec leurs clients ». [Corroy, Gonnet, 2008, p. 81].

« Répondre, du latin *respondere* –veut dire s'engager en retour, par rapport à la question, et veut également dire en répondre ». [Axelos, 2005, p. 68].

### **2.2.8. Invariabilité et prévisibilité de la réponse : le cas du dictionnaire**

Un mot du dictionnaire, son entrée dans le corpus du lexique est, en un sens, une question que l'on s'est posée au préalable ; la définition d'un terme, donnée de manière définitive (académique) et générale, en explicite les acceptions et équivaut de fait à une réponse. Le plus généralement forgée de certitudes définitives académiques. Invariabilité relative car dépendante de l'évolution de la langue et des modes, mais réponse pertinente puisqu'elle fera sens pour un temps et pour un groupe social déterminé.

On peut appliquer aux entrées des encyclopédies le même raisonnement en considérant cependant que la longueur d'un article d'une encyclopédie, donc de la réponse, n'est pas rédhitoire.

### **2.2.9. Variabilité et imprévisibilité de la réponse comme message informationnel**

« Shannon et Weaver en 1949, en travaillant sur la transmission des information ont souligné l'importance du degré d'imprévisibilité d'un message. Trop prévisible, un message apporte moins d'informations et suscite moins d'intérêt. Imprévisible, il devient difficile à comprendre car les informations sont trop nouvelles. Pour obtenir un message de qualité, la clé est souvent l'équilibre entre la redondance, c'est à dire la répétition des informations importantes pour qu'elles soient bien comprises et l'originalité des informations transmises », [Corroy, Gonnet, 2008 p.36-36].

## **2.3. La réponse comme élément du message informationnel**

Tout message, selon le schéma canonique shannonien subit des pertes, par son ambiguïté ou les conditions dans lesquels il est émis (bruit, perturbation, parasitage....), le récepteur tenant compte du contexte et de ce qu'il sait de l'émetteur. Notons, que lors de l'échange informationnel, (dialogue avec l'utilisateur), se forment des boucles de rétroactions, les positions d'émetteur et de récepteur

---

<sup>31</sup> Loi n°836 34 du 13 juillet 1983, article 27 : « Les fonctionnaires ont le devoir de satisfaire aux demandes d'information du public dans le respect des règles mentionnées à l'article 26, loi du 13/07/83 ».

varient constamment, et plus encore lorsque la durée de l'échange augmente, (demandes de précisions, ou d'éclaircissements).

### **2.3.1. Le principe théorique de non répétabilité des réponses**

En principe, une réponse est unique et adaptée à un interlocuteur. Il ne saurait être question (*sic*) de répéter à l'envi le résultat d'une requête. Donner des réponses, c'est donc tenter de proposer une *forme* de création renouvelée pour chacun des utilisateurs. C'est tenter d'agir pertinemment, d'ajouter de la pertinence, et pour le dire avec Leclerc-Reynaud d'apporter des « réponses inventives ». Documentaliste, Sylvie Leclerc-Reynaud cite le cas d'un usager qui demande ce qu'est un kolam, (représentations à base de pétales de fleurs, de poudres colorées, etc.....que l'on trouve en Inde). Non seulement le documentaliste, nous rapporte-t-elle, fournira la réponse sous forme de manuel, mais livrera en sus les produits nécessaires à la confection des kolams, [Leclerc-Reynaud, 2006, p.136-137]. Il s'agit ici d'adapter la réponse au contexte de la demande, qui est une plus value-créative.

### **2.3.2. Temporalité contextuelle des réponses**

On a vu qu'une entrée de dictionnaire ou une entrée d'encyclopédie, (et une encyclopédie), ou plutôt les définitions qui les accompagnaient, pouvait être considérée comme des réponses, et que celles-ci dépendaient du contexte temporel d'énonciation.

Une réponse à une question correspond à un moment du temps des connaissances. C'est une somme de connaissances. Elle est contemporaine des connaissances de son temps. (B. Croce). Comme une encyclopédie générale reflète un moment du savoir humain qu'elle synthétise. « Elles [les encyclopédies] jalonnent les décennies comme autant de miroirs, terme qui fut longtemps leur intitulé même : *Speculum majus* de Vincent de Beauvais (1244) ou *Miroir universel des lettres et des sciences* de Léonard Fioravanti. (1586) ». [Beaudiquez, 1989, p. 76].

Ainsi une réponse existe d'abord dans un contexte spatio-temporel donné, puisqu'elle est subordonnée et étroitement dépendante des connaissances admises à une époque donnée.

### **2.3.3. La réponse différée**

Différer une réponse, c'est admettre que le *rechercheur* n'est pas parvenu à la bonne *formule*, qu'il ne sait pas ou bien où et comment trouver, ou bien que le temps de la recherche lui fait défaut. Une réponse différée est une réponse laissée en attente mais dont le terme ou la résolution aura été négociée avec le questionneur.

#### **2.3.3.1. Délimitation de la réponse**

Selon Eric Sutter, *op.cit.*, p. 200-201, « Une partie de la réponse peut être hors champ de spécialisation de l'organisme (...); « ou nécessiter une interprétation ». Rendue indispensable par la complexité ou l'étendue du champ de certaines réponses, il apparaît parfois nécessaire d'indiquer le bornage de la réponse en interrogeant par exemple l'utilisateur sur le pourquoi de sa question.

### 2.3.4. Valorisation de la réponse fournie

Sutter distingue trois domaines de complémentarité ou valorisation qui accompagnent une réponse donnée en situation d'exercice professionnel :

1 pratique : le professionnel « complète ses premières informations par des conseils de mise en œuvre, (...), des recommandations »,

2 pédagogique : en cas de « lacunes » chez le demandeur,

3 médiatique : en cas de réponse partielle, le professionnel de l'information, s'enquiert auprès de l'utilisateur afin de savoir comment « il peut poursuivre son chemin », [Sutter, *op.cit.* p.201].

A ces trois domaines correspondent trois principes professionnels : « qualité : la réponse doit être de qualité, axée sur la complétude et la valeur intellectuelle de l'information ; rapidité : la réponse doit être donnée rapidement, pertinence : la réponse est en adéquation avec la demande », [Calenge, 1996, p.49].

#### 2.3.4.1. La réponse incertaine ou approximative

Le plus souvent à cause du manque de dialogue, ou de connaissance du sujet, il arrive parfois que les réponses conservent une marge d'incertitude. La réponse est alors incomplète. Cas des réponses comportant une marge d'incomplétude. Cas des variations des réponses produites par des professionnels : ils ne donnent pas la même réponse à une même question. (Comment l'esprit du *chercheur* d'information *indexe* la question ?)

« C'est tout ce que vous **avez** sur ce sujet ? » (Entendre, c'est tout ce que vous **savez**...).

« Vous n'avez pas répondu à ma question ».

#### 2.3.4.2. La réponse localisée

On appellera réponse *localisée*, toute réponse trouvée et fournie dans la bibliothèque ou le centre de documentation avec les seules ressources disponibles immédiatement, et sans qu'il soit procédé à des recherches *hors les murs*. Cette réponse peut-être sommaire ou complète, si elle est fournie par un organisme spécialisé dans le domaine, (champs de connaissance) de la question. Ainsi, est une réponse localisée spatialement, la réponse de « ce chercheur qui souhaite savoir si la bibliothèque de Charleville possédait, à l'époque où Rimbaud la fréquentait, une collection de *L'illustration*. (...) ou une étudiante en maîtrise d'histoire de l'art qui s'intéresse aux illustrateurs de l'œuvre de Rimbaud pourra, grâce à une recherche relativement simple dans le catalogue, disposer de l'ensemble des éditions bibliophiliques conservées à la bibliothèque ». [Perrin, 1999, p.106-107].

#### 2.3.4.3. La réponse modifiée par les débuts de la recherche

Il arrive souvent qu'une demande mal précisée ou incomprise, (ou pour d'autres raisons), par le *chercheur* voie modifiée ses premiers résultats. La soumission des premiers éléments de la réponse appelle des corrections de stratégies, de trajectoires et de domaines de recherche. Quand bien même l'anamnèse documentaire aurait été menée avec art. Cette réponse se construit simultanément avec la question, devenue alors un ensemble de questions et de propositions qui réajustent le but et les visées du demandeur.

#### 2.3.4. 4. La non réponse ou l'absence de réponse

Ne pas répondre, ne pas trouver de réponse à une question constitue paradoxalement une ...réponse, qui certes, ne *répond* pas à la question posée, mais en constitue en quelque sorte les linéaments ou les prémisses.

L'absence de réponse entraîne d'autres questions ou peut mettre fin à l'interaction.

#### 2.3.4.5. Type de réponse erronée

Démarche à contre sens. Cause de l'erreur : absence de questionnement, donc quiproquo.

« Je voudrais des informations sur les stratégies d'acteurs ». La stratégie d'acteur consiste à préserver sa liberté d'action. L'analyse stratégique estime que chaque acteur est en mesure de le faire dans les systèmes où il évolue. Tout repose, selon le sociologue Michel Crozier, (1922- ) sur le contrôle d'une zone d'incertitude, (celui qui contrôle cette zone d'incertitude, information, compétence, savoir-faire, détient de ce fait du pouvoir).

Le *rechercheur* d'information se mettrait vite dans l'embarras, s'il partait bille en tête, sans se faire préciser le sens de la question, vers les méthodes d'acteurs, (théâtre) et de comédiens, (C. S. Stanislavski, 1863-1938).

#### 2.3.4.5.1. La réponse experte ou spécialisée

On définira la réponse experte ou spécialisée comme celle qui est produite par une personne (ou un organisme) qualifiée répondant dans un domaine spécialisé de connaissance. Ou bien c'est la formation universitaire assortie des compétences du professionnel, ou bien c'est la pratique habituelle d'un fonds, d'un ou de plusieurs domaines de connaissance, qui font qu'une réponse rendue peut alors être déclarée experte.

C'est ce type de réponse qui demande le plus d'expertise, de temps de recherche et de savoir-faire du *rechercheur*. Il peut s'agir de fournir un détail bibliographique (une date d'édition, la référence d'une illustration, etc.....) ou de répondre à une problématique (ensemble de questions). Ces réponses sont fournies par des structures du type *Sindbad* de la BnF selon un modèle distribué en fonction des domaines de compétence des répondants,<sup>32</sup> ou par les centres de documentations spécialisés tels que par exemple : le Centre de documentation de l'urbanisme<sup>33</sup> (CDU) ou l'Institut national de la santé et de la recherche médicale, (Inserm).<sup>34</sup>

Néanmoins, ce type de réponses élevé se pratique de plus en plus couramment dans l'ensemble des structures informationnelles, comme nous

---

<sup>32</sup> Cf. *infra* Les services questions-réponses.

<sup>33</sup> Centre de documentation de l'urbanisme (CDU) <<http://www.urbanisme.equipement.gouv.fr>>.

<sup>34</sup> Institut national de la santé et de la recherche médicale, (Inserm).  
< <http://www.inserm.fr> >

l'indiquions plus haut en arguant du niveau des questions posées par l'ensemble des utilisateurs.

#### 2.3.4.6. Le niveau de réponse

Proche de la notion de pertinence, le niveau de la réponse : « Il s'agit d'une question épineuse, très débattue. Comment répondre ? S'agit-il de donner uniquement la réponse ou également des informations méthodologiques pour former le public à la recherche documentaire ? », [Galipienso Njongoue, 2007, p.52-53]. Un service de renseignement ne peut-il avoir un rôle pédagogique ? Pour illustrer son propos, Galipienso Njongoue donne un exemple qu'elle a eu à connaître : « Serait-il possible d'obtenir la liste des noms gravés sous les arches de la façade de la bibliothèque [Sainte Geneviève] ? » ; notons que c'est une réponse *localisée*.

« Deux réponses ont été données à intervalles différents. Dans la première, le « répondant » a uniquement communiqué les références du document où se trouve la liste des noms. La deuxième réponse contient en revanche des explications sur cette liste.

« La liste des noms figurant sur les façades de Sainte-Geneviève constitue un élément essentiel de la décoration voulue par l'architecte Henri Labrouste. Elle constitue, selon les propres termes de ce dernier, une sorte de "catalogue monumental" du contenu de la Bibliothèque, offrant à l'extérieur du bâtiment un échantillon de penseurs et de savants présents par leurs œuvres sur les rayonnages de l'intérieur; cette liste, qui s'ouvre avec Moïse et se clôt sur Berzélius (chimiste suédois mort le 7 août 1848) couvre toute l'histoire de l'humanité et illustre l'encyclopédisme des collections. Elle a été dressée par Labrouste lui-même, et la gravure assurée entre le 30 septembre et le 21 novembre 1848 par le sculpteur Johann Deutsch. On consultera avantagement, sur ce sujet, la contribution de Frédéric Barbier parue en 2001 sous le titre Autopsie d'une façade dans « Des palais pour les livres : Labrouste, Sainte-Geneviève et les bibliothèques », BSG / Maisonneuve et Larose, pp. 83-93 (notamment 92-93). »

La BSG sera amenée à choisir entre deux stratégies, [Galipienso Njongoue, 2007, p.53].

#### . 2.3.4.7 La construction de la réponse : éléments de réponses définitives

Pour résumer cette section consacrée à la réponse, nous pouvons observer que la réponse est un *objet* informationnel et un processus réactif construit en plusieurs étapes interdépendantes :

- 1) **processus psycho-cognitif** : écoute et analyse de la question,
- 2) **processus communicationnel** : dialogue, proxémie, empathie,
- 3) **processus documentaire** : recherche de l'information,
- 4) **processus informationnel** : restitution de l'information pertinente et validation ou non, (l'erreur ou l'incertitude font partie de la réponse ; une réponse approximative, un document non pertinent, une information erronée sont des réponses).

Enfin, une réponse est encore un fait culturel construit, et une reconstruction dépendante du cadre mental et des savoirs de son énonciateur.

La coopération et « l'intelligence collective [*en ligne*] : chacun sait quelque chose », concept de Pierre Lévy, ont ici raison des problématiques soulevées par le chantier des questions et des réponses...

Un autre point, peu abordé par la bibliothéconomie, concerne les problématiques de validation de l'information au sein des établissements

### 2.4. La validation de la réponse dans les bibliothèques et centres de documentation

#### 2.4.1. En bibliothèques publiques

En l'absence de regard critique, (les tutelles ne se soucient pas des réponses données), on peut se demander qui, dans ces lieux d'information que sont les bibliothèques et les centres de documentation, valide et décide de la qualité d'une réponse. (Exception des systèmes de questions-réponses qui mettent parfois en ligne leurs résultats consultables). Il serait intéressant d'imaginer, en tenant compte, par exemple, du concept d'intelligence collective de Pierre Lévy, de nouvelles procédures d'évaluation de la pertinence de l'information.

A cette question délicate, on peut avancer que le lieu (prestige de l'institution), et le statut (celui du professionnel) impriment au moins une marque d'autorité intellectuelle quand ils n'impressionnent pas l'interlocuteur. (*A priori* favorables). La source informationnelle est également partie prenante dans le processus de cette validation, puisque c'est le professionnel qui a acquis ce document (la donnée), et son contenu ayant été validé auparavant par un éditeur intellectuel et /ou un éditeur commercial.

#### **2.4.1.1. La réponse immédiate sur les livres ou la réponse médiatisée du bibliothécaire**

Quotidiennement, les bibliothécaires sont amenés à répondre aux demandes des lecteurs sur la présence ou non de tel ou tel « document », de sa date de retour, s'il a été emprunté, etc., La réponse donnée ne souffre pas d'équivoque : « Oui, le titre de l'ouvrage demandé est présent en rayon et donc, il est susceptible d'être emprunté. ». Parfois, il n'est plus disponible ou bien absent des collections et alors : « non, la bibliothèque ne peut le racheter ».

#### **2.4.1.2. La réponse médiatisée par les livres ou la réponse par le livre (document)**

Quelles que soient leur taille, les médiathèques sont également par nature des lieux de dissémination et de médiation de l'information scientifique et technique. Elles offrent, à destination de leurs divers publics, dans leur collection documentaire un arsenal d'ouvrages de vulgarisation scientifique. Parmi ces ouvrages, quelques-uns, qui font florès ces temps ci, abordent avec humour les questions (et les réponses) scientifiques les plus inattendues comme les plus insolites. Dans l'ouvrage (anonyme), *Pourquoi les manchots n'ont pas froid aux pieds ? et 111 autres questions stupides*, paru aux éditions du Seuil en 2007, dans une collection de référence : « *Science ouverte* », qui est une espèce d'anthologie des questions publiées dans la rubrique « *le dernier mot* » du magazine anglais *New Scientist*, la réponse qui est donnée à la question éponyme du titre de l'ouvrage, est rédigée sous forme de validation de contenu, d'abord par un océanographe et ensuite par un universitaire, [p.53-55].

Dans le cas que nous observons, on remarque que la validation de la réponse est le résultat antérieur de plusieurs processus superposés : le processus d'acquisition par l'organisme documentaire et un processus de choix éditorial et commercial, étapes qui s'insèrent dans d'autres dispositifs intellectuels. Un processus de sélection-acquisition de l'ouvrage par un éditeur intellectuel, lui même inscrit dans un processus d'inclusion dans une collection de référence, et

un processus de validation des informations de contenu par les auteurs et le *titre* de la publication.

Dans les deux cas cités, celui du bibliothécaire répondant et du livre « *cor – respondant* » se pose la problématique de la validité de la réponse donnée. Un autre aspect (qui explique le choix de ce livre) est que ce titre met l'accent sur les problématiques des questions apparemment « *bêtes* » et difficiles.

Les questions ne seraient *a priori* jamais difficiles, mais les réponses oui, puisqu'elles commandent, pour le cas que nous observons, des explications validées scientifiquement. Ce qui pose en conséquence les problèmes de la vérité scientifique (critère de réfutabilité). Dans un ouvrage qu'il consacre à la vulgarisation scientifique Lévy-Leblond fait précisément référence à la rubrique « *le dernier mot* » du *New Scientist*, et à ces problématiques de validation scientifique des réponses données. Il écrit : « (...) la validité d'un énoncé n'est jamais donné d'avance et résulte d'un long débat. Le caractère construit et non préétabli de la vérité scientifique, thèse encore parfois controversé, est ainsi démontré par cette initiative journalistique qui se révèle constituer une véritable expérience en sociologie des sciences ». [Lévy-Leblond, 2002, p.30].

#### **2.4.1.3. Dans les centres de documentation ou d'information spécialisés**

Dans un centre de documentation ou d'information spécialisée, une réponse a fonction de prestation. Cette prestation répond à une demande d'information, anticipée ou non. C'est un service, le plus souvent payant, qui peut prendre différentes formes : réponses orales ou écrites, renseignements spécialisés, (juridique, médecine, etc.) notes d'information, dossier documentaire, synthèse d'information, fichier électronique, etc.... Les modalités de la recherche d'information sont rendues opérantes par une organisation de la documentation, (communication) et la bonne connaissance des outils de repérage et des sources d'information.

Dans l'environnement concurrentiel, le centre de documentation répond aux besoins stratégiques internes (formation) et externes (information) de l'entreprise.

Ainsi, l'activité d'un centre de documentation est entièrement tournée vers la réponse et l'analyse des besoins, (enquête de besoin) à terme et à moyen terme. On peut dire que plus un centre de documentation est pertinent, dans ses prestations, et plus il est *répondant* et assuré de sa survie.

#### **2.5. La réponse par le web 2.0\***

Quand on interroge sur le web avec un moteur de recherche, on obtient une liste de réponses.

Quand on interroge « la communauté virtuelle » [Rheingold, cité par Breton, Proulx, 2006, p. 299] : *Digge like*,<sup>35</sup> *Scoopéo*,<sup>36</sup> *Del.icio.us*,<sup>37</sup> on peut espérer se connecter avec un spécialiste d'un domaine de connaissance et par conséquent obtenir une réponse des internautes.

Le web devient le *répondant* du cyberspace par le truchement des internautes hissés au rang d'experts. Les réseaux sociaux, les blogs spécialisés, les alertes

<sup>35</sup> Site web 2.0 de propositions par des internautes de partage de pages web.

<sup>36</sup> Site de *news*.

<sup>37</sup> Service web de partage de liens, encore appelés signets, favoris ou *bookmarks*.

Google, les agrégateurs de flux RSS\*etc., (l'exemple de *Del. icio. us*), fournissent également des aides précieuses dans l'apport de solutions à la réponse.

### **2.5. 1. La réponse selon Google : une expertise virtuose ou virtuelle ?**

Les moteurs présentent des méthodes de tri automatiques de résultats (réponses). L'objectif de ce classement est d'afficher une liste de dix à vingt premiers résultats censés répondre le mieux à la question.

Les concepteurs de Google, le plus connu et le plus utilisé des moteurs de recherche sur le web, poursuivent l'idée de pouvoir gérer intégralement le système complexe des connaissances sur la toile. La fourniture des réponses est donnée comme une vérité objective, alors qu'elles ne « sont rien d'autres que des parcours subjectifs, filtrés par les technologies de recherche, [Ippolita, 2008, p. 195]. Ce sont des réponses paramétrées par des algorithmes. Les temps de réponse de Google sont inférieurs à 3 secondes.

Consultable 24 heures sur 24 et sept jours sur sept, Google est un véritable service de questions-réponses en ligne.

### **2.5.2. Y a-t-il un lien entre une question et une réponse ?**

Un lien logique, sémantique ou phonétique ? Un lien psychologique, (réponse préconstruite, formatée). Un lien invisible, (*arlésien*) ? Un lien aveugle ? Un lien symbolique ? On peut raisonnablement soupçonner qu'il y a des réponses suggérées par le contexte culturel et les habitudes mentales, (cadre de référence), de celui qui reçoit la question, créant ainsi un type de réponses réflexes, (stéréotypes). L'idée de couple question-réponse renforçant l'idée de lien, de liaison. Une réponse n'est-elle pas une *partie séparée d'un tout* d'une question ? Une concaténation ? Si question et réponse sont des objets, existe-t-il pour parler comme Pierre Lévy, une *interface* entre ces objets, voire des *réseaux d'interfaces* [Lévy, 1993, p.199] qui feraient sens ? Ou encore des liens hypertextuels ?

Les linguistes parlent également de « la paire question-réponse », [Baylon, Mignot, 2005 p. 260], lors des interactions verbales, (tours de parole). Certains linguistes, qui analysent les conversations définissent les tours de paroles, comme une « paire adjacente », « contenant deux énoncés contigus, produits par des locuteurs différents [qui] fonctionne[nt] de telle sorte que la production du premier membre de la paire exerce un contrainte sur le tour le suivant, [Traverso, 1999, p.33]. Ces enchaînements d'actions sont repérables lors de nombreuses interactions verbales et notamment dans le couple question réponse : « le principe de la paire adjacente indique que le premier [énoncé] étant produit, le second est attendu », [Traverso, 1999, p.34]. Ainsi, de la paire demande de confirmation (question) et confirmation (réponse) : « Avez-vous ce document ? » –Oui, nous l'avons ! » fonctionne selon un principe de « dépendance conditionnelle ».

Quel lien, autre que la pertinence objective, (qui reposerait sur un calcul d'occurrences), peut-il exister entre les questions et les réponses ? Dans les systèmes de questions-réponses automatisés, leurs concepteurs parlent de « réponses attendues » pour vérifier le degré de pertinence et de validité du système, (stock des questions et stock des réponses).

Le lien le plus évident et perceptible est l'humain. C'est lui qui fait, qui décide qu'une réponse correspond à sa question. Ce lien est aussi qualifié de pertinent, puisqu'il corréle la réponse avec un utilisateur et son besoin informationnel.

Les langages documentaires indexent, connectent des questions avec des réponses. Un support peut également être un lien qui unit et la question et la réponse (cas du dictionnaire et des encyclopédies. Double lien :

Liens linguistique et mathématique : « combien font deux et deux ? Deux et deux font quatre ».

Lien de causalité : Nous sommes tentés d'explorer l'idée « d'attribution causale » de Fritz Heider (1896-1988), idée qui, bien que ne s'appliquant qu'aux personnes, peut être une piste de recherche, dans la mesure où l'attribution causale, considère l'homme comme une machine à inférer.

Raisonnement biaisé, sophisme et lien paradoxal binaire.

« Le sophisme des questions multiples. Inventé par les Mégariques dans l'Antiquité grecque, le plus célèbre est celui-ci : « Avez-vous cessé de battre votre père ? » Si l'on répond oui à cette question on répond implicitement que l'on battait son père. Si bien que celui qui n'a jamais battu son père ne peut répondre (ni par oui ni par non) », [Godin, 2006, p.498].

Le paradoxe autoréférentiel de *Sancho Pança*, tiré du *Don Quichotte* de Cervantès. A l'entrée d'un pays, des soldats demandent à l'étranger : « Pourquoi venez-vous ici ? » S'il dit la vérité, il peut aller librement mais s'il ment, il est aussitôt pendu (un gibet marque l'emplacement de cette inquiétante douane. Un jour, un voyageur plus malin que les autres déclarent aux soldats : « je viens pour être pendu. » S'il dit la vérité, il faut le pendre puisqu'il a dit : « je viens pour être pendu ». Mais ne sont pendus que ceux qui mentent. S'il ment, alors il doit être pendu comme tous ceux qui mentent. Cela dit s'il est pendu, il aura dit la vérité en disant qu'il est venu pour être pendu... [Godin, 2006, p.504].

(Il existe des réponses par habitude : quand on a déjà répondu à la même question).

Existe t il une logique, un enchaînement logique des questions et des réponses ? Questions et réponses connaissent-elles une forme de *rapport symbiotique* ?

Oui : le raisonnement. La logique induit des réponses logiques et elles seules. (Mais il existe plusieurs systèmes de logique), L'inférence induit des réponses programmées par la logique formelle, (Si.... alors...). La Logique : un exemple de raisonnement logique, Cf. Popper, déjà cité, *supra*, le syllogisme.

Les mathématiques. Question : « Deux et deux font-ils quatre ? Font quatre, réponse juste, mais en français ! « Un allemand répondrait non : deux et deux ne *font* pas quatre ; on dit *zwei und zwei ist vier*, deux et deux est quatre. Et en espagnol, *dos y dos son quatro*, deux et deux sont quatre », [Nordon, 1999, p. 8]. Fait avéré, (questions factuelles) : Quel est l'auteur d'*Une Saison en enfer* ? Une date : quelle est la date de la Bataille de Marignan ?

En médecine un syndrome répond à ensemble de symptômes qui permettent d'identifier une maladie, etc.,.....

NB : Faute de temps et de connaissances (linguistique<sup>38</sup> et logique) nous limitons notre recherche à ces rares exemples, qui visent à démontrer qu'il y a un lien logique, lequel enchâsse la question dans une réponse univoque et une seule.

## 2.6. Les services de questions-réponses en ligne

<sup>38</sup>Nous n'avons pu consulter : Coyaud, Maurice. *Linguistique et documentation : les articulations logiques du discours*. Paris : Larousse, 1972. 173p. (Langue et langage).

La définition des services de questions-réponses donnée par le *Dictionnaire de l'information* est relativement incomplète : « Prestation de base d'un centre de documentation ou d'un service d'information d'un organisme, (administration, association ou entreprise). Elle consiste en la fourniture d'informations simples dans des délais généralement courts (adresses, statistiques, références de documents, noms de personnes, etc.....), mais peut inclure aussi une prestation de recherche documentaire. Le service peut être assuré en face à face, par téléphone ou par courrier électronique. (...) », [Calcaly, 2008, p. 210-211]. En effet, au lieu de parler « d'informations simples », n'aurait-il pas mieux valu évoquer les informations « plus faciles » à se procurer ? En outre, cette définition ne permet pas de distinguer « services » et « systèmes » de questions-réponses.

Les premiers logiciels de systèmes de questions-réponses datent des années 1960-1970, [Grau, 2004, p. 80]. Ils furent développés dans un but de recherche d'information ou d'interface en langage naturel.

Les systèmes de questions-réponses en ligne utilisent les technologies du traitement automatisé de la langue. Nous ne traiterons pas dans cette section des systèmes de questions-réponses, (*Sqr*) entièrement automatisés, [Grau, 2004, chap.9, p. 189-218], tel *Qristal*, (*Sqr* de la société *Synapse*). Au croisement de la linguistique et de l'informatique, les *Sqr* anglophones s'affrontent régulièrement dans des campagnes TREC<sup>39</sup> d'évaluation de pertinence et lors de campagnes CLEF<sup>40</sup> pour les *Sqr* francophones, (et des autres langues). Les réponses sont issues de différents *corpus*, (internet) et de bases de connaissances.

L'être humain restant *l'interface* décisionnaire, d'où la distinction entre système, (machine) et service, (humain) de questions-réponses.

Ces services de références, parfois qualifiés de virtuels, sont apparus aux États-Unis dans les années 1999-2000. Ils utilisent des modes soit synchrones, (temps réel) soit asynchrones, (réponses différées).

Ce sont les bibliothèques américaines qui ont initié ce type de service avec le réseau de référence *Question Point*, [Jacquesson, Rivier, 2005, p. 536]. Ces services au public se développent timidement depuis quelques années en France dans de grandes bibliothèques et sont accessibles en ligne sur leurs sites, dès leur page d'accueil, où l'utilisateur dispose d'un formulaire pour poser sa question.

« Contrairement à ce qui existe en Grande-Bretagne ou aux Pays-Bas avec le service de réponses à distance unique « *Ask a librarian* » en France coexistent plusieurs services : « *Sindbad* », géré par la BnF, « [biblioSés@me](#) », géré par la Bpi, « *Le guichet du savoir* », à Lyon... », [Heurtematte, 2007, p. 59]. Le « retard français » est essentiellement dû d'une part, à l'intérêt porté pour le développement et la communication des collections au détriment des autres services orientés usagers et d'autre part, aux réticences d'une grande partie des bibliothécaires qui estiment que ces services relèvent de ceux des documentalistes, et enfin aux refus des édiles, qui ne jugent pas utiles de disposer de tels services, (coûts).

Ces services pourraient faire l'objet à eux seuls d'une étude. Il serait intéressant, par exemple, de poser des questions identiques à chacun des sites, de comparer et d'analyser les réponses fournies, puisqu'il est possible de consulter l'historique des questions. Voire d'analyser les réponses fournies par ces services et d'en évaluer la pertinence. (Copies d'écran datées). Nos propres

---

<sup>39</sup> TREC : *Text Retrieval Conference*.

<sup>40</sup> CLEF: *Cross Language Evaluation Forum*.

interrogations ayant permis d'observer une certaine *hétérogénéité* des réponses produites ...par des professionnels.

### **2.6.1. Les objectifs des services de questions-réponses dans les bibliothèques**

Selon Isabelle Copin, responsable de *Sindbad*, pour la BnF, (Bibliothèque nationale de France), et de Marie Montano, responsable de [biblioSés@me](mailto:biblioSés@me) à la Bpi, (Bibliothèque publique d'information), les services de questions-réponses correspondent à deux enjeux qui en fait s'entremêlent, enjeu interne : renforcer la vocation scientifique des bibliothèques, et enjeu externe : assurer la présence (visibilité) de la bibliothèque à distance.

#### **2.6.1.1. Les services de questions-réponses à distance des bibliothèques**

A distance, signifie que l'utilisateur peut (doit ?) interroger le service de n'importe quel lieu et poste connecté avec la bibliothèque. Néanmoins, Marie Montano et Isabelle Copin font très justement remarquer que les réponses à distance ont toujours existé dans les bibliothèques. C'est le média qui modifie l'interaction, mais non l'action de service. Ces services de haute qualité sont gratuits.

#### **2.6.1.2. Services de questions-réponses aux publics : les exemples de «*Sindbad* de la Bibliothèque nationale de France (BnF) et de [biblioSés@me](mailto:biblioSés@me) de la BPI (Bibliothèque publique d'information du Centre Pompidou), Paris<sup>41</sup>**

##### **2.6.1.2.1. *Sindbad* de la Bibliothèque nationale de France (BnF)**

La BnF a mis en ligne *Sindbad*, acronyme pour : Service d'information des bibliothécaires à distance. Créé en 2005 et d'abord réservé aux seuls bibliothécaires, ce service s'adresse et renseigne depuis 2007 ? toute personne qui lui en fait la demande. Consulter : « Poser une question à un bibliothécaire » sur la page d'accueil du site.

*Sindbad* tire profit de l'expertise des personnels de la BnF et de la richesse des collections qui bénéficient du dépôt légal. Le délai de réponse à une question est porté à trois jours, (Charte qualité). Le temps de traitement des questions par le personnel ne doit pas dépasser plus de trois heures. Les questions sont d'ordre factuel ou bibliographique. 14 départements de la BnF collaborent, dont parmi eux, environ 80 **répondants** (par internet), spécialisés dans un domaine. Les questions sont également traitées avec le logiciel *Question point*

##### **2.6.1.2.1.1. Règles et pratiques d'uniformisation et éléments d'activité du service : Indication quantitatives sommaires**

Adopter des pratiques communes afin d'agir en cohérence : règles de réponses (demandes de précision de la question, difficultés liées à l'absence de

---

<sup>41</sup> Source : Journée d'information du secteur éducation de l'ADBS, « *Les services questions réponses à distance* », Boulogne-Billancourt, le 20 février 2008. Exposé de Mesdames Marie Montano, (Bpi) responsable de [biblioSés@me](mailto:biblioSés@me) et d'Isabelle Copin, (BnF) responsable de *Sindbad*.

dialogue...). Charte de définition du service rendu (disponible sur le site). Bibliothèque de réponses *types*.

900 questions reçues par mois en moyenne. 57% des questions sont reçues par internet, 46% par téléphone et 1% par courrier. 6% des réponses ont trait à la philosophie, 15% au droit, à l'économie et à la politique, 14% aux manuscrits, etc. .... Beaucoup de questions viennent de l'étranger. Un projet de coopération avec [biblioSés@me](mailto:biblioSés@me) est à l'étude, ainsi qu'une extension à la francophonie.

#### **2.6.1.2.2. Le service questions-réponses [biblioSés@me](mailto:biblioSés@me) de la Bpi (Centre Pompidou)**

Ce sont 9 bibliothèques françaises, augmentées d'un partenariat avec une bibliothèque allemande, la Zentral-Landesbibliothek de Berlin, qui constituent le réseau de [biblioSés@me](mailto:biblioSés@me). 3 bibliothèques municipales et une bibliothèque départementale de prêt devraient intégrer ce réseau en 2008. Une charte de qualité fixe le délai de réponse maximum à 72 heures. Le logiciel *Question point* a permis de traiter en 2007 et sur douze mois : 2012 questions soumises au réseau<sup>42</sup>.

Marie Montano a rappelé au cours de son intervention, l'existence de « *Bpi tuyaux* », service de questions-réponses par téléphone, (accessible sur le serveur du journal *Libération*) et courrier des années 1980-1990, qui préfigurait celui de [biblioSés@me](mailto:biblioSés@me). La Bpi a adopté le même type de prestations et la même charte qualité que *Sindbad* de la BnF.

#### Coopération et résolution de l'*énigme* :

13 personnes, (modèle distribué en fonction des domaines de compétences) s'occupent du traitement des questions. Délais de réponses sous trois jours, aucune question n'est laissée sans réponse. Quand une question n'est pas résolue par le service, elle est renvoyée (réorientée) vers la BnF. Cette réponse donnée pourra, en retour, être lue par la Bpi. Les questions restées définitivement sans réponses sont alors envoyées sur le réseau mondial de *Question point* (réseau global) qui dispose d'une base de connaissance de réponses données. On peut également consulter sur le site la base de connaissance, (réservoir de questions-réponses) de [biblioSés@me](mailto:biblioSés@me) qui contient une sélection et un historique des questions traitées (réponses).

Marie Montano considère qu'il n'y a pas de questions difficiles et que les professionnels, pour s'assurer de la qualité de la réponse doivent : bien lire la question, la comprendre, la traiter, tenir compte du temps passé.

#### **2.6.1.2.3. Le *Guichet du savoir* de la Bibliothèque municipale de Lyon**

Seule grande bibliothèque publique municipale de prêt, (la BnF et la Bpi ne prêtent pas leurs documents), la bibliothèque municipale de Lyon a démarré en 2004, un service en ligne de questions réponses : « *Le Guichet du savoir* »<sup>43</sup>. Cinq

<sup>42</sup> Source : *Bibliothèque(s)*, déc. 2007, n°36 p.28.

<sup>43</sup> <<http://www.guichetdusavoir.org/>>.

règles précisent aux utilisateurs les modalités de ces systèmes : non discrimination du public et des questions, fourniture d'une réponse précise et argumentée, garantie d'une réponse sous 72 heures, publications de l'ensemble des questions et des réponses apportées, participation de l'ensemble des bibliothécaires à ce service.

A près un an d'exercice, le service avait répondu à 4 800 questions et depuis mars 2005, ce sont environ une vingtaine de questions/jour qui arrivent au *guichetdusavoir*. [Di Pietro, Calenge, 2005, p. 38-42].

Voir annexe n°2 : exemple de traitement d'une réponse par *Le Guichet du savoir*

#### **2.6.1.2.4. « Questions ? Réponses !<sup>44</sup> » le service de l'Enssib**

« *Questions ? Réponses !* » de l'Ecole nationale des sciences de l'information (Enssib), né en février 2008 est le dernier-né des services d'information à distance à destination du public. Un projet relativement avancé est en cours d'élaboration à la médiathèque de la Cité des Sciences de la Villette.

On peut augurer que les « *systèmes ou services* » de questions-réponses en ligne, corollaire du développement des technologies du traitement automatisé de la langue (TAL), sont appelés à s'implanter dans de nombreuses structures documentaires en raison de leur facilité d'emploi et de la précision toujours plus grande des réponses qu'ils proposent. Rien n'indique qu'ils devraient rester dans un proche avenir, un outil à la seule disposition des médiateurs professionnels et que le grand public ne finisse par s'en emparer, reléguant peut-être définitivement le bibliothécaire de référence au magasin de l'histoire du livre.

Pour résumer cette section, nous soulignerons que le principal avantage des services de questions-réponses en ligne tient au fait qu'il réintroduit et insère *visiblement* et intellectuellement la bibliothèque dans les dispositifs technologiques des nouvelles techniques de communication et d'information : « les réponses prenant le statut de document sinon de documentation », [Nguyen, 2006, p.68]. D'autres dispositifs de questions et de réponses existent, notamment ceux du web 2.0\* collaboratif, qui font appel aux internautes.

#### **2.6.1.2.5. La participation des internautes : Les *Faq* (*Frequently Asked Questions*) ou « *Foire aux questions* »**

« Issus des forums de discussions, les sites porteurs de *Faq* font appel aux internautes qui échangent questions et réponses dans un esprit de participation collaborative. De nombreuses questions trouvent ainsi leurs réponses fournies par la communauté élevée au rang d'expert ». [Foenix-Rioux, 2005, p.119].

Si le sérieux des *Faq* peut être mis en question parfois, il existe des outils professionnels telle la liste de diffusion de l'ADBS et celle de « *Biblio.fr* » qui peuvent être mises à contribution pour obtenir des informations.

---

<sup>44</sup> < <http://www.enssib.fr/questions-reponses/>>. Consulté le 23/08/2008.

Néanmoins, le parachèvement de la construction de *l'objet* réponse nécessite l'abord d'un concept très prégnant en documentation et à dimensions objectives et subjectives variables : la pertinence

## **2.7. De la question à la réponse : la pertinence comme réponse documentaire évaluée**

« (...), La pertinence est l'aspect le plus abordé des études de la recherche d'informations, et il n'existe aucune sorte de consensus en la matière », [Mkadmi, Saleh, 2008, p. 12].

### **2.7.1. Le triptyque pertinence- bruit - silence**

Répondre à une question d'un usager revient à communiquer à cet usager une information ou un document pertinents, c'est-à-dire correspondant à sa demande. En l'occurrence, il s'agit dévaluer la réponse fournie, soit par le logiciel, soit par le *rechercheur*, soit par l'usager, le plus souvent en interrelation circulaire. Cette évaluation reposant sur plusieurs critères, plus ou moins subjectifs ou objectifs. Par conséquent, il convient pour le bibliothécaire ou le documentaliste de fournir la réponse la plus complète possible, c'est-à-dire sans silence documentaire, (documents pertinents non retrouvés), en même temps que cette réponse sera la plus adaptée à la demande exprimée, c'est-à-dire également sans bruit documentaire, (documents retrouvés non pertinents). La pertinence s'apprécie encore par la qualité et la précision d'un descripteur, par la qualité d'une indexation d'un document. Enfin, la pertinence se mesure également mathématiquement, elle est surtout utilisée par des logiciels pour ordonner et classer les documents ramenés sous forme de tris en réponse à une requête.

#### **2.7.1.1. Pertinence : éléments définitionnels**

Le *Tlfi : Trésor de la langue française informatisé*, [en ligne] donne en première acception au terme pertinence, la définition suivante : « Qualité de ce qui est adapté exactement à l'objet dont il s'agit. *Synon.*, bien-fondé ». Le mot est attesté dès 1580.

Pour les auteurs de l'article *pertinent* in le *Grand dictionnaire de la psychologie*, [1991 p.568-569], l'adjectif pertinent, « qui convient exactement à son objet », dont l'équivalent anglais est *relevant*, « le sens du mot exigerait qu'il soit toujours utilisé avec un complément « pertinent par rapport à » : Cet article est plus précis puisqu'il relie ce syntagme « par rapport à un but, à l'intérieur d'une tâche ou d'une situation de problème.(...)».

La tâche implique l'accomplissement d'un comportement ou d'une suite de comportements, en vue d'atteindre un but dans une situation. En synthétisant cette définition, on peut dire que la pertinence est à la fois un processus (comportement) et un résultat (but). En l'occurrence, le processus est la résolution du problème générée par la recherche documentaire et les moyens mis en œuvre jusqu'à « sa découverte par essais et erreurs, par raisonnement, par insight ou par tout autre moyen ».

Il ne semble pas utile de multiplier les définitions, mais il est intéressant de rappeler que les mots « congru » et « congruence » sont des termes voisins de pertinence.

#### **2.7.1.1.2. Le concept de pertinence en documentation et ailleurs**

« Le concept de pertinence connaît plusieurs sens selon les points de vue envisagés : philosophique, logique, sémantique, pragmatique, etc. ». [Saracevic, 1970, cité par Le Coadic, 1997, p.56].

C'est Bradford, qui le premier a évoqué ce concept dans le contexte documentaire en 1948, dans son livre *Documentation*, en évoquant des « articles pertinents sur un sujet ».

Dans son ouvrage *Terminologie et documentation*, Maryvonne Holzem relève que « le concept de pertinence est opératoire chez les linguistes et cite « André Martinet [qui] appelait en 1960 « traits pertinents » les traits qui permettent de différencier les unités phonétiques au sein d'une langue », [Holzem, 1999, p.68], c'est-à-dire « dont la présence ou l'absence entraîne un changement de sens de l'unité », [*Dictionnaire historique de la langue française*, 1992, p. 1490]. On reconnaît ici la « perturbation homéostatique » dont nous parlions, dans notre seconde hypothèse.

Pour ce linguiste, la pertinence est affaire de description et de sélection dépendant de l'observateur : « Toute description suppose une sélection (...). Toute description sera acceptable à condition qu'elle soit cohérente, c'est à dire qu'elle soit faite d'un **point de vue déterminé**. [Souligné dans le texte]. Une fois ce point de vue adopté, certains traits dits pertinents, sont à retenir », [Martinet, 1991, p.31]. En extrapolant, on peut assigner l'observateur sur les traces de l'utilisateur ou du *chercheur* d'information, en quête eux aussi de pertinence plausible, observable, ou mesurable.

La pertinence est-elle une mesure ? Auquel cas, elle servirait à connaître. Car mesurer, c'est connaître. Mais les distances mesurées par les statistiques sont elles pertinentes ? Les calculs fondés sur la présence ou non des occurrences d'un texte sont-ils toujours judicieux ? On peut aussi imaginer, à la manière de G. Perec pour son roman *La disparition*, un texte sur la mer qui ne comprendrait aucune fois ce terme et qu'aucune recherche qualifiée ne remonterait dans ses filets !

Le neurobiologiste Jean-Pierre Changeux retient ce concept de pertinence. Il s'agit selon ce chercheur, « d'établir un lien causal pertinent entre structure et fonction [du cerveau] et [le monde] », [Changeux, 2004, p. 50]. En effet, et nous donnons ici un résumé de sa thèse : on ne peut comprendre comment fonctionne chaque partie du cerveau sans analyser et sans tenir compte des relations qui interagissent entre ces parties cérébrales et les individus. Il s'agit d'établir « une analogie possible entre les processus de communication intervenant entre organismes et ceux existant à l'intérieur de l'organisme (...). Cette interaction contextuelle nouvelle [la pertinence] et les effets amplificateurs induits pourraient expliquer les évolutions du cerveau, [Changeux, 2004, p. 50-52]. Changeux fait de la pertinence un lien de causalité biologique, qu'il pose comme hypothèse théorique. Il précise plus loin ce concept : « un changement *quantitatif* mineur concernant un seul composant chimique du cerveau suffit pour modifier *qualitativement* le comportement de l'organisme ». Il estime que ce concept de pertinence pourrait s'appliquer également à la « redoutable complexité du développement et de la différenciation de « notre appareil de connaissance » et pourrait nous aider à comprendre par exemple, l'apparition du phénomène de la dyslexie, [Changeux, 2004, p.252]<sup>45</sup>.

---

<sup>45</sup> Changeux et Pierre Lévy citent tous deux : l'ouvrage de Dan Sperber et Deirdre Wilson, *La pertinence*, Minuit, 1989 que nous n'avons pu consulter.

### 2.7.1.1.3. Pertinence et documentation

La réponse fournie est-elle toujours *vraiment* pertinente ou bien ne contient-elle pas un halo d'imprécision et d'indécision inévitables. En un mot, la pertinence est-elle de ce monde... de la documentation ? Les sciences de l'information retiennent actuellement « trois grandes catégories de pertinence : [la] pertinence système, [qui] concerne la capacité du logiciel documentaire, (...), la pertinence-utilisateur, (...), et la pertinence-thème, (...), Il s'agit de la pertinence du contenu, [du document], [Dalbin, 2002, p.46-50], [Mkadmi, Saleh, 2008, p.13].

Par conséquent, y a-t-il des réponses *vraies* ou plus *ajustées* à la demande les unes que les autres? En adéquation correcte ou exacte avec les questions ? Celles que soulèvent les problématiques concernant le concept de pertinence. Ce critère de pertinence, plus ou moins précisément défini par la norme ISO 11620, (il s'agit du degré de corrélation entre la réponse et la question posée). En fait, la pertinence « absolue n'existe pas ou du moins, est invérifiable car elle impliquerait que l'utilisateur sache parfaitement ce qu'il cherche au moment où il pose sa question et qu'il connaisse par avance la totalité des informations qui répondent à sa question », [Calcaly, 2008, p.202].

Ce critère de pertinence est-il encore judicieux et les professionnels de l'information ne sont-ils pas à la recherche d'une nouvelle congruence ? D'autant plus que c'est l'algorithmique qui fondera, au fur et à mesure des perfectionnements inéluctables du Web sémantique<sup>46</sup>, « la mesure de toutes choses ».

Selon Jean-Yves Prax et Simon Larcher, le critère de pertinence n'est plus valable. Dans leur ouvrage, *La gestion électronique documentaire*, les auteurs vont jusqu'à écrire, « qu'un agent de veille [documentaire] qui proposerait tous les jours à l'utilisateur la même séquence d'article serait inutile. Le seul critère de pertinence ne suffit donc plus et présente le risque d'un appauvrissement progressif du sujet ». [Prax, Larcher, 2004, p. 292]. En effet, la *répétabilité* d'une même information touche à sa valeur informationnelle et entraîne *ipso facto* sa démonétisation. C'est sa rareté qui fait le *prix* de l'information. Cette situation nouvelle obligera les professionnels de l'information à encore ajouter de la valeur informationnelle, (et intellectuelle), aux réponses fournies et à faire preuve de beaucoup d'esprit créatif. Enfin, ne faudrait-il pas accoler un critère de « validation critique » de l'information au concept de pertinence ? Que doit-on penser, par exemple, des conséquences des déclarations de certains scientifiques de renom, tel J. Watson, Prix Nobel de médecine et codécouvreur de L'ADN, dans les colonnes du *Sunday Times*<sup>47</sup> pour qui « les Africains sont moins intelligents que les Occidentaux » (*sic*) ?

### 2.7.1.1.4. Toute petite histoire de la pertinence documentaire

La masse d'information qui circule sur le web, et c'est un truisme que nous réitérons, pose également « le problème de l'accès à l'information pertinente : il est

<sup>46</sup> « (...) objectifs [du web sémantique] « faire comprendre » par les ordinateurs les informations en langage naturel existant sur internet et de transmettre des réponses adéquates » [Rivier, 2007, p.42].

<sup>47</sup> Vaulerin, Arnaud. Le prix Nobel James Watson, spécialiste de l'Adn..., *Libération*, 17 octobre 2007. [En Ligne] Disponible sur le world wide web :< <http://liberation.fr/actualite/societe/285446.FR.php?rss=true>> Consulté le 23/08/2008.

nécessaire de mieux cibler les résultats en fonction des besoins des utilisateurs afin d'effectuer un filtrage des ressources», [Pédauque, 2007, p. 111].

Pourtant, les problèmes liés à la recherche sur la pertinence et à l'information pertinente ont préoccupé, au moins depuis les années 50, les professionnels de l'information

Des documentalistes, soucieux d'informations adéquates et ajustées en fonction des utilisateurs, mettront au point de nouvelles méthodes « d'indexation qui utilisent le principe d'intersection entre les termes - soit à l'indexation (pré-coordination), soit à la recherche (post coordination)-(...) », [Chaumier, 2003, p. 32]. En développant ainsi des langages à structure combinatoire liés à la notion d'intersection, « cette notion d'intersection utilisée à propos des descripteurs facilite la pertinence de la recherche en limitant les mauvaises interprétations et le « bruit » », [Chaumier, 2003, p.33], les documentalistes ouvriront en même temps la voie aux thésaurus\*. (Cf. Première partie 1.8.1. ).

Les années 60 et 70 seront l'âge d'or des thésaurus\*. L'élaboration des thésaurus\* passe par la construction de tableaux thématiques dans lesquels les termes (descripteurs) sont reliés entre eux au moyen de flèches. Le but de ces représentations graphique est de caractériser les relations associatives et hiérarchiques au détriment des relations de synonymie afin d'une part, d'indexer des documents et d'autre part, de pouvoir retrouver ces documents dûment indexés. Ces représentations visuelles ont l'avantage de permettre d'embrasser le champ sémantique d'un sujet et « de déterminer les termes les plus pertinents lors d'une recherche ou d'une indexation ». [Chaumier, 2003, p.36].

L'introduction de la syntaxe au début des années 60, dans de nouveaux langages documentaires, (*Syntol*)\* pour remédier à l'imprécision parfois ponctuelle des thésaurus, préfigure les premiers travaux d'indexation automatisée.

A ces travaux précurseurs viendront s'ajouter les réflexions de Hans Peter Luhn (société IBM), sur les méthodes statistiques pour qui : « La fréquence d'un mot dans un article fournit une mesure utile de la signifiante de ce mot [...]. La co- occurrence relative, dans une phrase, de mots auxquels ont été affectés des poids de signifiante est une mesure utile de la signifiante de la phrase [...]. Plus certains mots sont trouvés souvent en compagnie les uns des autres dans une phrase, plus on peut dire que ces mots sont lourds de sens », [Hans Peter Luhn, cité par Chaumier, 2003, p. 48]. On le voit, la *signifiante* est une manière d'approcher davantage encore le critère de pertinence par le biais de calculs statistiques. Ces progrès seront rendus possibles par la puissance des systèmes informatiques.

Le développement de l'informatique procurera aux professionnels de l'information la possibilité de réaliser des recherches en texte intégral. Différentes méthodes d'indexation-recherche seront alors proposées dans les années 70. Ce seront tour à tour, les méthodes par assignations, les méthodes syntaxiques et les méthodes linguistiques. Ces dernières aboutiront « aux logiciels d'indexation et de recherche en langage naturel », [Des années 90], [Chaumier, 2003, p. 51-52]. Les années suivantes seront celles des moteurs de recherche et du calcul de pertinence, nommé « *Ranking* » en anglais.

#### **2.7.1.1.5. L'usage contemporain distingue deux types de pertinence :**

Une pertinence **objective**, c'est celle des systèmes et une pertinence **subjective**, (encore appelée pertinence psychologique), c'est celle du point de vue de l'utilisateur. Cette dernière devrait primer sur les autres puisque, en fin de compte, c'est à l'utilisateur qu'il revient d'évaluer l'intérêt et l'utilité de son information ou du document qui lui est proposé. Néanmoins, à ce stade, le professionnel de l'information, parce qu'il gère par habitude de l'information, a son mot à donner en forme de conseil d'expertise dans cette perception de la pertinence.

#### 2.7.1.1.6. La pertinence : une variable des réponses possibles ?

« L'homme est la mesure de toutes choses ». Cette déclaration attribuée à Protagoras, expose le savoir à une indétermination absolue.

Autant d'hommes autant de conceptions de la pertinence. C'est donc une variable.

La pertinence de l'information est l'adéquation de l'information délivrée aux besoins et capacités cognitives les plus *exactement* repérés d'un utilisateur.

Un collégien ou un lycéen recherchant des informations synthétiques sur les arbres producteurs d'huile végétale en France, trouvera pertinent un documentaire dans une édition *jeunesse* sur l'olivier, ou dans un guide spécialisé sur les arbres de l'éditeur Larousse<sup>48</sup>, par exemple, mais non pertinente, une thèse de botanique sur les *Olea europea* L.<sup>49</sup> et cependant, tous ces documents sont d'égale qualité informationnelle et documentaire.

Le fournisseur d'information doit être capable de se mettre à la place de l'utilisateur. Parfois, il doit faire coïncider besoin d'information avec *son* fonds d'information lacunaire, et proposer, faute de mieux, des documents « *approchants* » ceux des besoins de l'utilisateur.

#### 2.7.1.2. Efficacité de la recherche des moteurs et pertinence des réponses

Les moteurs de recherche procèdent, pour leurs classements, par des méthodes de tris automatiques de résultats. « Le but du classement est d'afficher dans les dix à vingt premières réponses les documents répondant le mieux à la question ». [Lardy, 2001, p. 55].

##### 2.7.1.2.1. Le tri par pertinence

« Les résultats de requête sont affichés selon un ordre déterminé par le calcul d'un score pour chaque réponse. La pertinence est basée sur les cinq facteurs suivants appliqués aux termes de la question.

**Le poids d'un mot dans un document** est déterminé par sa place dans le document. Il est maximum pour le titre et le début du texte (...);

**La densité** est basée sur la fréquence d'occurrence dans un document par rapport à la taille. Si deux documents contiennent le même nombre d'occurrences, le document le plus petit sera favorisé ;

---

<sup>48</sup> Brosse, Jacques. *Larousse des arbres et des arbustes*. Paris : Larousse, 2000. 576 p. ISBN 2-03-514302 .

<sup>49</sup> *Olea europea* L. : Nom botanique de l'olivier d'Europe. (*Olea europea sativa* : nom de l'espèce cultivée).

**Le poids d'un mot dans la base** est fondé sur la fréquence d'occurrence pour toute la base de données. Les mots peu fréquents dans le corpus sont favorisés. (...);

**La correspondance d'expression** est basée sur la similarité entre l'expression de la question et l'expression correspondante dans un document (...);

La relation de proximité est basée sur la proximité des termes de la question entre dans le document. Les termes proches sont favorisés ». [Lardy, 2001, p. 56].

#### **2.7.1.2.2. Le tri par popularité : l'exemple de Google**

Cette méthode est basée sur la co-citation, elle-même fondée sur le PageRank qui utilise le nombre de liens pointant sur les pages.

#### **Google Inc : un succès économique très rentable**

Google Inc est une société privée de droit américain, fondée en 1998 et cotée en bourse depuis 2004, qui génère plusieurs milliards de dollars de profits.

Un succès fondé sur une habile utilisation de la ressource publicitaire.

Cette société a produit un moteur de recherche performant basé sur un algorithme appelé *PageRank*, parce qu'il donne un rang aux pages web indexées par le moteur de recherche. Le classement de Google est le *PageRank*,

Google ne mélange pas la recherche et la publicité. Cette start-up s'enrichit pourtant en faisant commerce des mots-clés, les *Ads*, que l'annonceur achète. L'annonce publicitaire, liée avec le site apparaît à chaque requête du mot clé. Il existe ainsi une véritable bourse aux enchères des mots-clés. (L'annonceur peut également payer au coup par clic). Un système identique *Google AdSense*, est créé pour les éditeurs de sites Web. Ainsi, le référencement payant garantit une bonne place dans la liste des réponses.

#### **2.7.1.2.3. Le mode de calcul de la pertinence selon Google**

Olivier Andrieu, un des meilleurs spécialistes des moteurs de recherche et qui anime le site d'information *Abondance*<sup>50</sup>, explique le fonctionnement du plus connu et utilisé des moteurs de recherche actuel sur le Web dans un ouvrage grand public, auquel nous nous référons : « L'internaute saisit un mot clé, le moteur recherche les occurrences dans son index, (pages web sauvegardées). Une fois le « lot » de pages contenant le terme demandé identifié, le moteur classe les pages par ordre de pertinence, selon un ordre et un algorithme mathématique (basé sur certains critères de tri très précis) qui lui est spécifique. [Andrieu, 2004, p.13-15].

C'est ce tri qui est appelé *PageRank*. Il est fondé sur « l'indice de popularité » qui « est basé sur l'analyse des liens pointant vers des pages web ».

« L'indice de popularité d'une page est dépendant de la quantité et de la qualité des pages qui ont mis en place un lien vers elle. Plus cet indice est élevé, meilleur est le classement de la page en question dans les résultats du moteur ». [Andrieu, 2004, p. 15].

---

<sup>50</sup> <<http://abondance.com>>.

Google a amélioré son système de tri de résultats à l'affichage, « (...) en prenant en compte non seulement le nombre de liens pointant vers une page mais également l'indice de popularité des documents qui les affichent », [Andrieu, 2004, p.15].

#### 2.7.1.2.4. Critères de pertinence ou impertinence des critères ?

Les concepteurs et les soutiens de la technologie de Google affirment que les résultats sont corrects parce qu'ils expriment une démocratie directe. Pourtant, « la popularité ne peut pas être admise comme un critère de qualité objective. ». [Ippolita, 2008, p. 180]. Ce qui est en cause ici, c'est le principe du lien/vote. Le *ranking* se faisant l'expression de la qualité. Autrement dit, la quantité [d'information] devenant la qualité : ce qui manque de ...pertinence.

La pertinence comme valeur politique. Il semble bien que la pertinence, selon Google, soit un critère politique, puisque le calcul se fait sur un mode dit « démocratique », c'est-à-dire que plus un site est visité, meilleur sera son classement.

#### 2.7.1.2.5. La pertinence objective : les mesures de la pertinence

Plusieurs modes de calculs de pertinence, des plus simples au plus complexes existent. Aucun n'est décisif, au sens où il déclarerait irréfutablement la pertinence (*justesse, adéquation, vérité,...*), de l'information.

Ainsi les « méthodes de co occurrence qui utilisent les techniques de statistiques afin de classer les réponses par ordre de pertinence présentent en premier les documents considérés comme répondant le mieux (ou le plus ?) à la question » [Waller, Masse, 1999, p.113].

#### 2.7.1.2.6. Les mesures mathématiques simples de la pertinence documentaire

« Taux de rappel, de précision, de silence et de bruit [Rivier, 2007, p. 45-46]

**Taux de rappel** =  $\frac{\text{nombre de documents pertinents retrouvés}}{\text{Nombre total de documents pertinents (cible)}}$

**Taux de silence** =  $\frac{\text{nombre de documents pertinents non retrouvés}}{\text{nombre total de documents pertinents (cible)}}$

**Taux de précision ou de pertinence** =  $\frac{\text{nombre de documents pertinents retrouvés}}{\text{Nombre total de documents retrouvés (résultat)}}$

**Taux de bruit** =  $\frac{\text{nombre de documents non pertinents retrouvés}}{\text{Nombre total de documents retrouvés}}$

#### 2.7.1.2.6.1. Les mesures (calculs) complexes de la pertinence documentaire

Nous ne donnerons pas l'ensemble des calculs mais rappellerons quelques une des techniques utilisées en informatique documentaire. Parmi les méthodes statistiques les plus employées on retrouve : le modèle booléen pondéré\*, le modèle vectoriel\*, le modèle probabiliste\* etc..... (Cf. glossaire annexe 5).

Nous renvoyons à [Lefèvre, 2000, p. 163-179], pour plus d'explications. L'intérêt de ce rappel est de montrer que la pertinence mathématique vise à la ...

pertinence absolue et que les méthodes de calculs statistiques sont nombreuses et variées.

### Pertinence statistique et coefficient de similarité

La pertinence est conçue comme une fonction numérique qui « peut prendre des valeurs comprises entre une borne inférieure (0 par exemple) pour un document pas du tout pertinent, et une borne supérieure, qui indiquerait un document tout à fait pertinent. Pour les moteurs existants cette fonction numérique est obtenue comme résultat d'une *mesure de ressemblance*, ou de *distance*», entre la requête et les documents. Elle est aussi appelée *coefficient de similarité*. Cette mesure est effectuée non pas à partir des documents eux-mêmes, mais de leur représentation condensée, c'est à dire des index. En fonction du type d'index choisi, les résultats seront évidemment très différents ». [Lefèvre, 2000, p.164].

Sept critères permettent de juger de la plus ou moins grande pertinence, qui peuvent être associés de différentes manières.

« 1 correspondance booléenne (critère d'exclusion). Le premier critère traduit simplement *la condition logique de sélection* : pour qu'un document soit ramené, il faut qu'il satisfasse aux conditions de la requête (il contient tout ou partie des termes de la requête. (...).

2 Complétude (par rapport aux termes de la requête). Dans le cas de requête associant plusieurs termes par un OU *la complétude* ou *couverture* indique le nombre de termes de la requête retrouvés dans le document. Cela signifie que si une requête contient par exemple 5 mots associés par un OU, le logiciel placera en tête les documents qui contiennent les 5 mots. (...)

3 Termes discriminants (fréquence dans le corpus).

Les termes *discriminants* sont les termes peu fréquents dans le corpus. (...). Les termes fréquents dans le corpus ont un poids faible ; les termes plus rares dans le corpus, un poids élevé.

4 Fréquence ou densité dans le document.

La fréquence mesure le nombre d'occurrences d'un terme dans le document. (...) Un document sera considéré comme d'autant plus pertinent qu'il contient beaucoup de phrases traitant u sujet correspondant à la question (...).

5 Position dans le document

Dans le cas des documents structurés, il est possible de pondérer les termes en fonction des *zones* où ils apparaissent (tête du document ou zones significatives), (...).

6 Proximité et ordre dans le document

Ce critère prend en compte la notion de *termes composés*, ou la *proximité* dans le document des termes retrouvés. (...).

7 Fréquence sémantique ou contextuelle

Critère parfois utilisé dans les logiciels qui travaillent à un *niveau sémantique*. La notion de *fréquence* est appliquée ici non plus aux mots de la requête mais aux concepts qui leur sont associés. (...).» [Lefèvre, 2000, p. 164-165].

« Les concepts de bruit et de silence dans les systèmes documentaires son fonction d'un autre concept celui de la pertinence lors de l'indexation des documents (...) » [Holzem, 1999, p66].

Plusieurs études ont été réalisées pour évaluer la pertinence (réponse) des moteurs de recherche, « Le modèle *PageRank* n'apporte pas une réponse significativement meilleure que des stratégies ignorant les hyperliens », [Savoy, 2006, p. 49-59].

Cette interdépendance de la pertinence avec les concepts de bruit et de silence qui sont également des concepts subjectifs même s'il sont statistiquement mesurables, reste fortement corrélée et dépendante de l'interprétation humaine (évaluation) des résultats obtenus lors d'une interrogation.

## 2.8. Le monde du bruit et du silence

L'univers du *rechercheur* d'information est aussi celui du monde du bruit\* et du silence\*. Ces deux phénomènes sont causes de nombreux échecs documentaires lors des opérations de recherche d'information.

### **2.8.1. Le bruit documentaire**

Plusieurs définitions du « bruit\* » documentaire coexistent également au sein du monde de la documentation. Toutes tirent leur origine et...une certaine scientificité de la prestigieuse *Théorie mathématique de l'information* de Claude Elwood Shannon.

Le sens premier et d'usage courant du mot bruit, selon le *Trésor de la langue française*, [en ligne] est défini comme : « Ensemble de sons, d'intensité variable, dépourvus d'harmonie, résultant de vibrations irrégulières ».

Mais c'est le cybernéticien Norbert Wiener qui a promu ce terme au sens duquel se sont emparées les Sciences de l'information.

Le « bruit », (Wiener) dont le terme est emprunté aux télécommunications correspond à tout ce qui altère le message dans le procès de communication selon la théorie de l'information. [Baylon, Mignot, 2005, p. 41].

Pour la *Théorie de l'information* [Shannon], le « bruit est le nom donné à la modification d'un message entre son émission et sa réception : lorsque le canal qu'on utilise pour transformer un message est bruyant, le récepteur qui reçoit, par exemple, un 0 ne peut conclure à l'envoi d'un 0 avec certitude, mais seulement avec une certaine probabilité ». [Dion, 1997, p. 77].

Une acception plus contemporaine, qui relève de la linguistique et de l'informatique, est également attestée dans un alinéa du *Trésor de la langue française* [en ligne]: « Tout ce qui altère ou perturbe la transmission d'un message ».

Selon Noé Richter, *op.cit.*p.153, le bruit documentaire se définit comme : « Référence non pertinente fournie par un catalogue interrogé et ne correspondant pas à la question posée. Les causes du bruit sont multiples : analyse insuffisante, polysémie des descripteurs, déficience syntaxique du système d'indexation. Voir *homographie, Polysémie, Silence* ».

Le bruit\* représente l'ensemble des documents non pertinents ramenés ou retrouvés lors d'une interrogation

#### **2.8.1.1. Le bruit et la recherche plein texte ou texte intégral**

Les systèmes de recherche d'information qui font de la recherche d'information en texte plein (ou texte intégral), opèrent sur des chaînes de caractères et non sur des mots. Les textes des documents sont transformés en index et toute la recherche automatisée consiste à repérer, par les entrées de ces index, les documents ou parties de documents, dans lesquels se trouvent les formes, (caractères des lettres), recherchées.

Les principaux moteurs de recherche sur le web utilisent les technologies en texte intégral.

Les causes de bruit générées par ces systèmes sont nombreuses et nous n'en donnerons que les plus importantes : « Le bruit est dû en grande partie à l'homographie – deux formes différentes qui ici s'écrivent de façon identique- trouvent moult façons de résonner.

Homographies entre mots pleins d'une même langue : *convient* : *il convient qu'elles les convient à parler*, mais *le maçon de Mâcon* devient *le macon de macon* ; homographies entre mots pleins et mots vides d'une même langue : *or, car* : *or, il ne prenait pas le car, car il avait un matelas d'or* ; homographies entre mots pleins d'une langue et mots vides d'une autre : *but* (français) et *but* (anglais), *thé* (français) réduit à *the* (anglais). Il est donc risqué de rechercher les textes évoquant un *but en or* ; homographies suite à la suppression de mots vides : *manche gigot, manche à gigot, manche de gigot*. [Lallich-Boidin, Maret, 2005, p. 34].

### **2.8.1.2. Indexation libre et bruit**

« Les moteurs de recherche ayant focalisé l'attention des internautes sur le bruit. ». [Lallich-Boidin, Maret, 2005, p. 119].

L'indexation libre est sensible à l'évolution de la langue

### **2.8.1.3. Recherche d'information par des noms propres et bruit**

Les noms propres sont une des clés d'accès fréquemment utilisée pour s'ouvrir les portes des index de recherche notamment par exemple, sur les sites commerciaux et les serveurs de données.

Le système de recherche doit proposer un « certain degré d'approximation lors de l'appariement entre la forme saisie par l'utilisateur et celle enregistrée dans la base de données ». [Lallich-Boidin, Maret, 2005, p. 160]. Ces autorisations de formes approximatives introduites dans le logiciel et gérées par la machine produisent énormément de bruit\*.

### **2.8.1.4. La lutte contre le bruit et le silence dans la recherche d'information**

Le bruit documentaire résulte principalement d'une recherche peu précise, produisant un grand nombre de réponses peu pertinentes. Si l'on veut intervenir efficacement contre le bruit\*, il faut utiliser l'opérateur booléen ET afin d'ajouter un critère supplémentaire de recherche. On peut également utiliser l'opérateur booléen SAUF, pour exclure des réponses, enfin, il est nécessaire d'utiliser des termes spécifiques, qui précisent ou cernent au plus près la recherche.

Le silence\* documentaire résulte d'une recherche trop pauvre. La question est mal posée et les réponses obtenues sont peu nombreuses. Pour lutter contre le silence documentaire, il faut utiliser des synonymes, ou des termes génériques. On peut aussi utiliser l'opérateur booléen OU avec un autre mot.

## **2.9. Le silence**

Par ce terme, les professionnels de l'information désignent l'ensemble des documents jugés pertinents qui n'ont pas été retrouvés lors d'une interrogation, (question). « Le silence [documentaire] est dû essentiellement à la non concordance des termes utilisés pour l'interrogation et pour l'indexation ». [Richter, 1990, p.164].

### **2.9.1. Silence et indexation**

Les systèmes d'information qui pratiquent l'indexation contrôlée (*Rameau\**, liste de vedettes – matières, thésaurus, etc....), ont pour finalité documentaire de

réduire au maximum le silence, puisque l'indexation contrôlée permet l'accès aux seuls descripteurs autorisés.

Néanmoins ces modes d'indexations peuvent laisser échapper des pans entiers du savoir faute de connaître des descripteurs susceptibles de catégoriser des concepts nouveaux. (Cas des néologismes)

### **2.9.1.1. Silence et recherche d'information : le cas des suffixes**

La recherche d'information bute parfois sur les problèmes de traitement linguistique des suffixes, lesquels génèrent du silence à l'interrogation. Ils sont plus nombreux que les préfixes. « Les préfixes sont des affixes qui se greffent à l'extrémité droite d'une base » [Lallich-Boidin, Maret, 2005, p. 115].

La recherche d'information « est friande de morphologie dérivationnelle » (...) « En effet, les discours scientifiques et techniques en sont riches. La suffixation est donc source de silence car elle permet d'employer des formes voisines », [Lallich-Boidin, Maret, 2005, p. 119].

On appelle morphologie dérivationnelle, cette partie de la linguistique qui traite des mots nouveaux à partir de mots existants.

### **2.9.1.2. Le silence et la recherche plein texte ou texte intégral**

Dans les systèmes d'information qui utilisent la recherche d'information en texte intégral ou plein texte, le silence\* est également dû, comme on l'a évoqué plus haut pour le bruit\*, à des causes linguistiques et plus spécialement de synonymie, c'est-à-dire ici, au niveau de traitement automatisé de la langue, où « il englobe toutes les variations graphiques, du sigle à son expansion : *sgbd* ou *dbms* ou *système de gestion de bases de données* ou *data base management system*; du singulier au pluriel : *bonhomme* ou *bonshommes*; du masculin au féminin : *cancer malin* ou *tumeur maligne*; des formes verbales : *résout*, *résolvait*, *résoudra*, *résolu*, *résoudre* ». [Lallich-Boidin, Maret, 2005, p. 34].

## **Conclusion de la deuxième partie**

Dans cette seconde partie nous avons interrogé, tour à tour, la question et la réponse, comme objets de connaissance. On a vu que l'origine étymologique du terme question fait retour sur...la recherche. En posant, en premier lieu, la question de la question dans l'espace documentaire, nous avons tenté de savoir ce qu'était une question, et de connaître les rapports qu'elle entretient avec la réponse lors des contextes pragmatiques et multiples d'énonciation et de recherche d'information. Nous avons postulé que la question était, non seulement une information, mais également en attente de *quelque chose*, qui peut être un besoin d'information, et en tout cas, une réponse. Objet de médiation entre la connaissance d'une part et d'autre part entre un utilisateur, la question est encore une représentation mentale, un *objet médiateur* dans le cerveau du *rechercheur*.

En tentant de catégoriser la nature des questions, et en reprenant les typologies établies par des professionnels, on s'aperçoit que ces typologies

soulèvent des problématiques spatio-temporelles, lesquelles ont nécessairement des incidences sur les réponses, particulièrement dans la délimitation des réponses (chronologie). Inscrite dans un processus communicationnel, la question obéit à un certain nombre de fonctions linguistiques et intellectuelles. Inscrite dans un processus intellectuel de recherche d'information, la question interroge aussi les connaissances du *rechercheur* d'information et s'ouvrent alors les débats à propos de l'existence ou non des questions complexes ou difficiles....

En second lieu, notre intérêt s'est porté bien évidemment sur la réponse, en tant que but informationnel et résolution d'une question. Objet informationnel, la réponse existe, toujours en *réaction* à une question. (Ce qui permet de supputer en retour, qu'une question est *action*). Si la question documentaire soulève des aspects de catégorisations spatio-temporelles, la réponse en sus, révèle des problématiques de contenus informationnels et de connaissances plus élargies qui induisent des difficultés lors de la restitution de l'information demandée, (une réponse fait sens ou non pour un utilisateur et dépend d'une époque donnée, d'un état des connaissances donné, etc.....). Nous avons souhaité montrer que les réponses fournies au public par le bibliothécaire, n'était jamais évaluées, (à notre connaissance) ni validées, à l'exception notable des services de questions-réponses en ligne, où elles participent d'un processus collaboratif. Par ailleurs, nous avons tenté, peut-être maladroitement, de savoir s'il existait un lien, autre que la pertinence, entre les questions et les réponses et quelle pouvait être la nature de ces liens. L'idée de liens étant renforcée, entre autre, par les systèmes de questions-réponses, (*Sqr*) qui fournissent des résultats calculés et logiques, (appariement du couple questions-réponses) et les services de questions-réponses en ligne qui affichent des réponses, lesquelles prennent alors le statut nouveau de documents.

En troisième lieu, notre attention s'est portée vers les problématiques de pertinence documentaire, qui distinguent selon l'usage, la pertinence subjective (psychologique) des humains et la pertinence objective des systèmes, (algorithmique) des machines. Nous avons dit que la pertinence créait une somme, ou *un jeu* de réponses variables possibles, qu'elles soient obtenues par un médiateur humain ou une machine (parfois les deux) et que c'était en dernier ressort, à l'utilisateur de décider du sort de la réponse ou des réponses. Enfin, nous avons évoqué le « bruit » qui est l'obtention de réponses hors sujet en regard à une question posée, alors que le « silence » est le fait de ne pas obtenir de réponse tandis que les documents pertinents sont bien présents dans la base de données. Bruit et silence étant causés, en principe, par une mauvaise indexation et un choix de mots-clés non pertinents. Les technologies les plus récentes des systèmes d'information ne suppriment complètement ni le bruit ni le silence, qui participent de l'échec documentaire, dont nous allons présenter quelques unes des causes dans la troisième partie qui suit.

## **Troisième partie**

### **3. De quelques causes de l'échec documentaire**

Cette troisième partie est la plus difficile, parce que la plus fondamentale au sens scientifique. Elle appelle des travaux et des résultats d'études qui sont rares, des lectures nouvelles, des réflexions, du temps pour réfléchir, de pratiquer un « hors soi » d'observateur qui interroge son propre univers de recherche...De même, cette étude fait appel à des domaines de savoirs interdisciplinaires dont nous ne faisons qu'entrouvrir les portes, sans en visiter jamais en totalité les espaces tant sont vastes et inconnues ces terres émergentes et incessamment renouvelées de la connaissance.

Le temps et la place nous manquent, nous l'avons dit, nous aurions aimé utiliser le questionnaire que nous avons réalisé mais, faut-il à nouveau le souligner, l'absence de recherche et de publications, (au moins en langue

française), sur cet objet d'étude qui nous préoccupe, ne nous aura pas facilité la tâche. Si les études d'usages sont nombreuses<sup>51</sup> qui explicitent les pratiques des usagers en recherche d'information, il n'en est pas de même pour les études portant sur l'exercice des professionnels en situation de non réponse ou d'échec documentaire. C'est pourquoi nous souhaiterions vivement aller plus loin dans la recherche qui nous a conduit jusque là... (D'après nos lectures très récentes sur le sujet, il semble que les problématiques des *Sqr* soient de celles qui nous préoccupent).

Une question a-t-elle toujours sa réponse ? Plus précisément, pour ce qui nous occupe ici, une question a-t-elle sa réponse *en attente* d'être trouvée ? Une question s'imbrique-t-elle dans une réponse ? Existe-t-il des questions sans réponse ? Ce que le *chercheur* cherche à appréhender, c'est bien une réponse. Pourquoi semble-t-elle lui échapper ? Pourquoi parfois ne trouve-t-on pas ? Il est possible que l'on cherche quelque chose qui n'existe pas ou plus, ou bien encore est inaccessible : « Beaucoup d'informations essentielles ne sont enregistrées que dans la tête des gens » [Line, 2004, p.49].

Il y a un parallèle évident à faire avec les enquêtes (en quête) policières et la non élucidation de certains faits divers. Rappelons un élément de notre citation, (Cf. *supra*) du dictionnaire de philosophie *Lalande* : le terme « question » a connu un sens primitif « d'enquête et d'investigation » qui a disparu, [Lalande, 1991, p. 872]. Ainsi, il arrive qu'un lecteur vienne en bibliothèque avec un unique mot d'un titre de livre dont il demande au bibliothécaire de retrouver l'exactitude. Souvent, la mémoire du demandeur est défaillante. Si le bibliothécaire finit par retrouver le titre recherché, ce ne sera pas toujours avec le mot de départ. On se rappelle que dans la perspective de l'ethnométhodologie, « tous les énoncés, toutes les actions comportent une marge d'incomplétude et sont indexicaux. (...), » [Kilborne, 2006, p.47-48]. La solution quand elle est heureuse, appartient au hasard. Le hasard (aléas= hasard). Mais avant de s'avouer *vaincu*, le bibliothécaire aura fait appel à ses collègues.

On peut ne pas trouver une information parce qu'elle a été tout simplement retirée de la circulation (ou qu'elle n'a pas été rendue accessible). On ne trouve pas tout sur internet. Parallèle à faire avec les découvertes et les inventions. On ne trouve pas parce que l'on ne *sait* pas chercher. On ne trouve pas parce que l'on est ignorant du sujet demandé (incompréhension). Problèmes d'accès aux sources. On ne trouve pas parce que l'information demandée est complexe : exemples des questions (problématiques) à tiroir.

### 3.1. Des causes d'échec documentaire

En première analyse, il y a parmi les causes qui font qu'on ne (re)trouve pas l'information demandée ou attendue, celles qui dépendent de nous (*chercheur*) et pour qui est demandé un droit à *l'erreur*, et celles qui ne dépendent pas de

<sup>51</sup> A titre d'exemple, signalons : Renty, Emilie. [et al]. *Whist :enquête web sur les habitudes de recherche d'informations liées à la santé sur internet*. Paris ; Inserm UMR-S 707-faculté de médecine Pierre et Marie Curie, 2006-2007. 20 p. Disponible sur < [http://www.inserm.fr/fr/questionsdesante/rapports/att00003484/enquete\\_whist\\_2007.pdf](http://www.inserm.fr/fr/questionsdesante/rapports/att00003484/enquete_whist_2007.pdf) Consulté le 07/09/2008.

nous, (parmi cette dernière catégorie, nous rangerons le vol de documents ou toute dégradation). Mais ce départ est arbitraire et simplificateur, puisque nous y avons insisté, les technologies de l'information ont été élaborées pour l'essentiel, par des professionnels de l'information pour des professionnels de l'information. Pourtant, si aujourd'hui la science informatique imprime sa marque et a pris la main informationnelle et technologique, beaucoup de ces causes d'échec procèdent concurremment des techniques et des outils d'indexation classiques, tels les langages pré coordonnés<sup>\*52</sup>.

Les sciences cognitives, nous l'avons dit dès la première partie, appellent connaissances procédurales\*, les connaissances liées à l'action, (savoir-faire) et connaissances déclaratives, (savoir) les connaissances sur le monde. La plupart des causes d'échec documentaire trouvent ici leur schéma explicatif. Ce sont celles qui tirent leur origine du *rechercheur* d'information et celles qui relèvent de la conception des outils, de leurs pratiques et usages dans l'exercice professionnel.

Est considéré ici, « échec documentaire », une partie de l'ensemble des processus, linguistiques, sémantiques, psychologiques, technologiques, techniques, (bibliothéconomiques), ergonomiques, professionnels, communicationnels, (masse d'information) etc., qui entravent la recherche d'information et induisent des échecs dans les résultats, (non réponse). En un mot, il y a échec documentaire, quand on ne trouve pas l'information souhaitée.

Les causes d'échecs documentaires proprement dites sont innombrables et variées. Nous n'aborderons pas, faute de place, les causes liées aux erreurs de stratégies de recherche. Quelques causes seulement seront donc abordées maintenant, le plus souvent dans le cadre de la recherche documentaire informatisée.

Dans sa thèse de doctorat, [Ihadjadene, 1999, p.151 *et sqq.*] apporte des éléments de définition concernant la notion « d'échec » dans une recherche d'information. Il signale d'une part, « qu'aucun consensus n'est établi » en matière de définition « d'échec » et rappelle d'autre part, citant lui-même les travaux de [Tonta, 1992] qui avait distingué, en les synthétisant, « trois approches » différentes qui pouvaient être qualifiées d'échec, à savoir : « les mesures de performance précision/bruit, les mesures de satisfaction : il ya échec lorsque le document trouvé ne correspond pas aux besoins de l'utilisateur, l'inexistence des notices » [bibliographiques]. Dressant une typologie des échecs, Ihadjadene, [Ihadjadene, 1999, p.152-153] placé du côté des usagers, avait observé que les échecs d'ordre orthographique et syntaxique étaient les plus nombreux, « principale cause d'échec pour les usagers », [Ihadjadene, 1999, p.153].

Rappelons que nous soutenons que le *rechercheur* d'information est toujours un *apprenant* et qu'il a souvent la figure ou le profil de l'utilisateur. Ce critère est d'autant plus prégnant, qu'une partie des analyses et du corpus de la thèse d'Ihadjadene se porte vers l'étude des pratiques d'élèves-bibliothécaires de l'Enssib en situation de *rechercheurs* d'information, [Ihadjadene, 1999] laquelle précise que l'efficacité d'une recherche d'information dépend de la connaissance des procédures utilisées, des outils utilisés, de la manière dont la base est

---

<sup>52</sup> La pré coordination est « la combinaison de deux ou plusieurs concepts dans un descripteur synthétique faite pendant l'élaboration du langage documentaire. Antérieure à l'indexation du document, cette combinaison est donnée toute faite par le thésaurus ou la liste d'autorité. » [Richter, 1990, p.161].

indexée, etc. Selon lui, le taux d'échecs des catalogues informatisés avoisinerait les 30%. Mais, à notre connaissance, la plupart des études utilisables ne portent que ...sur des usages estudiantins ou de public en général, et non sur des pratiques de professionnels observés *in situ*.

Parmi toutes les problématiques rencontrées par le *rechercheur* d'informations, les plus délicates sont celles qui procèdent des problématiques du langage. « Les difficultés à retrouver l'information tiennent principalement aux caractéristiques du langage naturel », [Lefèvre, 2000, p.233].

### **3.1.1. Principales causes linguistiques inhérentes à la langue française**

Il est nécessaire pour la bonne intelligence du sujet de ce mémoire de faire un retour à quelques concepts de la linguistique sans toutefois espérer en approfondir les thématiques, qui risqueraient d'en éloigner. On examinera quelques unes des spécificités de la langue française desquelles viennent sourdre de « ces énigmes » qui sont l'objet de cette étude. En effet, d'une part, tout mot de la langue française peut-être compris, interprété différemment par chacun des différents locuteurs, et d'autre part, les techniques de traitements automatisés de la langue, malgré de notables progrès ne sont pas encore parvenues à faire en sorte qu'une question précise trouve sa réponse idoine ou correspondante. On se limitera donc à une manière de passage en revue sans trop y insister des principales difficultés langagières, afin d'éviter de dresser un catalogue hétéroclite, tel un inventaire à la Prévert

#### **3.1.1.1. Principales difficultés linguistiques et sémantiques**

Le langage naturel, celui que nous utilisons tous les jours n'est pas univoque. Un langage est dit univoque quand chaque mot et chaque expression possèdent un unique sens et une seule interprétation possible.

Parmi les principales difficultés d'une recherche qui peuvent venir se heurter sur la barrière de la langue naturelle, on repère principalement la polysémie et les problématiques de sémantiques telles la connotation et la dénotation.

La sémantique, qui peut se définir comme « une discipline qui a pour objet la description des significations propres aux langues et leur organisation théorique », [Tamba-Mecz, 1991, p.4], interroge le sens. « Mais qu'est ce que le sens d'un mot ? (...) Qu'est ce qu'une explication de sens ? [Wittgenstein, 1996 p. 35]. Et le sens n'est-il pas fonction et du contexte d'énonciation, en sus de l'énonciateur et de ses représentations (*signes*)? Le sens est donc bien une variable productrice d'*incertitude*. « Une co relation observée de variations », [Baltz]. C'est pourquoi, le sens *bruit* également et est cause d'échecs documentaires.

#### **3.1.1.2. Polysémie des mots, dénotation et connotation**

La polysémie est, avec l'homonymie, un des embarras majeurs du langage naturel Ils créent des phénomènes d'ambigüités sur les sens des mots. Exemple d'ambigüité polysémique : feuille de papier et feuille d'arbre. « Les poules du couvent couvent » est un exemple classique d'homonymie. Les traitements automatisés de la langue n'ont pas encore résolu ces difficultés syntaxiques et sémantiques

Par dénotation, on entend donner la définition d'un mot comme dans un dictionnaire. Sa signification est *scientifique*, pour le moins, *objectivée*. Dénoté un mot devient définir un terme, lui assurer son univocité. La dénotation repose sur des faits, des définitions, des descriptions. Ces termes sont majoritaires dans les langages contrôlés.

Au contraire, on entend par connotation, le fait pour chaque mot d'être *subjectivé*, interprétable, c'est à dire de faire le *plein de sens*, d'être soumis à une pluralité de points de vue. La connotation dépendant étroitement du contexte. Les mots connotés sont en principe issus du vocabulaire poétique ou du langage naturel. D'une part, lors des interactions verbales ou autres, ils ne manquent pas de produire de préjudiciables effets de flou sémantique et d'autre part, leur forte polysémie les fait rejeter des langages contrôlés. La folksonomie\* participe de la connotation. C'est pourquoi cette pratique collaborative, d'indexation réalisée par les internautes du web 2.0\* ne permet pas de véritables recherches documentaires.

### **3.2. Causes linguistiques et langages documentaires**

L'Afnor définit les langages documentaires comme « des langages artificiels constitués de représentation de notions et de relations entre ces notions, destinés, dans un système documentaire, à formaliser les données contenues dans les documents et dans les demandes des utilisateurs » [Afnor, 1987]. Ces langages sont tous des métalangages.

#### **3.2.1. Causes techniques inhérentes à la langue anglaise**

L'une des principales causes techniques induisant des échecs documentaire lors d'interrogation dans un contexte de recherche d'information automatisée est due au fait que les systèmes logiciels plein texte ou texte intégral sont d'abord conçus pour la langue anglaise et adaptés ensuite aux autres langues. « Le français est privé de ses accents, de ses cédilles [etc..] [Lallich-Boidin, Maret, 2005, p. 34].

### **3.3. Causes liées à l'analyse et à l'indexation des documents**

L'indexation est l'opération intellectuelle qui consiste à « affecter au document un ou plusieurs descripteurs, mots-clés ou indices » [Accart, Réthy, 2003 p.401]. Indice : du latin *dicere, dictus*, dire, puis de *indicium*, « ce qui montre, ce qui fait comprendre ».

Elle est toujours précédée de l'analyse documentaire. Mal réalisées, analyse et indexation peuvent être sources de bruit\* ou de silence\*. Indexer, c'est aussi montrer comment s'organise l'état des connaissances, comment ces connaissances sont hiérarchisées. C'est montrer où se situe ce que l'on cherche dans un ensemble beaucoup plus vaste, comment se tissent les liens entre *spécifique* et *générique*.

Indexer, c'est aussi (avec le « voir aussi »), ouvrir des portes vers d'autres disciplines en proposant des associations auxquelles on ne pense pas d'emblée. Toujours cette même idée de cartographier le savoir pour le localiser et d'en dessiner les routes et les chemins qui se croisent, se coupent, se rejoignent.

### 3.3.1. L'analyse documentaire

L'analyse est la première étape de l'indexation documentaire et improprement effectuée, une analyse documentaire peut engendrer des perturbations lors d'une recherche d'information. Selon l'Afnor, « l'analyse documentaire est l'opération qui consiste à présenter sous une forme concise et précise des données caractérisant l'information contenue dans un document ou un ensemble de documents ». [Waller, Masse, 1999, p. 16].

Ainsi, l'analyse est la recherche de l'information contenue dans les documents en vue du choix des concepts à représenter. Il s'agit avant tout de cerner les intentions de l'auteur. « Une histoire de la peinture est analysée par le mot clé *Peinture*, et l'analyse d'une histoire de l'art comprenant trois parties consacrées à la peinture, à la sculpture, à l'architecture, dégage le seul mot-clé *Art*, terme générique qui englobe ces trois disciplines », [Richter, 1990, p. 26-27].

On le comprend dans l'exemple cité, le choix des descripteurs est déterminant si l'on veut retrouver le document indexé, (réponse).

C'est cet usage professionnel qui peut entraîner avec lui le risque d'échec documentaire. En effet, la réduction à un seul descripteur va générer du silence documentaire.

Mal menée, une analyse documentaire engendrera des résultats incomplets. Au moins trois facteurs de risques principaux peuvent être évoqués, qu'il ne sera pas nécessaire de développer ici tant ils paraissent prégnants.

Ce sont principalement : la jeunesse de l'analyste, son manque d'expérience dans la maîtrise des langages contrôlés ou son manque de formation dans un domaine du savoir; sa lecture informative trop rapide du document. Cette dernière phase conditionnant le succès de l'extraction de mots clés-pertinents de l'indexation. On indexe mieux ce que l'on a préalablement compris de la teneur du document, mais l'étendue des champs du savoir constitue autant de chausse-trappes.

#### 3.3.1.1. L'indexation documentaire : une pratique à haut risque

L'indexation est une opération bibliothéconomique usuelle qui commande d'attribuer un ou plusieurs sujets (descripteurs, ou vedette-matière), deux en règle générale, à un document afin de pouvoir le retrouver.

La norme Afnor NF Z 47-102 en précise ainsi les modalités d'applications pour l'indexeur : « L'indexation est l'opération qui consiste à décrire et à caractériser un document à l'aide de représentations des concepts contenus dans ce document, c'est à dire à transcrire en langage documentaire les concepts après les avoir extraits du document par une analyse. La transcription en langage documentaire se fait grâce à des outils d'indexation tels que les thésaurus, classification, etc. ». [Afnor, N.F. Z 47-102, 1986, p. 2].

La difficulté fondamentale pour toute personne qui indexe un document est de bien saisir la différence entre les mots qui décrivent le contenu d'un document et ceux qui seront utilisés à l'interrogation d'une base de données. Les écarts d'analyses documentaires sont des écueils importants dans le processus de recherche documentaire.

#### 3.3.1.2. L'indexation au risque de l'indexeur et de l'idiosyncrasie

Suzanne Waller et Claudine Masse rappellent que si en théorie, « quel que soit l'indexeur un même texte devrait toujours être indexé de la même façon par le même indexeur » et que cet indexeur « devrait choisir les mêmes mots clés pour le même texte à plusieurs jours ou plusieurs mois d'intervalle », [Waller, Masse, 1999, p. 161], il en va tout autrement dans la réalité.

Un même texte lu par des indexeurs différents sera indexé selon les compétences de chacun et la compréhension (représentation), qu'ils ont du texte.

Plus l'indexeur est expérimenté et moins le risque de bruit et de silence ne risquent de survenir au cours de l'interrogation. Inversement, une connaissance pointue du sujet par l'indexeur expose au risque d'une représentation spécialisée ou poussée que ne pourrait atteindre en retour, un *rechercheur* d'information.

On vient de le dire, le sens varie selon différentes variables afférentes à l'indexeur (idiosyncrasie). C'est encore parce que le sens varie selon le contexte, que le concept d'indexicalité, issu de l'ethnométhodologie intéresse ici le professionnel de l'information.

### 3.3.1.3. Métalangage et ethnométhodologie

Le terme de « ethnométhodologie \* » est formé à partir de « ethnométhode » et « logos » et non de « ethno » et « méthodologie ». Cette remarque est d'importance car l'ethnométhodologie ne désigne pas une méthode particulière appliquée à l'ethnologie. Cette science consigne des ethnométhodes, c'est-à-dire les savoir-faire pratiques que mobilisent les individus dans le cours ordinaire de leur vie.

L'ethnométhodologie et son concept clé d'indexicalité, explique en partie ces variations et les fluctuations des représentations sémantiques construites par les indexeurs au cours de leur activité professionnelle.

Il n'y a pas loin, semble-t-il, de l'indexicalité à l'indexation. Les deux termes, étymologiquement proches, renvoient à la fabrication et à l'analyse du sens.

La définition de l' « indexicalité\* » est donnée par Garfinkel dans le chapitre 1 de son ouvrage : *Recherches en ethnométhodologie*. Il y revient également, à la fin de son livre, à travers un article écrit en collaboration, « *Les structures formelles des actions pratiques* », [Garfinkel, 2007, p. 429-473].

Le contexte topographique ici, c'est le terrain de la bibliothèque, de la médiathèque ou du centre de documentation. Sur ce lieu, les énoncés sont connectés, situés, ou bien par le bibliothécaire avec le document qu'il indexe ou bien avec le document indexé et d'autres bibliothécaires, supposés être ensemble à la recherche de la pertinence du document à indexer.

On peut raisonnablement supposer que l'indexeur subit des effets de contexte et variations de sens au moment d'indexer.

Indexer un document, un livre, une revue, un support numérique, cela a été dit plus haut, revient après une attentive analyse, à en extraire et à *coder*, (le plus souvent au moyen d'un descripteur ou d'un symbole numérique doté de sens, indice Dewey, par exemple) le contenu intellectuel et ou documentaire afin de pouvoir retrouver, (localiser) ce document aux fins de communication.

Indexer, c'est donc également donner du sens, le sens du contenu d'un document. C'est donc faire sens (action). C'est redonner du sens sur du sens,

puisqu'il s'agit d'en synthétiser le contenu, (sens), par condensation. Cette création aboutit à un *méta* sens, donc à un métalangage.

C'est en outre, pour le bibliothécaire en situation professionnelle, faire l'exercice de sa subjectivité.

Il s'agit encore de faire sens pour intégrer ce document dans une collection de documents. C'est une fois de plus, établir une relation *indiciaire* potentielle avec l'utilisateur, mais réelle aussi avec la collection de documents présents dans la bibliothèque.

Il faut savoir qu'un même ouvrage peut être indexé (analysé, c'est à dire *compris*, voire *interprété*) différemment par deux bibliothécaires. On estime qu'un même bibliothécaire pourrait attribuer plusieurs côtes différentes à un même titre. En effet, l'établissement d'une côte relève plus de pratiques d'habitudes et d'un usage professionnel que de techniques scientifiques. Ces variations de sens sont dépendantes de variables telles la culture générale et la bonne connaissance du sujet par le bibliothécaire-indexeur.

Cependant l'analogie ou la proximité avec le concept de l'indexicalité se nuance ici, car l'indexation connecte un document avec un utilisateur à condition que ce dernier sache faire sens, ce qui ne devrait pas poser de problème insurmontable, étant donné que l'ethnométhodologie postule que la vie sociale exige des individus qu'ils soient en éveil constant pour mettre « en œuvre un raisonnement sociologique pratique, une sociologie spontanée ». Chaque individu possède donc une compétence sociale spéciale pour comprendre et évaluer en le connaissant le monde environnant.

Pierre Lévy a nuancé fortement la notion d'indexicalité en ajoutant que « le contexte » est également « un enjeu, un objet perpétuellement reconstruit et négocié ». [Lévy, 1993, p. 26]. Sens et contextes sont mouvants, construits et... fluides. Ce qui ne semble pas novateur. Héraclite, (vers 500 av J.C.- 480 ? A.J.C.), avait imprudemment et sentencieusement déclaré « On ne peut pas se baigner deux fois dans le même fleuve », [Héraclite, cité par Brun, 1969, p. 136], s'était fait rabrouer, par Cratyle qui lui objectait qu'on ne pouvait même pas s'y baigner *une* seule fois !...

#### **3.3.1.4. L'indexation comme reformulation du document.**

Si indexer un contenu sémantique documentaire consiste à lui attribuer en retour des descripteurs, cette pratique de condensation du sens du contenu, en contractant précisément le sens entraîne *de facto* une perte d'information génératrice à l'interrogation, de silence.

#### **3.3.1.5. La représentation du document : l'étude d'Annick Bertrand**

Maryvonne Holzem dans son ouvrage *Terminologie et documentation*, [Holzem, 1999, p.149], s'arrête pertinemment sur une étude d'Annick Bertrand, membre du laboratoire *Travail et Cognition* de l'Université de Toulouse, intitulée « Compréhension et catégorisation dans l'indexation des documents », [Bertrand, Annick. 1994].<sup>53</sup>

L'étude d'Annick Bertrand s'intéresse aux problématiques soulevées par les pratiques des indexeurs qui utilisent le langage *Rameau*, très employé en bibliothèques et le texte qui suit en est un résumé indicatif.

4 groupes d'indexeurs, des débutants aux professionnels, ont eu pour tâche d'indexer des ouvrages en utilisant *Rameau*\*.

L'objet de cette étude consistait à indexer des ouvrages d'économie, de psychologie et de physique. 5 variables étaient utilisées : les trois premières concernaient les indexeurs, (1 : l'expertise et la tâche d'indexation, 2 : l'utilisation de *Rameau*, 3 : la familiarité avec le domaine des ouvrages). Les deux dernières relevaient du domaine et à la langue de l'ouvrage.

Le résultat de cette étude a permis d'isoler trois stratégies d'indexations différentes.

« 1 La première, celle du groupe d'experts professionnels indexant habituellement en *Rameau* et connaisseurs des domaines à indexer. (...). C'est à partir du langage documentaire qu'ils formaliseront les idées contenues dans les ouvrages. Cette stratégie pose la question du rapport entre indexation et catégorisation. L'extraction des termes est effectivement fonction de catégories préétablies dans la liste *Rameau* lors d'activités d'indexation antérieures, autrement dit par des indexeurs nord-américains.

2 La seconde stratégie mise en place relèvera d'un processus qualifié d'*ascendant*. Elle sera le fait des indexeurs débutants qui, ne connaissant pas *Rameau*, partiront des concepts identifiés dans les ouvrages pour ensuite remonter au langage documentaire qui sert alors d'outil d'aide à l'indexation. Annick Bertrand constate alors *un grand décalage entre les termes choisis pour accéder au langage et les termes d'indexation sélectionnés*, (p.69) Cette stratégie pose pour l'auteur le problème de l'indexation/ compréhension, un sujet étant capable de resituer une information sans comprendre le contenu du document dans laquelle elle est exprimée. (...).

3 La troisième stratégie, celle des professionnels indexant déjà dans ce domaine mais ne connaissant pas *Rameau*, relèvera comme pour le premier groupe, de processus descendants, guidés par des connaissances sur le domaine représenté par le document et orientés par l'objectif d'effectuer une indexation pointue. (p.69). Il en résultera une distorsion entre les termes proposés et le langage d'indexation. (...). Le fait d'indexer pour un public d'utilisateurs spécialisés dans un champ disciplinaire (le champ de la psychologie) (p.71), conduit les professionnels à extraire un nombre beaucoup plus élevé de concepts que dans les autres cas, et surtout à ne pas retenir les vedettes sélectionnées dans *Rameau* ».

En outre, les résultats de cette étude mettaient en évidence le fait que le travail d'indexation est une activité sociale normée, et que l'indexeur « construit une représentation non pas tant en fonction du contenu du document qu'il a entre les mains que du langage documentaire qu'il a coutume d'utiliser (...) ». [Holzem, 1999, p. 152].

Il faut noter que les conclusions de cette étude soulèvent des problématiques quasi universelles et absolument récurrentes lors de l'utilisation de listes normalisées de type *Rameau*\* et des variations des pratiques d'indexations en milieu professionnel.

### **3.3.1.6. . La « surindexation » documentaire comme source de bruit**

---

<sup>53</sup> Compréhension et catégorisation dans l'indexation des documents. *Cahiers du LCPE. Linguistique et ergonomie*.1994, n°1, p. 53-71.

« La surindexation consiste à choisir et à retenir plus de mots clés que d'informations contenues dans le texte étudié », [Waller, Masse, 1999, p. 162].

Ce phénomène se produit lorsque l'indexeur tentant de se prémunir à l'excès du silence, oublie d'un mot-clé pertinent, multiplie les mots clés, augmentant ainsi le bruit. Plusieurs explications sont possibles à cet état de fait : outre l'inexpérience de l'indexeur, qui estime que la multiplication des mots-clés permettra de retrouver l'information recherchée, il faut ajouter une notion psychologique liée à des considérations d'ordre déontologiques. « Le silence est proprement insupportable pour le professionnel de l'information qui se doit d'avoir réponse à tout (ou presque !», [Waller, Masse, 1999, p.164].

En faisant volontairement ou non, le choix du bruit contre celui du silence, l'indexeur, décide en quelque sorte « qu'il y ait quelque chose plutôt que rien », ce rien silencieux, agaçant et déstabilisant professionnellement.

Le silence place le documentaliste dans une posture si inconfortable, qu'elle est pour lui quasiment inadmissible.

Cette « surindexation », est une des causes possibles d'incapacité à retrouver des documents. Devenue plutôt rare chez les professionnels qui en connaissent les écueils, il semble qu'elle puisse connaître un renouveau avec les pratiques de flou sémantique des « *Tags\** », issus de la « *Folksonomie\** » du web 2.0\*.

En s'étendant quelque peu sur les pratiques d'indexation documentaire, on a tenté de démontrer qu'elles pouvaient, dans certaines circonstances, être productrices à leur tour d'erreurs, et par conséquent fournir au *chercheur* d'information autant de bruit\* que de silence\* qu'il n'attendait sûrement pas. Ces pratiques bibliothéconomiques sont de moins en moins réalisées manuellement, les professionnels de l'information recourant de plus en plus souvent aux technologies de l'informatique documentaire, et les documents (notices) téléchargés sont souvent pré-indexés, (externalisation des procédures).

Enfin, les problèmes de variation diachronique (émergence de mots nouveaux dans une langue), empêchent de faire des requêtes sur ces néologismes, puisqu'ils tardent à y être introduits, notamment dans les sources officielles.

Ainsi, les résultats d'une recherche effectuée en avril 2008 avec le terme « métadonnée(s) » sur le catalogue *Bn Opale plus* en ligne de la Bibliothèque nationale de France.

### **3.3.1.7. Recherche sur le terme « métadonnée »**

Métadonnée est un terme relativement connu des milieux de l'information-documentation. Il signifie donnée (*data*) sur une donnée (*data*). Il est curieusement peu employé comme descripteur. Il a pourtant valeur de terme générique.

A titre d'exemple, une seule référence est ramenée sur le catalogue en ligne de la Bibliothèque nationale de France, *Bn - Opale plus* [recherche simple sujet] avec le terme « métadonnée » au singulier d'abord et ensuite avec la marque du pluriel « métadonnées ».

1) Recherche « simple sujet » : *métadonnée* : aucune réponse<sup>54</sup>.

2) Recherche « simple sujet » : *métadonnées* : une seule réponse<sup>55</sup>.

À comparer avec la base de données commerciale, « *Electre* », du Cercle de la librairie, qui donne des résultats plus nombreux.

Cet exemple illustre la difficulté qu'il y a à retrouver parfois des informations simples. On retiendra également qu'il faut penser à utiliser la marque du pluriel, parce que dans *Rameau\**, les substantifs y sont toujours exprimés au pluriel, sauf cas exceptionnels, exemple : *Correspondance*, (la correspondance et l'art épistolaire).

### 3.3.1.8. *Rameau\** et *Le Métier de bibliothécaire*

Dominique Lahary, conservateur territorial et l'un des contributeurs de l'édition de 2003 du *Métier de bibliothécaire*, y notait que « *Rameau* est d'un maniement complexe et nécessite un apprentissage assez lourd. Le choix des autorités et la composition des vedettes prennent un certain temps, sauf quand on l'applique à un domaine limité par exemple [la musique] ». [Association des bibliothécaires français, 2003, p. 243]. Ces problèmes de repérage de l'information affectent ce que les bibliothécaires nomment les autorités, que sont « les listes des vedettes auteurs ou titres (catalogues descriptifs), des descripteurs décimaux (indices) (...), [obligatoires], [Richter, 1990, p. 159]. Ainsi pas d'existence du poète **Lorca** en bibliothèque, sauf à la forme du renvoi vers l'autorité, (donc l'entrée dans le catalogue) retenue : **Garcia Lorca**, Federico (1848-1936). (On trouvera des indications et des interrogations sur les modes de classements) dans : [Lahary, 2007, p. 229-260].

### 3.3.2. Principales causes liées aux modes et usages professionnels de classement des références

Cas du catalogue auteurs et titres d'anonymes des bibliothèques. On pouvait, (les « catalogues auteurs » manuels disparaissant) ne pas trouver les références aux documents si l'on ignorait les usages qui avaient présidé à la réalisation des catalogues de bibliothèques. Il fallait savoir, quand la recherche par le nom de l'auteur était impossible, soit parce que l'auteur d'un ouvrage était inconnu (ouvrage anonyme), ou soit parce qu'il y avait plus de trois auteurs (ouvrage dit anonyme par excès d'auteurs), il fallait savoir donc que la fiche de l'ouvrage était classée au titre de l'œuvre et intégrée dans le catalogue auteurs et titres

---

<sup>54</sup> Disponible sur le world wide web [en ligne] réf. du 11/04/ 2008. <<http://catalogue.bnf.fr/servlet/RechercheEquation;jsessionid=0000OGJv9-oK5ktVne0PfaQtZ6x:-1?host=catalogue>>.

<sup>55</sup> Disponible sur le world wide web [en ligne] réf. du 11/04/ 2008. <<http://catalogue.bnf.fr/servlet/RechercheEquation;jsessionid=0000OGJv9-oK5ktVne0PfaQtZ6x:-1?host=catalogue>>.

d'anonymes et qu'elle était classée (on disait intercalée), alphabétiquement derrière tous les auteurs...

Un usage spécifique aux bibliothèques municipales faisait limiter la constitution des catalogues des titres et partant, la seule recherche aux seuls ouvrages de fiction (romans). Il était quasi impossible de retrouver l'auteur d'un documentaire, sauf si l'on savait utiliser les catalogues systématiques matières.

La recherche par titre obéissait à des contraintes de classement alphabétiques quasi sibyllines : le classement alphabétique continu, et le classement alphabétique discontinu (mot par mot).

Des règles, toujours en vigueur dans les catalogues informatisés, régissant la position des articles, compliquaient la tâche du *chercheur* d'information.

Ces modes de classement de l'information concernent également les titres de périodiques.

### **3.3.2.1. Cas des périodiques**

Pour retrouver un article du journal *Le Monde*, il ne faut pas tenir compte de l'article (classement discontinu) mais « chercher *Le Monde* à la lettre *M* puis à *Mo*, etc. ». (...) Dans un catalogue informatisé ils [[les titres de périodiques avec chiffres] sont classés en début de liste. *Bulletin* tout court vient avant tous les bulletins de..., Le *Bulletin de l'Académie* est avant le *Bulletin des Lettres* qui précède le *Bulletin du Ministère de l'Éducation nationale*. », [Darrobers, Le Pottier, 2000, p.102].

### **3.3.2.2. L'information cachée**

On appelle information cachée, les références bibliographiques qui figurent en notes de bas de pages ou sont éparées dans les ouvrages et à la fin des articles de périodiques, par exemple. Leur repérage nécessite hasard heureux et temps de recherche allongé, [Darrobers, Le Pottier, 2000, p.80]. Une mauvaise manipulation, un saut de pages inopiné et l'information convoitée ne sera pas mise au jour.

L'indexation systématique Dewey d'une part, répartie les documents en fonction des disciplines et non selon les sujets, et d'autre part, dissémine physiquement les documents. Ainsi, « au sujet enfant », correspond un indice en psychologie :155.4 ; en sociologie : 305.23 ; en médecine : 618.92 ; en puériculture : 649.... » [Béthery,1998, p.22]. Nous ne développerons pas davantage ces problématique que soulève l'indexation Dewey dans la mesure où elles recoupent celles de l'indexation matière dont nous avons abondamment parlé lesquelles génèrent bruit\* et silence\* documentaires : « Si deux ou plusieurs sujets voisins sont traités dans l'ouvrage, on pourra choisir l'indice générique qui les englobe tous : un ouvrage de mathématiques en trois volumes, traitant respectivement d'algèbre (512), d'analyse (515), et de géométrie (516), se verra attribuer l'indice général des mathématiques 510 ». [Béthery,1998, p.33]. On ne retrouvera donc pas les ouvrages spécifiés par leur contenu.

Dans tout système homme-machine, l'interaction avec des outils cognitifs, génèrent des problématiques nouvelles. Ces problématiques sont d'autant plus prégnantes que « Le repérage, l'identification et la localisation des documents numériques

sont des techniques documentaires naissantes qui ne bénéficient pas encore de la longue tradition du repérage bibliographique que connaît l'imprimé », [Jacquesson, Rivier, 1999, p.239].

### **3.3.2 3. Le traitement automatisé de la langue : indications sommaires**

Pour être interrogeable au moyen des logiciels, la langue subit un certain nombre d'opérations qualifiées de traitement automatisé de la langue (Tal).

La langue française contient environ 50 000 entrées ou lemmes. « Ces 50 000 lemmes sont en fait un réservoir dans lequel les textes puisent 300 000 formes différentes. C'est la raison pour laquelle on dit que la langue française est une langue à morphologie riche. (...) C'est la raison pour laquelle les logiciels de traitement du français et d'autres langues comme l'allemand, l'italien, sont généralement conçus autour du couple dictionnaire de morphèmes et analyseur morphologique, et non d'un simple dictionnaire de formes, comme la plupart des logiciels dédiés à la langue anglaise. [Lallich-Boidin, Maret, 2005, p. 54].

#### **3.3.2.4. La lemmatisation**

Ce néologisme provient de la communauté des informaticiens qui se préoccupent des problèmes de traitement automatisé de la langue. La lemmatisation\* est un procédé qui consiste à donner à chaque forme du lemme, « d'une part des informations portées par le lemme en lexique et, d'autre part, le cas échéant, des valeurs de flexions (genre, nombre, personne...) ». [Lallich-Boidin, Maret, 2005, p. 64].

La lemmatisation\* d'un texte n'est pas une technologie entièrement efficace. « Suite à la seule analyse morphologique, environ 70% d'un texte est lemmatisé correctement. (...). Grâce à ces méthodes (méthodes de levée des ambiguïtés, et autres méthodes spécifiques), on obtient 95% des formes mais jamais 100% ». [Lallich-Boidin, Maret, 2005, p. 65].

Pour réaliser une lemmatisation\*, il faut un dictionnaire électronique et déterminer un ensemble de règles linguistiques d'une langue donnée. *Index* et *Ikaos* sont des lemmatiseurs, par exemple.

#### **3.3.2.5. L'indexation automatique**

On se rappelle que l'indexation est un processus intellectuel qui consiste à construire une représentation du contenu de chaque document, au sein d'un corpus de texte, d'une collection d'objets de toute nature ou d'un fonds de bibliothèque. Pour les professionnels de l'information, ce processus est encadré par des langages contrôlés de type *Rameau*\* ou des thésaurus, par opposition à l'indexation libre.

« Il existe deux méthodes pour indexer automatiquement un document à l'aide d'une liste d'autorité:

Rechercher chaque descripteur de la liste dans le document à indexer et conserver tous les descripteurs retrouvés ;

extraire des candidats descripteurs du document, chercher à les apparier avec des descripteurs de la liste d'autorité et conserver les descripteurs ainsi retrouvés ». [Lallich-Boidin, Maret, 2005, p ; 248].

Les ordinateurs parviendront-ils à comprendre le langage naturel ? Toute requête en langage naturel sera-t-elle jamais comprise ? C'est l'un des enjeux actuels majeurs des travaux des chercheurs sur le web sémantique.

### 3.4. De quelques causes liées à l'évolution incessante des technologies de l'information

#### 3.4.1. Surabondance informationnelle

La production de savoir, l'explosion informationnelle continue et exponentielle sont des facteurs générateurs de risques informationnel. Couplée à la complexité d'utilisation des systèmes et aux difficultés de repérage de l'information et à la vérification de sa validité.

Usabilité : On sait par la loi de Moers, « qu'un usager, tend à ne pas utiliser un système d'information lorsqu'il lui apparaît qu'il est plus pénible d'obtenir l'information demandée grâce à ce système que de ne pas l'avoir en ne l'utilisant pas », [Calcaly, 2008, p. 160]. Le *rechercheur* dont nous avons dit qu'il était également un usager, se trouve parfois confronté à la complexité des systèmes d'information, « à cause de leur difficulté d'apprentissage, de la mémorisation ardue de leur mode d'emploi, **sont à l'origine de nombreuses erreurs** », [souligné par nous], [Le Coadic, cité in Calcaly, 2008, p. 269]. Des études d'usage existent et nous y renvoyons notre lecteur.

Ergonomie des systèmes d'information : pour certains professionnels, les catalogues en ligne des bibliothèques ne seraient pas construits pour l'utilisateur qui cherche mais pour le lecteur qui trouve.

#### 3.4.2. Erreur de manipulation

Causes techniques<sup>56</sup>. Quelques fois, c'est le système et ceux qui l'ont paramétré qui est en cause.

« 30 à 40 % des recherches sur Opac aboutissent à des résultats qui ne correspondent pas aux cibles recherchées. (...) les membres du personnel [de la Bibliothèque du Congrès, Washington] préfèrent faire leurs recherches sur l'Opac. [Marchionini, cité par Wright, 1998, p.82]. L'Opac\*, la version externe du catalogue de la **bibliothèque universitaire de Paris 8**, montre quelques faiblesses relevées par des enseignants de cette université en 2005. « Il est n'est pas rare d'obtenir comme réponse à une interrogation des ouvrages qui ne correspondent pas à l'interrogation posée ». [Papy, Chauvin, 2005, p. 9]. Pour les auteurs de cet article, qui donnent en exemple les résultats infondés sur ....Pierre Lévy, les causes de l'échec documentaire sont à rechercher dans « Les logiques des concepteurs informatique [qui] déterminent le fonctionnement des logiciels. Rarement décrites, elles abandonnent l'utilisateur (...). [Papy, Chauvin, 2005, p. 9]. En pointant ainsi frontalement les insuffisances techniques des Opacs\*, comme « un système d'information primitif », [Papy, Chauvin, 2005, p. 9], les auteurs mettent le doigt sur la limite de ces interfaces, lesquelles pourraient bientôt céder la place à des moteurs de recherche plus performants et ou des Opacs dits de seconde génération.

### 3.5. Principales causes psychologiques et intellectuelles

Le manque de connaissances, induit évidemment des non réponses, des réponses partielles, voire erronées. Nous ne nous étendrons pas sur cet aspect délicat qui relève de la formation personnelle de chacun. *L'inculture générale* de quelques bibliothécaires peut induire des surprises et des non réponses : « Jusqu'à

---

<sup>56</sup> Sur ces questions voir : Printz, Jacques. Anatomie des erreurs humaines et actions préventives. In *Puissance et limites des systèmes informatisés*. Paris : Hermès, 1998. Chap.3, p.51-114.

la fin des années 80, la bibliothèque municipale d'Y. coule des jours tranquilles. (...) où l'on trouve par exemple *Le Rire*, de Bergson, classé dans la rubrique « humour », [Kuhlmann, Kuntzmann, Bellour, 1989, p. 194]. Nous manquons de données pour évaluer les facteurs d'âge, d'expérience et de niveaux de diplômes, etc. dans le processus de recherche.

Le raisonnement peut manquer de pertinence. Damasio, estime que « la finalité d'un raisonnement est d'amener à une prise de décision et que prendre une décision consiste, par essence à sélectionner une réponse, c'est à dire à choisir une action (non verbale), ou un mot, ou une phrase (...) en rapport avec une situation donnée. » [Damasio, 2008, p. 215]. Pour raisonner et prendre une décision, il est nécessaire de disposer de stratégies logiques. Or selon, Damasio, « les processus biologiques qui conduisent à la sélection d'une réponse ne s'inscrivent pas forcément tous dans le cadre de raisonnements », [Damasio, 2008, p. 216].

La notion de « biais cognitif » : un biais cognitif « est une erreur dans la prise de décision et /ou le comportement adopté face à une situation donnée résultant d'une faille ou d'une faiblesse dans le traitement des informations disponibles », [Wikipédia, 2008].

### **3.5.1. Erreur de navigation et surcharge cognitive**

La surcharge cognitive\* et la désorientation\* sont deux nouveaux symptômes de fragmentation de la connaissance repérés et constatés, par exemple, par [Ertzscheid, 2003, p. 133-142] lors des processus de navigations hypertextuelles sur le Web. Il apparaît, toujours selon ce jeune chercheur, que cette surcharge cognitive survient à cause de nos faibles capacités mnésiques dans le traitement simultané de l'information au-delà de cinq à neuf unités d'information. (On l'a vu *supra* avec [Dion, 1997, p.131]. qui parle de six à sept unités). La désorientation est due au manque de cohérence lors de la navigation. Prolongeant sa recherche, Ertzscheid en vient à évoquer, dans une terminologie médicale, le syndrome d'Elpenor, «ou réveil incomplet pour désigner les phénomènes d'amnésie topographique dont est victime l'utilisateur au cours d'une période de temps délimitée durant la session de navigation. [Ertzscheid, 2003, p. 133].

Ces divers avatars participent de l'échec documentaire, en tant qu'ils sont complémentaires de l'activité incessante du *chercheur* d'information.

### **3.5.2. Causes psychologiques et psychosociales**

On peut ne pas trouver parce que l'on perd la mémoire, (perte provoquée par divers facteurs : stress, fatigues, troubles visuels). On peut être déstabilisé par certains regards ou comportements d'utilisateurs. L'erreur (l'hypnotisation). Il y a plusieurs manières de se tromper.

Illusion d'optique (cas du bâton coudé dans l'eau, droit hors de l'eau). Dissonance cognitive (Festinger). Excès de confiance et /ou de certitude. Erreurs et préjugés. Précipitation.

Parce que l'on ignore tout de la question ou que l'on ne sait pas comment s'y prendre ...pour appréhender cette question.

Le manque de pugnacité dans la recherche, (entêtement), désintérêt pour la question, absence de motivation et implication moindre du *chercheur*.

## **3.6. Difficultés sémantiques et dialogiques**

On l'a vu plus haut, le dialogue constitue une fonction essentielle dans le processus interactionnel. La parole, élément constitutif du dialogue, est une source de difficultés importantes et de problèmes multiples. L'indexicalité le démontre, le sens varie en fonction du contexte mais aussi des individus (culture, niveau de langue, etc.). Pierre Levy a raison de souligner que le sens est toujours « transitoire, perpétuellement reconstruit et négocié » [Lévy, 1993, p. 26]. Nous suivons Merleau-Ponty qui déclare que « l'orateur ne pense pas avant de parler, ni même pendant qu'il parle », [Merleau-Ponty, 1945, p. 209]. Les risques de quiproquos, les lapsus, etc., sont nombreux et anticiper la demande d'un usager avant de l'avoir bien comprise peut être déclencheur de bien des malentendus.

### **3.6.1. Un dialogue imparfait, partiel ou l'incompréhension mutuelle**

L'échange de paroles doit s'articuler dans le meilleur des cas en un esprit de coopération. Pour cela, il est nécessaire que cet échange dialogique « contienne autant d'information qu'il est requis pour les visées de l'échange, mais ne contienne pas plus d'information qu'il n'est requis ». [Grice, cité par Breton, 2003, p. 27]. Trop d'information tue la demande d'information comme le manque d'éléments informationnels en oblitère la demande. Ainsi manque de concision, et précision...

#### **3.6.1.1. Les obstacles à la communication interpersonnelle**

Il arrive que des comportements, des conduites personnelles difficiles provoquent des perturbations dans les étapes du processus communicationnel et en modifient les états jusqu'à le paralyser. Si ces situations sont rares, elles peuvent néanmoins enrayer toute communication. Plus insidieux, car moins immédiatement perçus par les protagonistes, des processus psychologiques risquent d'inférer également si l'on n'y prête pas attention. Ils se fondent essentiellement sur le manque d'attention et de considération pour l'autre.

#### **3.6.1.2. Le respect de l'autre : l'absence d'écoute, d'empathie et de symétrie**

Dans un contexte relationnel, une mauvaise écoute, distraite ou partielle impliquera sinon des échecs de résultats, ou à tout le moins, une relation susceptible d'incompréhension mutuelle qui peut inférer négativement sur les réponses. Le manque d'empathie ou d'intérêt pour la question de l'usager n'est pas sans conséquence sur le déroulement de l'interaction. Si ce processus paraît évident, entretenir une relation asymétrique avec un usager, c'est à dire d'autorité, (intellectuelle) ou de hauteur (dédain ou mépris), ne promet rien de bon quant à l'issue du processus informationnel.

#### **3.6.1.3. Management de la performance : stress et pénibilité au travail**

L'introduction de la *culture* anglo saxonne de l'évaluation et du résultat dans le modèle français a provoqué une véritable modification en profondeur des rapports sociaux au travail, [Gaulejac, 2005]. Des sociologues parlant même « d'usure mentale », [Dejours, 2008.] Le professionnel formé, (formaté ?) aux concepts de l'idéologie gestionnaire de l'entreprise et confronté au pouvoir

managérial, et sur qui reposent des exigences de rentabilité et de performance (culture du résultat) est soumis à des pressions de plus en plus fortes de sa hiérarchie ou de sa tutelle. Les conditions de travail et leur pénibilité produisent des ratés professionnels et pas seulement dans la recherche d'info.

### **3.6.2. Questionné/ Questionneur : une relation ambivalente**

L'utilisateur, pour diverses raisons (timidité, défi lancé au professionnel, ...) n'explique pas la totalité de son besoin.

#### **3.6.2.1. La proxémie ou la gestion des distances**

« La distance sociale (...), marque selon les mots d'un de nos sujets, « la limite du pouvoir sur autrui » [Hall, 1978, p. 152].

On ne saurait ignorer que l'interaction cognitive qui unit un *rechercheur* d'information et un utilisateur est aussi une interaction située spatialement et socialement. Une mauvaise gestion de la distance nuit à l'anamnèse avec l'utilisateur d'autant plus que s'instaurent rapports de pouvoir et de séduction. Une jeune femme pourra se sentir mal à l'aise si elle sent son espace vital envahi et inversement, une personne d'origine étrangère ou handicapée, se sentira rejetée si la distance physique est anormalement augmentée.

La prise en compte de la gestion des distances sociales comme marqueurs, permet d'insister sur des causes indirectes des processus d'échecs, d'autant plus que des facteurs culturels plus généraux induisent des perceptions différentes des situations de contexte. Hall montre bien que même dans un contexte européen, « (...) les cultures réciproque divergent en de nombreux points. [Hall, 1978, p.161 et *sqq.*]. D'autres paramètres, tels le ton de la voix, la qualité du regard et l'empathie, par exemple, participent ou peuvent être causes directes ou indirectes d'échecs dans le processus d'interrogation. Beaucoup sont des causes d'origines psychologiques.

### **3.7. Causes économiques : aperçu rapide**

#### **La valeur de l'information : de la valeur immatérielle à la valeur économique**

« La valeur symbolique de l'information créée reste sans valeur tant qu'elle ne s'incarne dans la réalité physique d'un service », [Milon, 1999, p. 34].

« Intrinsèquement, la valeur d'une marchandise, d'un bien ou d'un service se définit selon les modalités de l'échange » [Milon, 1999, p. 34].

L'information est aussi une marchandise, un bien consommable. Qui a un coût, un prix, une valeur.

Si généralement l'information en soi n'a pas de valeur sur le plan de la théorie économique, elle donne quand même lieu à une industrie de l'information, [Milon, 1999, p. 87].

Que sert de savoir que le puits informationnel est là sous les pieds si l'on ne peut accéder à la source des données en raison de leur coût ? La plupart des bases de données importantes sont payantes. Exemples, la base de données textuelles Frantext disponible sur abonnement [<http://www.ciril.fr/~mastina/FRANTEX>].

Electre, la base de données bibliographique en ligne du Cercle de la librairie, n'est accessible qu'aux seuls abonnés. C'est un instrument de recherche

extrêmement utile, dans la mesure où l'on peut identifier, un document et savoir par exemple s'il est disponible dans le commerce. La recherche y est curieusement plus facile que sur le site de la BnF *Opale plus* relevant pourtant du service public.

### **3.8. De quelques incidences politiques sur la recherche d'information**

Même si ce paragraphe peut sembler éloigner des problématiques de ce mémoire, il nous a paru inévitable de faire allusion à la toute puissance hégémonique de Google Inc et à cet esprit missionnaire, qui caractérise souvent et « au nom du bien », la démocratie américaine.

« *Our mission is to organize all the information in the world* » Cette phrase prononcée au printemps 2005 par le représentant de Google Europe et rapportée par [Cassin, 2006, p. 11] illustre les enjeux qui se profilent au plan mondial dans la maîtrise et la gestion (organisation) des flux informationnels. Outre que l'information accessible par le seul Google n'est pas neutre, en ce sens qu'elle procède d'un calcul statistique et que le web n'est pas la seule source d'information, la tendance à une forme d'hégémonie culturelle « pour notre bien » n'est pas sans poser de profondes questions sur l'espace démocratique d'information ainsi mis en coupe réglée par les stratégies de Google.

#### **3.8.1. La manipulation de l'information :**

Les agences de presse qui diffusèrent des informations partiales et partielles pendant les deux guerres du Golfe, et les médias français alors sous contrôle militaire direct, ont démontré, s'il en était encore besoin, les limites de la liberté d'expression dans les Etats démocratiques. De nombreux médias aux Etats-Unis ont, après les attentats du 11 septembre 2001, adopté une posture « patriotique » que de nombreux experts ont évalué comme « propagandiste », [Breton, Proulx, 2006, p. 84]. Les traitements subis par l'information en de telles circonstances rendent la diffusion d'informations fiables extrêmement difficiles. Tout *rechercheur* d'information se doit d'être circonspect, car en périodes de troubles, de tensions politiques et sociales, l'information objective ou si l'on préfère, honnête, ne passe pas. (Cas des chiffres réels du chômage en France 2008). Si la rétention d'information peut-être considérée comme une action stratégique-politique des pouvoirs, les rumeurs, intox (hoax), pullulent sur le Net et créent autant de tintamarre informationnel que de bruit documentaire.

#### **3.8.2. La censure : un filtre politico-moral de l'information**

Censurée, l'information devient quasi introuvable ; ou bien il y faut mettre des moyens importants et astucieux pour contourner cette interdiction administrative ou juridique, et retrouver l'information recherchée. Sur l'internet *français*, des considérations morales, protection de la jeunesse, et politiques, interdiction des sites néo-nazis et négationnistes, par exemple, masquent mal les tentatives « de museler de façon globale l'expression sur internet. Derrière certains censeurs se profilent d'importantes compagnies de télécommunications, [Jacquesson, Rivier, 1999, p. 336]. Les attentats du 11 septembre 2001, puis ceux de Madrid, (2004) et de Londres,

(2005), issus de la mouvance islamiste ont ravivé les craintes d'une censure et d'une surveillance policière plus active.

Dans les bibliothèques publiques françaises, malgré les recommandations *libérales* et le *Manifeste*<sup>57</sup> de L'Ifla [*International Federation of Library Association and Institutions*] pour Internet, (mars 2002) et autres instances professionnelles en matière de liberté d'expression et de libre accès à l'information sur le web, un courant dominant parmi les personnels, orienté vers le contrôle actif des accès à certains sites, imprime sa marque militante et délimitante, prescriptive et prohibitive. Ce qui nous conduit à penser que l'on est passé insensiblement de la bibliothèque publique à la bibliothèque pudique, (multiplication de « Chartes d'usage Internet » locales restrictives). (Songeons, par comparaison, aux nombreuses scènes quotidiennes de meurtres et de sexe sur le petit écran).

On ne peut trouver que dans l'existant, (mise à disposition de l'information). L'exemple chinois : en 2005, Google a obtempéré aux autorités chinoises en acceptant de censurer l'accès à certains sites qui leur déplaisaient, instaurant par là « une nouvelle grande muraille de Chine. » [Cassin, 2006, p.176]. En Août 2008, à Pékin en pleine période olympique, les choses n'ont guère changé et le droit à l'information reste bafoué et contrôlé.

### 3.8.3. Les droits d'auteurs et l'accès aux documents

La question du droit d'auteur, qui a pour objet la protection des œuvres de l'esprit, constitue à elle seule une autre grande barrière juridique qui place souvent le *rechercheur* d'information dans l'embarras. Nous ne développerons pas davantage ces problématiques et leurs conséquences juridico-informationnelles qui nous entraîneraient trop loin.....

### 3.9. Les problématiques des mises à jour des ressources informationnelles liées à la durée de vie des pages web

Sur le web, le temps est compté pour les documents hypermédias. La durée de vie d'une page web oscille, selon les estimations, entre « 44 jours et un maximum de deux ans, (...) et par ailleurs une proportion significative des pages qui survivent connaissent des modifications », [PrévotEAU, Utard, 2005, p. 60]. **Error 404 page not found (on a server)** est la réponse automatisée du serveur, lorsque la page demandée a disparu. (on peut tenter de la récupérer en utilisant la fonction *cache*).

#### 3.9.1. Difficultés liées aux particularités fonctionnelles des moteurs

Les moteurs de recherche n'indexent que 16% environ des informations disponibles sur le web.

La méthode du *PageRank* (Google) ne permet pas d'accéder aux pages récentes « car logiquement, peu de liens pointent vers elles et les pages traitant d'un domaine très pointu, donc assez peu connu (leur indice est faible même si elles sont pertinentes » [Andrieu, 2004, p. 15].

---

<sup>57</sup> *Manifeste Ifla pour Internet* [en ligne] Disponible sur : < <http://www.ifla.org/III/misc/im-f.htm> >. Consulté le 05/08/2008.

Cette méthode conduit également à privilégier, à classer « un peu trop souvent dans les premières positions les mêmes pages, issues de sites « dinosaures » du web ». [Andrieu, 2004, p. 15].

De même, la mise à jour mensuelle de l'index de Google, la « *Google dance* », entraîne une relative perte d'information disponible sur le web. C'est un élément qu'il faut avoir présent à l'esprit quand on cherche de l'information très récente.

### **3.9.1.1. Google : une gestion des connaissances et de l'information accommodées**

Parmi les causes de non réponses possibles avec les moteurs de recherche, il faut signaler les *accommodements* avec l'information réalisés par les concepteurs de la technologie algorithmique, toujours plus ou moins secrète, de Google. Dépendant étroitement de la théorie des graphes, les moteurs de recherche ne proposent jamais une information neutre ou *objective*.

La théorie des graphes est la base mathématique de tous les algorithmes des moteurs de recherche sur le réseau web. (*Grphe* voir article de *Wikipedia*).

« L'un des éléments fondamentaux des algorithmes sur les réseaux est le facteur temps rapporté au nombre de nœuds examinés. Le temps d'achèvement d'un parcours (ou d'une recherche) qui relie un nœud du réseau à un autre dépend du nombre d'éléments du réseau, et varie toujours entre une valeur minimum et une valeur maximum, qui peuvent à l'occasion être très différentes, en fonction de l'algorithme qu'on utilise pour générer le parcours. Dans le réseau des pages web, chaque page est un nœud du graphe global et chaque lien est un arc. Si l'on prend le facteur temps comme prémisses, il apparaît avec évidence que les **résultats proposés par Google** comme résultat de n'importe quelle interrogation (en termes techniques, une *query* en fonction de données) ne peuvent pas s'appuyer sur une **consultation de « tout » l'internet** ». [Ippolita, 2008, p. 136].

Lors des interrogations et pour obvier au problème du temps de la recherche, des filtres sont mis en place, qui assurent « la vitesse du résultat final au moyen d'une série d'écarts et de choix étudiés expressément pour limiter l'intervalle (*range*) des possible blocs de données à analyser », [Ippolita, 2008, p. 137].

Google, ne pouvant pas parcourir tout le web à la recherche de l'information pertinente, réalise des sélections d'informations, et « engrange dans d'immenses banques de données de recherches préfabriquées, où des réponses standard sont fournies à des interrogations standard, mais aussi grâce au profil de l'utilisateur » [Ippolita, 2008, p. 137].

Ainsi, un utilisateur « (...), qui accomplit habituellement des recherches en français, n'interrogera pas toute la banque de données de Google, mais seulement la partie qui est en français », [Ippolita, 2008, p. 138].

L'information, ici son résultat d'interrogation, est pliée aux nécessités fonctionnelles de l'algorithmique et du traitement restitué de l'information. Cette torsion provoquée, artificialise les résultats informationnels et nuit à leur *scientificité*. Cette stratégie de la rapidité du résultat induit des soupçons de manipulation de l'information aux fins de marketing.

### **3.9.2. Critères de visibilité des moteurs de recherche sur internet : les problématiques du référencement**

Pour être retrouvé sur le web, il convient d'effectuer un bon référencement.  
Un bon référencement est un « processus qui s'effectue en 3 phases essentielles

1. Référencement : votre site doit être « trouvable » (« en rayon » dans l'index du moteur de la façon la plus complète possible.
2. Identification : une ou plusieurs des pages de votre site doivent se trouver « dans le lot » des pages identifiées car contenant les mots clés constituant la requête de l'internaute.
3. Positionnement : vos pages doivent être optimisées en fonction des critères de pertinence des moteurs afin d'être classées au mieux dans les pages de résultats pour vos mots clés choisis au préalable. (...) », [Andrieu, 2007, p.8].

Mal référencé, un site web n'apparaîtra pas et par conséquent, l'information de la page web demandée. Il existe des sites extrêmement intéressants qui, parce qu'ils sont mal référencés, restent invisibles.

Enfin, il faut savoir que les concepteurs de sites web peuvent interdire, le passage des moteurs à l'aide d'un fichier appelé « *Robots.txt* », rendant par là même le site ou la page web invisibles.

### 3.9.3. Le web invisible et ou profond

Reçoit cette appellation de web invisible toute la partie informationnelle du web qui n'est pas indexée par les moteurs de recherche, que ce soit pour des raisons techniques, sites accessibles par mots de passe, pages interdites aux moteurs de recherche, pages des intranets et extranets protégés par des coupes feu, banques de données payantes, pages dynamiques, etc.....On devine que ce web inaccessible aux moteurs constitue un immense ensemble de données non exploitables.

## Conclusion de la troisième partie

Nous avons consacré cette troisième partie à la présentation -seulement -de quelques causes principales d'échec documentaire, tant les possibilités de non-réponses sont innombrables. Un questionnaire pourrait nous apporter des réponses documentées, que faute de temps, nous n'avons pu adresser à ceux auxquels il était destiné. (Sauf erreur de notre part, le sujet des échecs documentaires des professionnels ne semble pas avoir retenu l'attention des chercheurs en Sciences de l'information, au moins pour la bibliographie en langue française).

La notion même d'échec documentaire est une *idée neuve* en Sciences de l'information et nous renvoyons sur ce point aux travaux de [Tonta, 1992, cité par Ihadjadene, 1999, p.151 et sqq.] qui avait distingué, en les synthétisant, « trois approches » différentes qui pouvaient être qualifiées d'échec. Nous soutenons que le *rechercheur* d'information, parce qu'il n'est jamais assuré de son résultat, est toujours un *apprenant* et donc, qu'il encourt d'autant plus de risques d'échecs, qu'il aura de questions à résoudre, et donc de réponses à soumettre, en même temps que naissent des savoirs nouveaux et multiples, et que se hausse concomitamment le niveau de formation des utilisateurs. Les technologies de la recherche d'information autant que le *rechercheur* d'information sont causes ou sources d'échec documentaire.

Parmi les causes d'échecs documentaires, on pourrait repérer celles qui dépendent directement du *rechercheur* d'information de celles qui ne dépendent pas de lui Mais, à l'évidence, ces causes sont plus complexes et s'imbriquent ou

s'engendrent les une les autres. Certaines sont d'ordre technique, (nous ne faisons que les effleurer ici) d'autres psychologiques et sociales. D'autres encore, politiques, juridiques, économiques etc.,.

Nous avons vu que la communication est un processus changeant, en perpétuel mouvement et réinterprétations. Or, La construction du sens et du contexte est à l'œuvre également dans les processus de construction des représentations et d'indexation, qu'elles soient documentaires ou non. Sens et contextes éclatent en une myriade de représentations qui font à nouveau sens et ainsi de suite, créant des infinités d' hypertextes peuplés de sens et de représentations nouveaux, induisant des espaces d'interprétations plausibles conférant à l'échec documentaire sa toute puissance génératrice.

## Conclusion générale

« Où donc est-il le vieux bibliothécaire, qui surveille, grognon, une salle de lecture, **insensible aux questions qu'on lui pose**, attentif seulement aux livres qu'il feuillète, et à l'heure du départ qu'il attend ? », [Bonnerot, 1951, p. 7-9]. La question de Jean Bonnerot, (1882-1964), conservateur en chef de la bibliothèque de la Sorbonne, (1939-1952), n'est pas qu'un simple clin d'œil. Elle vient rappeler qu'au delà de la caricature récurrente, pour ne pas dire facile, des générations de

professionnels de l'information ont eu le désir de répondre au mieux des besoins de leurs demandeurs.

Pour nous, qui nous sommes lancés dans cette entreprise d'interroger la question et de ses rapports avec la réponse et des *impédimenta* multiples qui entravent les processus de recherche et de restitution de l'information dans le cadre documentaire professionnel, sur quoi débouchons-nous ? Certes, le langage, écrit ou oral, est la cause principale de l'échec documentaire. La question (émetteur), induit-elle l'échec documentaire, quand il y a échec ou bien la réponse (récepteur) reste-t-elle insaisissable et toujours en attente d'être trouvée ? « A une question donnée ne peut-être apportée une réponse idéale. L'inverse supposerait que le contexte de la question est suffisant pour déterminer l'unicité de LA réponse. Or, ceci est rarement (pour ne pas dire jamais) le cas. » [El-Bèze, 2006, p. 277-297].

Il nous semble pertinent de soutenir et l'étymologie du mot nous y invite, qu'une question est une recherche, en tant que la recherche d'information est une série de questions et de réponses. Une recherche de *quelque chose*, dont elle est l'outil réflexif et cognitif, qui s'inscrit dans un processus intellectuel, *en soi* et *hors soi*.

Si la question est un fait nouveau et perturbateur de l'économie intellectuelle du sujet, l'apprentissage est un *élément* rendu nécessaire et permanent, dans les processus de recherche d'information, dans la mesure où celle-ci dépend de l'évolution incessamment recommencée des technologies, toujours nouvelles de l'information.

La question mène à la connaissance, le questionnement, qui est une série de questions, amène la connaissance, en tant que réponse *provisoire*, car contemporaine. La connaissance est porteuse de l'échec comme de la réussite.

Elucidation de l'énigme ? Au terme de ce travail, nous ne sommes pas parvenus à élucider complètement notre « énigme ». Une gageure. Ne nous sommes-nous pas posés à nous-mêmes avant tout une question qui nous intrigue aussi bien comme personne qu'en tant que professionnel ? Savoir ce qu'il en était des questions sans réponse, laissées en quasi *abandon* n'était peut-être qu'une interrogation personnelle relevant de préoccupations intellectuelles. Hormis les travaux de Le Coadic en France, les études au sujet des questions sont rares et ...américaines quand elles existent.

Notre résultat reste empirique et parcellaire. L'impossibilité de dépouiller et de traiter notre questionnaire, (Cf. Annexe 4) nous a contraint à limiter notre ambition initiale. Nous estimons que l'interrogation des pratiques professionnelles est un domaine mal connu, et que pour d'évidentes raisons, (situation dans l'entreprise) les études ou les recherches sur les causes d'échec documentaires au sein de la profession, quasi nulles. Pourtant, nous avons souvent entendu nos confrères faire part de leurs difficultés à répondre aux utilisateurs.

Une question est certes un fait nouveau. Cet objet en regard de sa réponse éventuelle que nous avons tenté d'identifier dans le contexte informationnel. Les causes d'échec, comme d'erreur, sont multiples. Le langage est la source principale des échecs. Source orale, on ne trouve pas à cause des formulations des énoncés de la question et de leurs nécessaires interrelations avec l'utilisateur. Source écrite, les langages documentaires, s'ils sont une tentative de réponse aux problèmes des ambiguïtés (polysémie, etc., ...) du langage naturel, n'en sont pas moins des réponses partielles, qui créent en retour des ...*phénomènes* de non

réponses. « L'information recherchée n'est pas externe et objective, mais plutôt construite en permanence et interprétée en permanence par l'utilisateur », [Thivant, 2002, p.61].

Nous avons montré que les indexations sont plutôt, d'abord des constructions sémantiques à haute variabilité produites par des indexeurs, et ensuite plus ou moins des traductions de contenus documentaires.

Nous avons vu que si une indexation *rapide* perturbe le résultat de la recherche d'information et génère bruit et silence, inversement, une indexation en profondeur, c'est à dire quand l'indexation se fait plus précise et prend en compte moult caractéristiques informationnelles, cette précision *indexicale* entraîne à son tour, quasi automatiquement du bruit documentaire lors de toute recherche.

Les réponses aux questions sont elles-mêmes des constructions de sens et de perception, additionnées des représentations de connaissances susceptibles de varier et de variations, produites par leurs énonciateurs et les contextes d'énonciation. Ainsi, un enchaînement subtil de représentations conduit-il à produire d'autres enchaînements de significations. Une réponse est alors une somme de réponses.

« Jusque au début des années 80, la recherche documentaire était affaire de spécialistes ». [Jacquesson, Rivier, 2005, p.314]. L'irruption de l'internet et des moteurs de recherche a modifié considérablement la donne de la recherche documentaire et informationnelle. Les moteurs de recherche produisent toujours des résultats pertinents même fortement *bruités*, malgré le fait qu'ils traitent des quantités énormes d'information. Ils sont devenus de précieux et indispensables auxiliaires en recherche d'information. Les quelques imperfections qui persistent lors des interrogations à l'aide des moteurs sont essentiellement dues d'une part, à l'approche syntaxique du document, à l'absence de contextualisation et à la qualité du contenu des bases de données, et d'autre part à la prédominance des sites commerciaux.

C'est pour quoi, dans le sillage de Tim Berners-Lee est née l'idée de Web sémantique. Le web sémantique vise à rendre les contenus informationnels non plus uniquement accessibles, mais également exploitables par les machines entre elles au moyen d'une « méthode de raisonnement, d'un format de standardisation, des métadonnées et des systèmes d'organisation des connaissances (ontologies). [Ghouas Dziri, 2006, p.193].

Sur quoi débouchons-nous ? *That is the question*. Et notre réponse (recherche ) pourrait manquer de pertinence.

Plusieurs problématiques s'entortillent dans notre sujet :

a) en premier lieu, celles des questions et des réponses. Il y a parfois des réponses univoques et automatiques, mais quoi qu'il en soit, elles existent toujours en fonction de contextes particuliers et du destinataire et sont donc susceptibles de variations. N'existent pas la question et la réponse indépendamment d'un contexte spatio- temporel et d'un but. Toutes les réponses sont incomplètes, en devenir permanent et laissent place à une part d'indécidable, (au sens de Gödel)<sup>58</sup>.

---

<sup>58</sup> Selon Kurt Gödel, (1906-1976), mathématicien et logicien américain, toute théorie laisse place à des affirmations indécidables.

Il n'y a pas une réponse univoque mais des ensembles de réponses possibles, des jeux de réponses plausibles, des réseaux sémantiques et hypertextuels de réponses.

Il est possible que des difficultés (échecs) viennent de ce que l'esprit cherche constamment à apparier questions et réponses, (attribution causale), ou de problématiques de pertinence documentaire.

b) En second lieu, celles des échecs documentaires des professionnels. Savoir ce sur quoi butent les professionnels de l'information, comment réagissent-ils face aux échecs documentaires, reste pour nous une question sans réponse. A moins que la plupart des difficultés rencontrées par les non professionnels soient de même nature que celles sur lesquelles viennent se heurter les professionnels.

Pourtant, aussi surprenant et paradoxal que cela puisse paraître, « **Collecter les informations inexistantes** » [Dalbin, 2008]<sup>59</sup> risque de devenir une préoccupation nouvelle et majeure des *chercheurs* d'information. En effet, « Le National Health Service (NHS) du Royaume Uni, considérant que les informations publiées sont facilement accessibles estime qu'il appartient aux médiateurs de se focaliser sur la mise à disposition d'informations inexistantes. Il s'agit en l'occurrence des questions posées par les patients et les médecins généralistes et qui ne trouvent pas de réponses dans les publications et les travaux des chercheurs. La création d'une base de données des « ignorances » ou des réponses incertaines scientifiquement sert même à déterminer des priorités en matière de politique de recherche dans le secteur médical ».<sup>60</sup>

## Annexes

---

<sup>59</sup> Dalbin, Sylvie. Collecter les informations inexistantes. In *l'œil de l'Adbs*, juillet-Août 2008, n°32.

<sup>60</sup> Communication faite au congrès de L'European Association for Health Information and libraries (EAHIL), Helsinki 23- 28 juillet 2008. < [Http://www.eahil.net/](http://www.eahil.net/)>. Consulté le 19/08/08.

## **Annexe 1**

Réponse envoyée par le service de réponses à distance de la BnF (SINDBAD) à notre question sur les sources bibliographiques concernant les questions réponses. [juin 2007].

Monsieur,

Pour votre recherche, nous vous recommandons de consulter la base de

données bibliographiques : LISTA (Library, information science & technology abstract), accessible gratuitement à l'adresse suivante : <<http://www.libraryresearch.com/>>.

Vous pouvez également interroger le catalogue de la bibliothèque de l'Ecole nationale supérieure des sciences de l'information et des bibliothèques, en ligne sur le site Internet de l'école :

<[http://www.enssib.fr/article.php?id=74&cat=La+biblioth%C3%A8que&id\\_cat=74](http://www.enssib.fr/article.php?id=74&cat=La+biblioth%C3%A8que&id_cat=74)>.

Par ailleurs, l'ENSSIB propose sur son site, des ressources en texte intégral, notamment le "Bulletin des bibliothèques de France" ainsi que les travaux de ses étudiants et de ses chercheurs, dans lesquelles vous trouverez peut-être des informations relatives au sujet qui vous intéresse.

Enfin, pour compléter votre recherche, vous pouvez venir consulter, à la Bibliothèque d'étude de la Bibliothèque nationale de France, plusieurs bases de données bibliographiques. L'une: LISA (Library and information science abstracts) est spécialisée dans les sciences de l'information et la bibliothéconomie ; l'autre : "Francis" est une base de données bibliographiques en sciences humaines, sociales et économiques, mais aussi en bibliothéconomie et en sciences de l'information.

Vous trouverez sur le site Internet de la BNF, à l'adresse suivante : <<http://www.bnf.fr/pages/zNavigat/frame/pratic.htm>>, des informations pratiques (conditions d'accès, tarifs et horaires) pour préparer votre visite, si vous le souhaitez.

**Annexe 2** : Document « *Guichet du savoir* » Bibliothèque municipale de Lyon

## Composition sous-jacente

Guichets du savoir :

de l'art, définition terme technique

[Suivre ce sujet](#) | [Envoyer ce sujet](#) | [Imprimer ce sujet](#)

**papyrus2008** Écrit le : 05-01-2008 21:43

Bonjour,

Usager



Groupe :  
Membres  
Messages : 2  
Inscrit le : 05-01-2008

Je souhaite retrouver un terme technique en art : quel terme désigne l'action de peindre à nouveau sur une toile déjà peinte, autrement dit sur un tableau déjà existant ? (Les peintres, faute de matériel, avaient parfois recours à cela.)

Merci d'avance de votre aide.



**bml art** Écrit le : 07-01-2008 10:00

### Réponse du département Arts & Loisirs

Usager  
confirmé



Groupe :  
BmL  
Messages :  
1 044  
Inscrit le :  
21-02-2004

Il s'agit du terme repentir, voir la définition dans [Universalis.fr](http://Universalis.fr).

D'après le [Dictionnaire de la peinture](#), le **repentir** : "*Changement de composition au cours de l'exécution picturale. Les repentirs ne doivent pas être confondus avec les repeints, dus à une main différente de celle du peintre*". Toujours dans le même dictionnaire, le **repeint** : "*Ce terme désigne la restauration locale pratiquée pour emplacer ou marquer une partie altérée de la peinture originale ou encore pour modifier les détails d'une composition qui n'est plus au gout du jour*".

Le terme repentir dans le site de [Wikipedia](http://Wikipedia).



**papyrus2008**



Écrit le : 11-01-2008 16:16

Bonjour,

Usager



Groupe : Membres  
Messages : 2  
Inscrit le : 05-01-2008

J'ai posé la semaine dernière une question sur ce thème à laquelle vous avez répondu, et vous en remercie. Toutefois, ma formulation manquait peut-être de précision, si bien que je n'ai finalement toujours pas trouvé le terme technique que je cherche...

Je renouvelle donc ma demande.

## Annexe 3 : Quelques services de questions-réponses complémentaires et leurs domaines, (vérifiés en juillet 2008)

### 3. 1. Sites de bibliothèques

BIUM Bibliothèque Interuniversitaire de Médecine, Paris

**BIUMINFO le service questions réponses de la BIUM**

[http://www.bium.univ-paris5.fr/biuminfo/refvirt\\_deb.htm](http://www.bium.univ-paris5.fr/biuminfo/refvirt_deb.htm)

Cité de la musique : <http://www.cite-musique.fr/francais/services/questions/index.asp> [musique]

SCD Lyon 1 Service de renseignement en ligne <http://butemp.univ-lyon1.fr:8080/uly1/>

Le site du Ministère de la Culture recense **10 services de questions réponses** sous la rubrique : Des

bibliothécaires répondent à vos questions :

[http://www.culture.fr/fr/sections/themes/bibliotheque\\_mediatheque](http://www.culture.fr/fr/sections/themes/bibliotheque_mediatheque)

Les Signets de la BnF recensent **une dizaine de services de questions-réponses** internationaux

Accueil>>Généralités>>Guides de recherche d'information>>Fenêtre sur les services questions-réponses à distance :

[http://signets.bnf.fr/html/categories/c\\_025\\_524s-a-d.html](http://signets.bnf.fr/html/categories/c_025_524s-a-d.html)

**Source** : *Les services de référence virtuels en France* : Fiche pratique #4 Réalisée par Catherine Jackson, Enssib, 2007

### 3.2. Sites commerciaux

Ca m'intéresse : [http://www.caminteresse.fr/questions\\_reponses/les\\_qr](http://www.caminteresse.fr/questions_reponses/les_qr) [généraliste]

Cybersciences.com [http://www.cybersciences.com/cyber/fr/questions\\_de\\_sciences/retour\\_au\\_sommaire.html](http://www.cybersciences.com/cyber/fr/questions_de_sciences/retour_au_sommaire.html) [sciences]

Explic : <http://www.explic.com/moteur.php> [généraliste]

Pourquoi.com <http://www.pourquoi.com/> (généraliste et science)

Questions et réponses.free .fr : <http://question.et.reponses.free.fr/> [généraliste]

Sqr onisep.fr : <http://sqr.onisep.fr/qr/index.php> [site de l'Onisep : information sur les métiers et les formations]

Svp: <http://www.svp.com/fr/index.php> [conseil en management]

Twikeo.com : <http://www.twikeo.com/> [généraliste]

Webchercheurs.com : <http://www.webchercheurs.com/> [généraliste]

Web-libre.org : <http://www.web-libre.org/questions/> [généraliste]

Wondir.com : <http://www.wondir.com/wondir/jsp/index.jsp> [en anglais]

Yahoo ! Questions/réponses, <http://fr.answers.yahoo.com/> [généraliste]

**Attention Google answer et Ketady.fr sont périmés !**

## Annexe 4

Questionnaire réalisé pour ce mémoire à l'intention des professionnels de l'information concernant les questions et sa lettre de présentation. Il a été réalisé en collaboration avec Jean Marie Firdion, sociologue à l'Ined, (Institut national des études démographiques) et chercheur associé à l'équipe ERIS, (Equipe de Recherche sur les Inégalités Sociales), au Centre Maurice Halbwachs du CNRS.

**N.B.** : faute de temps, ce questionnaire n'a pu être soumis aux enquêtés auxquels il était destiné.

Paris, le

Bonjour,

Etudiant en Master 2 de recherche en Sciences de l'information à l'Université de Paris 8 Saint-Denis, je souhaite réaliser une étude en direction des professionnels de l'information sur les causes de l'échec documentaire et sur quelques facteurs qui viennent perturber le processus de recherche d'information.

Je vous propose dans ce but de répondre au présent questionnaire réalisé à cette seule fin dont je garantis l'entière confidentialité.

Merci de bien vouloir le compléter avec vos éventuelles remarques, et de me le retourner rapidement à l'adresse mail suivante : [operrinbonnerot@gmail.com](mailto:operrinbonnerot@gmail.com)

Ajouter vos éventuelles remarques.

Cordialement,

Olivier Perrin

## Questionnaire à l'intention des professionnels de l'information :

- 1-sur la liste biblio.fr : <http://listes.cru.fr/sympa/info/biblio-fr>  
(bibliothécaires et documentalistes) et
- 2-sur celle de l'ADBS : <http://listes.adbs.fr/sympa/arc/adbs-info>

(documentalistes et professionnels de l'information).

Pourquoi ce questionnaire?

Ce présent questionnaire est réalisé pour collecter des informations auprès, d'une part, des professionnels de l'information, (bibliothécaires, documentalistes, ...) et, d'autre part, de toute personne en situation de recherche d'information inaboutie. Il s'agit en premier lieu de faire émerger des causes qui perturbent le processus de recherche, en second lieu de savoir pourquoi certaines des questions émises ne trouvent pas de réponses (silence), et troisièmement de savoir si réponse (solution) ou résolution il pouvait y avoir. Par conséquent de tenter de *ça-voir* où, comment et pourquoi ce processus cognitif est mis en échec.

Comment ?

Par un questionnaire standardisé : une vingtaine de questions soumises aux différents destinataires et adressées par mail.

A qui ?

Echantillon : professionnels de la profession

Sélection de l'échantillon ?

Au moyen de listes de personnes relevant de certaines catégories professionnelles, de listes d'établissements comportant un service de documentation...

Au moyen d'un réseau de connaissances (usagers de certains services, réseau amical, à partir d'un site Internet...)

Sur une zone géographique délimitée : agglomération parisienne,...

Hypothèses :

Un critère discriminant le niveau d'études et de qualification professionnelle, (niveau de formation).

Variabilité des réponses en fonction des niveaux d'études /qualifications.

Toute question posée à un service d'information est un fait nouveau, lequel peut induire des perturbations cognitives allant jusqu'au silence documentaire

Un professionnel chevronné parce qu'il est plus qualifié ne court-il pas de plus de risque de difficultés ou d'erreurs en raison même de son savoir faire, (des missions difficiles lui seront plutôt dévolues) ?

L'âge du *rechercheur* d'information n'entre t-il pas en compte dans la non résolution des besoins d'information des usagers ? (quelle est l'hypothèse sous-

jacente ? un effet de génération (ceux qui ne sont pas nés avec les nouveaux médias informatiques), un effet d'âge (difficulté à se mettre à *jour* quant aux nouvelles technologies et ressources documentaires, ou bien lassitude, attitude blasée, curiosité émoussée...), un effet de cycle professionnel (en fin de carrière les possibilités d'évolution, de promotion peuvent se réduire dans certains milieu et modérer en retour l'investissement dans le travail)...

Pourquoi certains -professionnels de l'information ou non -sont-ils en situation d'échec ?

S'agit-il des problématique liées à la nature des questions ? De la difficulté à *leur* apporter une réponse ?

Des problématiques liées à l'accès aux ressources informationnelles ?

Aux stratégies de recherche ? Aux différentes interrelations humaines et homme/machine ? Aux processus psychologiques, (anxiété, doute,...), et linguistiques (quiproquo, dialogue, reformulation), qui interfèrent ? Etc.

## Questionnaire sur la recherche documentaire

Olivier Perrin (Université Paris 8)

*Veillez cocher les cases correspondant à votre choix, ou répondre sur la ligne prévue à cet effet... Merci d'avance !*

### **A1** Quel est votre sexe ?

1. Homme.....  1  
2. Femme.....  2

### **A2** Quelle est votre année de naissance ?.....

19 |\_\_|\_\_|

### **A3** Quel est votre diplôme le plus élevé ?

1. Aucun diplôme .....  1  
2. CEP (Certificat d'études primaires).....  2  
3. BEPC, brevet des collèges .....  3  
4. CAP, brevet de compagnon .....  4  
5. BEP .....  5  
6. Baccalauréat, brevet supérieur, brevet professionnel.....  6  
7. Diplôme de 1<sup>er</sup> cycle universitaire, BTS, Bac +1 à 3.....  7  
8. Diplôme de 2<sup>ème</sup> cycle universitaire, Bac + 4 et 5.....  8  
9. Diplôme de 3<sup>ème</sup> cycle universitaire, doctorat.....  9

## *Partie B : Vie professionnelle*

### **B1** Quelle est votre profession ?

1. Bibliothécaire.....  1  
2. Assistant de bibliothèque, bibliothécaire adjoint.....  2  
3. Documentaliste.....  3  
4. Autre, précisez ⇒ [ \_\_\_\_\_ ]  4

### **B2** Quel est votre secteur d'activité professionnelle ?

1. Agriculture, Sylviculture, Pêche.....  01  
2. Industrie, Construction.  02  
3. Commerce, Artisanat .....  03  
4. Transport, Communication.....  04  
5. Finances, Economie .....  05

6. Administration publique .....	<input type="checkbox"/> 06
7. Education .....	<input type="checkbox"/> 07
8. Santé, Services sociaux .....	<input type="checkbox"/> 08
9. Tourisme .....	<input type="checkbox"/> 09
10. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 10
<b>B3 Etes-vous membre d'une association professionnelle ?</b>	
1. Oui.	<input type="checkbox"/> 1
2. Non.	<input type="checkbox"/> 2
<b>Partie C : La recherche d'information</b>	
<b>C1 Quel est le lieu dans lequel vous effectuez vos recherches documentaires ?</b>	
1. Centre de documentation ... ..	<input type="checkbox"/> 1
2. Bibliothèque .....	<input type="checkbox"/> 2
3. Médiathèque .....	<input type="checkbox"/> 3
4. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 4
<b>C2 Depuis combien de temps exercez-vous dans le champ de l'information-documentation ?</b>	
1. Moins d'un an .....	<input type="checkbox"/> 1
2. De un à cinq ans .....	<input type="checkbox"/> 2
3. Plus de cinq ans et moins de 10 ans .....	<input type="checkbox"/> 3
4. Plus de 10 ans .....	<input type="checkbox"/> 4
<b>C3 Dans quel(s) domaine(s) principaux effectuez-vous ces recherches ?</b>	
1. Sciences Humaines et sociales .....	<input type="checkbox"/> 1
2. Economique .....	<input type="checkbox"/> 2
3. Politique .....	<input type="checkbox"/> 3
4. Juridique .....	<input type="checkbox"/> 4
5. Information, media.....	<input type="checkbox"/> 5
6. Médical .....	<input type="checkbox"/> 6
7. Scientifique .....	<input type="checkbox"/> 7
8. Technique .....	<input type="checkbox"/> 8
9. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 9
<b>C4 Quels outils professionnels sur support papier utilisez-vous pour vos recherches d'information ?</b>	<i>Ici, plusieurs réponses sont possibles</i>
1. Bibliographies.....	<input type="checkbox"/> 1
2. Catalogues, fiches.....	<input type="checkbox"/> 2
3. Dictionnaires.....	<input type="checkbox"/> 3
4. Dossiers documentaires .....	<input type="checkbox"/> 4

5. Encyclopédies .....	<input type="checkbox"/> 5
6. Ouvrages de référence .....	<input type="checkbox"/> 6
7. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 7
<b>C5</b> Quels outils informatiques utilisez-vous pour vos recherches d'information ?	<i>plusieurs réponses sont possibles</i>
1. Bases de données documentaires .....	<input type="checkbox"/> 01
2. Catalogues en ligne, Opac.....	<input type="checkbox"/> 02
3. Forum de discussion, blogs.....	<input type="checkbox"/> 03
4. Moteurs de recherche généralistes.....	<input type="checkbox"/> 04
5. Moteurs de recherche spécialisés .....	<input type="checkbox"/> 05
6. Agents intelligents (Copernic @...) .....	<input type="checkbox"/> 06
7. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 07
<b>C6</b> Quel outil d'indexation et quel type d'indexation utilisez-vous ?	<i>plusieurs réponses sont possibles</i>
1. CDU .....	<input type="checkbox"/> 1
2. Dewey .....	<input type="checkbox"/> 2
3. Langage Rameau .....	<input type="checkbox"/> 3
4. Plan de classement .....	<input type="checkbox"/> 4
5. Thésaurus .....	<input type="checkbox"/> 5
6. Autre, précisez ⇒ [ _____ ]	<input type="checkbox"/> 6
<b>Partie D : Les questions et leur résolution</b>	
<b>D1</b> Pensez-vous qu'une question posée dans un contexte documentaire est un fait nouveau, un événement ?	1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non 3. <input type="checkbox"/> nsp
<b>D2</b> Estimez-vous que répondre à une question, c'est résoudre un besoin d'information ?	1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non 3. <input type="checkbox"/> nsp
<b>D3</b> Existe-t-il selon vous des types de questions de nature plus difficile que d'autres ?	1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non 3. <input type="checkbox"/> nsp
<b>D4</b> Pensez-vous qu'il existe des domaines de savoir (ou de la connaissance) plus difficiles à appréhender ?	1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non 3. <input type="checkbox"/> nsp
<b>D5</b> Vous est-il arrivé de ne pas pouvoir/savoir répondre ?	
1. Oui .....	<input type="checkbox"/> 1
2. Non ⇒ <i>aller en E1.</i>	<input type="checkbox"/> 2
<i>Si oui :</i> La dernière fois que vous n'avez pas pu ou su répondre... ?	
<b>D5.1</b> Quelle était la nature de la question ?	
1. Factuelle .....	<input type="checkbox"/> 1

<p>2. Thématique ..... <input type="checkbox"/> 2</p> <p><b>D5.2 Le domaine couvert par cette question ?</b></p> <p>1. Sciences Humaines et sociales ..... <input type="checkbox"/> 1</p> <p>2. Economique ..... <input type="checkbox"/> 2</p> <p>3. Politique ..... <input type="checkbox"/> 3</p> <p>4. Juridique ..... <input type="checkbox"/> 4</p> <p>5. Information, media..... <input type="checkbox"/> 5</p> <p>6. Médical ..... <input type="checkbox"/> 6</p> <p>7. Scientifique ..... <input type="checkbox"/> 7</p> <p>8. Technique ..... <input type="checkbox"/> 8</p> <p>9. Autre, précisez ⇒ [ _____ ] <input type="checkbox"/> 9</p>	
<p><b>D6</b> Pouvez-vous donner un exemple (bref) de question auxquelles vous n'avez pas pu apporter de réponse ?</p> <p>[ _____ ]</p>	<p>[ _____ ]</p>
<p><b>D7</b> Pensez-vous qu'il existe des questions sans réponse ?</p>	<p>1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non 3. <input type="checkbox"/> nsp</p>
<p><b>D8</b> Quand vous ne parvenez PAS à répondre à une question, cela tient ... ?</p> <p>1. ...aux difficultés à trouver des sources d'information ?</p> <p>2. ...aux structurations de l'information (organisation) ?</p> <p>3. ...aux repérages de l'information (indexation) ?</p> <p>4. ...à la difficulté pour vous à comprendre la question ?</p> <p>5. ...à la difficulté d'utilisation des systèmes d'information ?</p> <p>6. ...aux interrelations dialogiques de personne à personne ?</p> <p>7. ...aux interrelations dialogiques de personne à ordinateur ?</p> <p>8. ... à la gestion du temps avec l'utilisateur ?</p> <p>9. ... au comportement de l'utilisateur ?</p> <p>10. ... au manque d'intérêt de la question ?</p> <p>11. ... aux problèmes soulevés par les critères de pertinence ?</p> <p><b>12</b> autre, précisez ⇒ [ _____ ]</p>	<p><i>plusieurs réponses sont possibles</i></p> <p>1. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>2. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>3. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>4. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>5. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>6. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>7. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>8. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>9. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>10. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>11. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p>
<p><b>D9</b> Lorsque vous n'avez pas répondu à une question, quelle pratique avez-vous ?</p> <p>1. ...réorientation de l'utilisateur ?</p> <p>2. ...recours à un tiers professionnel ?</p> <p>3. ...recours à un réseau ?</p> <p>4. ...énoncer l'échec à l'utilisateur ?</p> <p>5. ... autre, précisez ⇒ [ _____ ]</p> <p>[ _____ ]</p>	<p><i>plusieurs réponses sont possibles</i></p> <p>1. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>2. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>3. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>4. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>[ _____ ]</p>
<p><i>Partie finale E</i></p>	

<p><b>E1</b> Pour vous, qui valide la réponse que vous donnez ?</p> <p>1. La source d'information .....</p> <p>2. Votre expertise .....</p> <p>3. L'utilisateur .....</p>	<p><b>E1</b> <i>plusieurs réponses sont possibles</i></p> <p>1. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>2. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p> <p>3. <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> nsp</p>
<p><b>E2</b> Sauf dans le cas d'une réponse binaire (de type vrai ou faux), êtes vous d'accord avec cette affirmation : « une réponse à une question n'est jamais qu'une approximation »</p>	<p>1. <input type="checkbox"/> tout à fait d'accord</p> <p>2. <input type="checkbox"/> plutôt d'accord</p> <p>3. <input type="checkbox"/> plutôt pas d'accord</p> <p>4. <input type="checkbox"/> pas d'accord du tout</p>
<p><b>E3</b> Vous est-il arrivé qu'un usager vous déclare que vous n'avez pas répondu à sa question ?</p>	<p>1. <input type="checkbox"/> oui 2. <input type="checkbox"/> non</p>
<p><b>E4</b> Etes vous d'accord avec cette affirmation : « si l'utilisateur s'adresse à vous, c'est qu'il a déjà effectué des recherches, ce qui augmente la difficulté de la recherche »</p>	<p>1. <input type="checkbox"/> tout à fait d'accord</p> <p>2. <input type="checkbox"/> plutôt d'accord</p> <p>3. <input type="checkbox"/> plutôt pas d'accord</p> <p>4. <input type="checkbox"/> pas d'accord du tout</p>
<p><i>Merci d'avoir pris le temps de nous répondre !</i></p>	
<p><b>F1</b> Si vous souhaitez ajouter des informations ou des remarques sur les sujets dont nous avons parlé ou sur d'autres, ces lignes sont à votre disposition :</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p> <p>.....</p>	<p>[ ]</p>

## Annexe 5

### Glossaire des expressions et termes techniques cités. Ils sont porteurs d'une \* dans ce mémoire

**Bruit** : On appelle bruit l'obtention de réponses hors sujet lors d'une interrogation. « Il conditionne la pertinence des réponses ». [Accart, Réthy, 2003, p.394].

**Champ** : Terme documentaire utilisé en informatique qui désigne un des éléments de la notice bibliographique et qui sert à la saisie et à l'interrogation. Champ auteur, champ titre,... [Accart, Réthy, 2003, p.395].

**Champ sémantique** : regroupement d'un ensemble de descripteurs d'un thésaurus autour d'un même domaine d'application.

**Cognitivism** : Le cognitivisme est un courant de recherche multidisciplinaire, (linguistique, neuroscience, psychologie, intelligence artificielle, informatique, cybernétique, etc...), dont l'objet est l'étude du fonctionnement de l'activité intellectuelle.

**Connaissances déclaratives** : « Font référence à des faits, à des informations qu'il est possible de décrire verbalement et dont l'organisation est flexible et contrôlé par le sujet ». [Le Maire, cité par Dennecker, Kolmayer, 2006, p. 292].

**Connaissances procédurales** : « Qualifie un type de connaissance peu accessible à l'observation et qui est de l'ordre du savoir-faire ». [Baron, cité par Dennecker, Kolmayer, 2006, p. 292].

**Dégradation d'une question** : La dégradation d'une question consiste à proposer une suite de reformulations possibles plus générales (moins contraignantes), que la question de départ. [Lallich-Boidin, Maret, 2005, p.203]. Dans le contexte du traitement automatisé de la langue, cette dégradation est contrôlée par des règles appelées *patrons linguistiques*.

**Désorientation** : « Effet cognitif produit chez l'utilisateur qui perd la liaison entre son projet de navigation et les cartes ou zones d'information qu'il est entrain de lire ». [Rhéaume, 1993, cité par Ertzscheid, 2003, p.137].

**Dublin Core** : Ensemble de 15 métadonnées utilisées pour décrire les documents électroniques sur le web. Norme Iso depuis février 2003 ; <<http://www.dublincore.org>>.

**Folksonomie** : Ce mot, contraction de folk (peuple) et taxonomie, désigne « une forme de classification collaborative (indexation) spontanée », basée non sur un vocabulaire contrôlé et standardisé, mais sur des termes choisis par les utilisateurs eux-mêmes. C'est un des usages actuels du web 2.0.

Heuristique : « Qui sert à la découverte. Se dit d'une hypothèse que l'on adopte à titre provisoire pour sa fécondité dans la recherche, sans se préoccuper de sa justesse ».[Grawitz, 1990, p. 1079].

**Hybride** : Nouveau concept appliqué aux bibliothèques « unissant bibliothèque matérielle [traditionnelle] et bibliothèque immatérielle [numérique] » [Eboli, 2007, p.9]

**Hypertexte** : Fonction logicielle qui, dans un texte, associe à certains mots ou groupe de mots grâce à un lien défini à l'avance, des compléments d'information. [Duchemin, 2000, p. 538]. L'hypertexte est par nature non-linéaire.

**Indexicalité** : Propriété qu'ont les énoncés ou les comportements d'être toujours rattachés à un contexte ou en rapport avec un référent, d'où il découle que nous construisons en permanence le sens des mots. Dans la perspective de l'ethnométhodologie, tous les énoncés, toutes les actions comportent une marge d'incomplétude et sont indexicaux. (...). [Kilborne, 2006, p.47-48].

**Information spécialisée** : « L'information spécialisée est une information destinée aux spécialistes (industriels, chercheurs, enseignants) qui leur est nécessaire dans leur activité professionnelle et qui conditionne une prise de décision ou une action particulière. Il faut la différencier de l'information pour le grand public notamment par la nature des informations transmises ainsi que le canal de communication emprunté » (Ghislaine Chartron, *De l'information spécialisée à l'information élaborée : problèmes de modélisation*. 8<sup>e</sup> congrès SFCIC, Lille, 21-22 mai 1992, p. 39).

**Lemmatisation** : La lemmatisation consiste à reconnaître chaque forme fléchie du texte afin de lui faire correspondre une forme canonique (forme infinitive d'un verbe, forme au singulier d'un nom, etc.), et une catégorie morphosyntaxique (nom, adjectif, etc.).

**Modèle booléen pondéré** : « Modèle statistique dans lequel des poids sont préalablement attribués aux mots du corpus et les termes de la requête peuvent également être pondérés ». [Lefèvre, 2000, p. 166].

**Modèle probabiliste** : « Entre une requête et un document, est calculée à partir de l'estimation de la probabilité que le document a d'être pertinent pour cette requête ». [Lefèvre, 2000, p. 170].

**Modèle vectoriel** : « Consiste à représenter un document par un *vecteur*, qui comporte autant de composantes qu'il y a de mots significatifs dans le corpus (les mots vides ne sont pas pris en compte) ». [Lefèvre, 2000, p. 168].

**Opac** : *On Line Public Acces Catalog*. En français accès public en ligne ou encore catalogue interactif en ligne : *Ciel*. C'est l'interface permettant à l'utilisateur de formuler sa requête en vue d'obtenir une réponse. (« C'est un système de gestion documentaire. On distingue deux modes d'utilisation des SGD : le *mode production* réservé aux professionnels et le *mode consultation* accessible à tous »). [Rivier, 2007, p. 70].

**Pertinence** : « Qualité qui convient à l'objet dont il s'agit, de ce qui est conforme à la raison, au bon sens ». (*Le Robert 1*). Le taux de pertinence est le rapport du nombre de documents pertinents au nombre total de documents obtenus en réponse à une question

**Pondération** : attribution d'un poids ou valeur à un descripteur en fonction de son importance (fréquence), dans un texte indexé.

**Post-coordonné** : La post-coordination est établie au moment de l'indexation en associant deux termes avec l'opérateur booléen ET.

**Pré-coordonné** : Dans certain langage documentaire, les termes sont combinés avant l'indexation afin d'associer des notions proches en les reliant par la syntaxe à l'aide d'un tiret, d'une étoile. (Ex : *Rameau*).

**Rameau** : Acronyme pour Répertoire d'autorités matières encyclopédique et alphabétique unifié (Rameau). Langage pré-coordonné issu de la liste des vedettes matières de l'université Laval du Québec, elle même issue de la *LCSH* (Library of Congress Subject Heading). Rameau est le langage d'indexation de la Bnf, des bibliothèques universitaires, de la plupart des bibliothèques et médiathèques publiques. < <http://rameau.bnf.fr>>.

**RDF** : *Ressource Description FramWork*, RDF est un standard qui sert à décrire de manière structurée l'information contenue dans les documents. C'est une de ces métadonnées qui visent à promouvoir le web sémantique.

**Silence** : Désigne l'ensemble des documents pertinents qui n'ont pas été retrouvés lors de l'interrogation d'une base de données. S'oppose au bruit (un trop plein de documents).

**Surcharge cognitive** : « Effort additionnel et concentration nécessaire pour maintenir plusieurs tâches ou plusieurs parcours [*trails*] en même temps ». [Conklin, 1987, cité par Ertzscheid, 2003, p. 135].

**Syntol** : Langage documentaire construit dans les années 60. « Les relations y sont associatives, prédicatives consécutives (...) » [Chaumier, 2003, p. 46].

**Tag** : Tags Cloud (Nuage de mots). Mot clé. Liste de tags, souvent en ordre alphabétique, dont la taille dépend de la fréquence d'utilisation. Mode d'indexation des hypermédias par les internautes dans le cadre du web 2.0.

**Thésaurus** : Liste de termes normalisés. Un thésaurus est une liste organisée de termes descripteurs sélectionnés pour indexer une documentation relative à un domaine de la connaissance, ou à un secteur d'activité. [Richter, 1990, p.165]. Selon l'AFNOR, les descripteurs sont « des termes, des mots ou des expressions retenus parmi un ensemble de synonymes, de quasi synonymes et de termes apparentés, pour représenter de manière univoque un concept (...) ».

L'ensemble des termes qui constituent le thésaurus est réparti en champs sémantiques, à l'intérieur desquels les descripteurs sont reliés par des relations de

significations qui sont de 3 types : 1 la relation d'équivalence, 2 la relation hiérarchique 3, la relation d'association

**Usabilité** : « L'usabilité mesure jusqu'à quel point, un produit d'information, un système d'information, un service d'information, une information est prête à l'usage », [le Coadic, 1997].

**Web 2.0** : Le terme a été créé en 2004 par Dale Dougharty et popularisé par l'éditeur Tim O'Reilly. La notion de Web 2. 0 désigne une évolution qualitative des modes d'usages et d'appropriation des services web. Cette évolution repose sur une succession d'améliorations technologiques, ergonomiques et sémantiques. Elle se traduit par l'émergence de nouvelles formes d'interactivité. Le Web 2.0 se définit comme social et collaboratif.

**World wide web ou W 3 ou Web** : Système d'information réparti hypermédia. C'est une application client/serveur. Le client et le serveur dialoguent grâce au protocole *HyperText Transfert Protocol* (http). Le client doit disposer d'un logiciel de navigation (« *browser* »), pour accéder aux services du Web.

**Z 3950** : Norme américaine ANSI Z 3950 devenue la norme ISO 239.50. Elle permet l'interrogation des bases de données bibliographiques indépendamment des systèmes qui les gèrent.

## Annexe 6

### Bibliographie et références des sources citées dans ce mémoire

**N.B.** : les références bibliographiques sont établies selon le fascicule de documentation **Z44 -005** de l'AFNOR pour les documents imprimés

#### Monographies

ACCART, Jean-Philippe, RETHY, Marie-Pierre. *Le métier de documentaliste*. 2<sup>ème</sup> éd. Paris : Electre-Ed. du Cercle de la librairie, 2003. 451 p. (Bibliothèque-Documentation). ISBN 2-7664-0872-6

AFNOR : Voir : ASSOCIATION FRANCAISE DE NORMALISATION

ANDRIEU, Olivier. *Google*. Paris : Micro Application, 2004. 191 p. (Guide Microapp) ISBN 2-7429-3092-2

ANDRIEU, Olivier. *Réussir son référencement web*. Paris : Eyrolles, 2007. 302 p. ISBN 978-2-212-12264-0

ARISTOTE. *Organon IV : les seconds analytiques*. Nouv. éd. Paris : Vrin, 1970. 250 p. (Bibliothèque des textes essentiels)

ASSOCIATION DES BIBLIOTHECAIRES FRANCAIS. *Le métier de bibliothécaire*. 11<sup>ème</sup> éd. Paris : Electre-Ed. du Cercle de la librairie, 2003. 454 p. (Bibliothèque-Documentation). ISBN 2-7654-0866-6

ASSOCIATION FRANCAISE DE NORMALISATION. (Afnor). *Documentation*, Vol.1. Editions de l'Afnor, 1986. 286 p.

ASSOCIATION FRANCAISE DE NORMALISATION (Afnor). *Vocabulaire de la documentation*. 2<sup>ème</sup> éd. Paris : La Défense : Afnor, 1987

ASTOLFI, Jean-Pierre. *L'erreur : un outil pour enseigner*. Issy les Moulineaux : ESF, 2008. 117 p. (Pratiques et enjeux pédagogiques) ISBN 978-2-7101-0212

ASTOLFI, Jean-Pierre, DEVALAY, Michel. *La didactique des sciences*. 5<sup>ème</sup> éd. Paris : Presses universitaires de France, 1989. 127 p. (que sais-je ; 2448) ISBN 2-13-043469 -X

AXELOS, Kostas. *Réponses énigmatiques : faille, percées*. Paris : Ed. de Minuit, 2005. 94 p. (Arguments) ISBN 2-7073-1899-X

BAKHTINE, Mikhaïl. *Le marxisme et la philosophie du langage : essai d'application de la méthode sociologique en linguistique*. Paris, Minuit, 1977. 233 p. (Le sens commun) ISBN 2-7073-0151-2

BALTZ, Claude. In-formation. In *Hypertextes et hypermédias : créer du sens à l'ère numérique*, actes de H<sup>2</sup>PTM'03, 24 sept.-26 sept. 2003, Université Paris VIII, Paris : Hermès Sciences, 2003, p.363-377

BARAQUIN, Noëlla, BAUDART, Anne. [et al.] *Dictionnaire de philosophie*. 3<sup>ème</sup> éd. rev. et augm. Paris : A. Colin, 2005. 377 p. ISBN 2-200-26986-2

- BAYLON, Christian, MIGNOT, Xavier. *La communication*. 2<sup>ème</sup> éd. Paris : Armand Colin, 2005. 416 p. (fac) ISBN 2-200-34408-2
- BEAUDIQUEZ, Marcelle. *Guide de bibliographie générale : méthodologie et pratique*. Nouv. éd. Paris : München : London : K.G. Saur [etc.], 1989. 277 p. ISBN 3-59820-454X
- BESNIER, Jean-Michel. *Les théories de la connaissances : un exposé pour comprendre, un essai pour réfléchir*. Paris : Flammarion, 1996. 127 p. (Dominos) ISBN 2-08-035458-2
- BETHERY, Annie. *Abrégé de la classification décimale de Dewey*. Nouv. éd. Paris : Electre-Editions du Cercle de la Librairie, 1998. 317 p. (Bibliothèques) ISBN 2-7654-0682-0
- BLANC-MONTMAYEUR, Martine, DANSET, Françoise. *Choix de vedettes matières à l'intention des bibliothèques publiques*. Paris : éditions du Cercle de la Librairie, 1984. 283 p. (Bibliothèques) ISBN 2-7654-0318-X
- BONNEROT, Jean. Avant-propos. In COTTON DES HOUSSAYES, J.-B. *Des devoirs et des qualités du bibliothécaire: discours prononcé dans l'Assemblée générale de Sorbonne le 28 décembre 1780*. Paris : E. Baudelot imprimeur, 1951 [p] 6-9
- BOUGNOUX, Daniel. *Sciences de l'information et de la communication*. Paris : Larousse, 1993. 808 p. (Textes essentiels) ISBN 2-03-741010-7
- BRETON, Philippe. *Eloge de la parole*. Paris : La Découverte, 2003. 190 p. ISBN 2-7071-3600-X
- BRETON Philippe, PROULX, Serge. *L'explosion de la communication : introduction aux théories et pratiques de la communication*. Paris : La Découverte, 2006. 382 p. (Grands repères. Manuels) ISBN 2-7071-4808-3
- BRIANT, Vincent de, PALAU, Yves. *La médiation : définitions, pratiques et perspectives*. Paris : Nathan, 1999. 128 p. (Collection 128) ISBN 2-09-191052-X
- BRUN, Jean. *Héraclite ou le philosophe de l'éternel retour*. Paris : Seghers, 1969. 188 p. (Philosophes de tous les temps)
- CALCALY, Serge, (dir.). *Dictionnaire de l'information*. 3<sup>ème</sup> éd. Paris : Armand Colin, 2008. 295 p. ISBN 978-2-200-35132-8
- CALENGE, Bertrand. *Accueillir, orienter, informer : l'organisation des services aux publics dans les bibliothèques*. Paris : Electre-Ed. du Cercle de la Librairie, 1996. 429 p. (Bibliothèques) ISBN 2-7654-0625-1
- CASSIN, Barbara. *Google-moi : la deuxième mission de l'Amérique*. Paris : Albin Michel, 2006. 258 p. (Banc public) ISBN 978-2-226-17259-4
- CHANGEUX, Jean-Pierre. *L'homme de vérité*. Paris : O. Jacob, 2004. 402 p. (Odile Jacob poches) ISBN 2-7381-1383-4
- CHAUMIER, Jacques. *La technique documentaire au fil de l'histoire : 1950-2000*. Paris : ADBS, 2003. 179 p. (Sciences de l'information .Série Etudes et techniques) ISBN 2-84365-064-X
- CHAUMIER, Jacques. *Travail et méthodes du documentaliste : pour une exploitation méthodique et optimale de l'information*. 7<sup>ème</sup> éd. Issy les Moulineaux : ESF, 2007. 174 p. (Collection formation permanente. Série entreprise) ISBN 978-2-7101-1892-3
- CORROY, Laurence, GONNET, Jacques. *Dictionnaire d'initiation à l'info-com*. Paris : Vuibert, 2008. 351 p. ISBN 978-2-7117-2560-1

COTTIN, Stéphane, MOYRET, Sophie. *Petit guide d'accès à l'information juridique française : pratique de la recherche documentaire juridique* Paris : ADBS, [S.d.]. 108 p. ISBN 2-84365-036-4

Dalbin, Sylvie. Instruments de recherche sur le web. In *La recherche d'information sur les réseaux : cours INRIA, 30 septembre- 4 octobre 2002*, Le Bono, Morbihan, p.46-50

DAMASIO, Antonio R. *L'erreur de Descartes : la raison des émotions*. Paris : O. Jacob, 2008. 365 p. (Bibliothèque) ISBN 978-2-7381-2070-0

DARROBERS, Martine, LE POTTIER, Nicole. *La recherche documentaire*. Paris : Nathan, 2005. 159 p. (Repères pratiques ; 25). ISBN 2-09-183208-1

DARROBERS, Martine, LE POTTIER, Nicole. *La recherche documentaire*. Paris : Nathan, 2000. 159 p. (Repères pratiques ; 25). 2-09-182481-X

DEBRAY, Régis. *Introduction à la médiologie*. Paris : Presses universitaires de France, 2000. 223 p.(Premier cycle) ISBN 2-13-050105-2

DEJOURS, Christophe. *Travail, usure mentale : essai de psychopathologie du travail*. Nouv. éd. Paris : Bayard, 2008. 298 p. ISBN 978-2-227-47780-3

DELEUZE, Gilles, GUATTARI, Félix. *Mille plateaux : les postulats de la linguistique*. Paris : éd. de Minuit, 1980.

DENNECKER, Claire, KOLMAYER, Elisabeth. *Éléments de psychologie cognitive pour les sciences de l'information : avec exercices corrigés*. Villeurbanne (Rhône): Presses de l'Enssib, 2006. 334 p. (Les Cahiers de l'Enssib ; 4) ISBN 978-2-910227-65-4

*Dictionnaire historique de la langue française : M-Z*. Paris :Dictionnaire La Robert, 1992, 2383 p. ISBN 2-85036-187-9

DION, Emmanuel. *Invitation à la théorie de l'information*. Paris : éditions du Seuil, 1997. 156 p. (Points Sciences). ISBN 2-02-029940-2

DRILLON, Jacques. *Traité de la ponctuation française*. Paris : Gallimard, 1991. 472 p. (Tel) ISBN 2-07-072198-1

DUPRIEZ, Bernard. *Gradus : les procédés littéraires : dictionnaire*. Paris : 10 /18, 1984. 541 p. ISBN 2-264-00203-4

EI-BEZE, Marc. Systèmes de question-réponse. In *La compréhension des langues et interaction*. Paris : Hermès-Sciences publications, Chap.10, avril 2006, p.277-297

ERMINE, Jean-Louis. *Les systèmes de connaissance*. 2<sup>ém</sup> éd. rev. et augm. Paris : Hermès science Publications-Lavoisier, 2000. 236 p. ISBN 2-7462-0159-3

ERTZSCHEID, Olivier. Syndrome d'Elpenor et sérendipité : deux nouveaux paramètres pour l'analyse de la navigation hypermédia. In *Hypertextes et hypermédiat : créer du sens à l'ère numérique*, actes de H<sup>2</sup>PTM'03, 24 sept.-26 sept. 2003, Université Paris VIII, Paris : Hermès Sciences, 2003, p.133-142

FAVIER, Laurence, MARTIN-JUCHAT, Fabienne. La science de l'information face à de nouveaux paradigmes : prise en compte de la dimension sociale de la recherche d'information et remise en cause de la figure de l'utilisateur. In *Recherches récentes en sciences de l'information : convergences et dynamiques. Actes du colloques MICS-LERASS, 21-22 mars Toulouse*. Paris : ADBS, 2002, p.255-267

GALIPIENSO NJONGOUE, Claire. *Le développement des services de renseignement à distance à la bibliothèque Sainte Geneviève : perspectives d'aide à la décision*. Villeurbanne : Enssib. Projet professionnel personnel, 144 p. 2007

GANASCIA, Jean-Gabriel. *Les sciences cognitives : un exposé pour comprendre, un essai pour réfléchir*. Paris : Flammarion, 1996. 127 p. (Dominos) ISBN 2-08-035418-3

GARFINKEL, Harold. *Recherches en ethnométhodologie*. Paris : Presses universitaires de France, 2007. 473 p. (Quadrige. Grands textes) ISBN 978-2-13-056150-7

GARRIC, Nathalie, CALAS, Frédéric. *Introduction à la pragmatique*. Paris : Hachette supérieur, 2007. 2007 p. (Hu linguistique) ISBN 978-2-01-145716-5

GAULEJAC, Vincent de. *La société malade de la gestion : idéologie gestionnaire, pouvoir managérial et harcèlement moral*. Paris : Seuil, 2005. 275 p. (Economie humaine) ISBN 2-02-068912-X

Ghernaoui-Hélie, Solange, Dufour, Arnaud. *De l'ordinateur à la société de l'information*. 2<sup>ème</sup> éd. Paris : Presses universitaires de France, 2001. 125 p. ( Que sais-je ? ) ISBN 2-13-051721-8

GHOUS DZIRI, Anissa. Ontologies et thésaurus : des systèmes d'organisation des connaissances différents et complémentaires. In *Indices, index, indexation : actes du colloque international organisé les 3 et 4 novembre 2005 à l'université Lille 3 par les laboratoires CERSATES et GERICO*, p. 191-202

GODIN, Christian. *La philosophie pour les nuls*. Paris : First, 2006. 537 p. (Pour les nuls) ISBN 2-87691-998-2

*Grand dictionnaire de la psychologie*. Paris : Larousse, 1991. 862 p. ISBN 2-03-501030-6

GRAU, Brigitte. L'évaluation des systèmes de question-réponse. In *Evaluation des systèmes de traitement de l'information*. Paris : Hermès science Publication-Lavoisier, 2004. [chap.3], p. 80-98.

GRAU, Brigitte. Les systèmes de question-réponse. In *Méthodes avancées pour les systèmes de recherche d'information*, Paris : Hermès science Publication-Lavoisier, 2004. Chap. 9 p. 189-218,

GRAWITZ, Madeleine. *Méthodes des sciences sociales*. 8<sup>ème</sup> éd. Paris : Dalloz, 1991, 1140. (Précis Dalloz) ISBN 2-247-01100-4

GUINCHAT, Claire, SKOURY, Yolande. *Guide pratique des techniques documentaires : traitement de l'information*, Vol. 2. Nouv. éd. Vanves : EDICEF/AUPELF, 1996. 302p. ISBN 2-841-29205-3

HALL, Edward T. *La dimension cachée*. Paris : éditions du Seuil, 1978. 254 p. (Points civilisation) ISBN 2-02-004776-4

HOLZEM, Maryvonne. *Terminologie et documentation : pour une meilleure circulation des savoirs*. Paris ADBS, 1999. 292 p. (Sciences de l'information Série Recherche et documents) ISBN 2-84365-032-1

*Les hypermédias : conception et réalisation*. (Imad Saleh, dir.). Paris : Hermès sciences Publication-Lavoisier, 2005. 327p. (Cognition et traitement de l'information) ISBN 2-7462-1037-1

IHADJADENE, Madjid. *La recherche et la navigation dans un système d'information grand public : le cas des hypercatalogues sur internet*. Thèse de Doctorat de Sciences de l'Information et de la Communication. Université Claude Bernard-Lyon1, 1999.

IPPOLITA. *La face cachée de Google*. Paris : Payot, 2008. 236 p. (Manuels Payot) ISBN 978-2-228-90288-5

JACQUESSON, Alain, RIVIER, Alexis. *Bibliothèques et documents numériques : concepts, composantes, techniques et enjeux*. Nouv. éd. Paris : Electre-Editions du Cercle de la Librairie, 2005. 573 p. (Bibliothèques) ISBN 978-2-7654-0915-3

KILBORNE, Yann. *Regarder le monde : filmer l'invisible, approche ethnométhodologique de la démarche documentaire*. Paris : Les Presses du Léma, 2006. 209 p.

KUHLMANN, Marie, KUNTZMANN, Nelly, BELLOUR, Hélène. *Censure et bibliothèques au XX<sup>e</sup> siècle*. Paris : Editions du Cercle de la Librairie, 1989. 349 p. (Bibliothèques) ISBN 2-7654-0418-6

LAHARY, Dominique. Techniques de base relatives à l'organisation d'une collection et à son accès par les usagers. In *Le métier de bibliothécaire*. Paris : Electre-Editions du Cercle de la Librairie, 2007, p.229-260

LAHARY, Dominique. Des bibliothèques traditionnelles aux « bibliothèques virtuelles ». In *La recherche d'information sur les réseaux : cours INRIA, 30 septembre- 4 octobre 2002*, Le Bono, Morbihan, p.178

LALANDE, André. *Vocabulaire technique et critique de la philosophie*. 16<sup>è</sup>m éd. Paris : Presses universitaires de France, 1991. 1323p. ISBN 2-13-043911X

LALLICH-BOIDIN, MARET, Dominique. *Recherche d'information et traitement de la langue : fondements linguistiques et applications*. Villeurbanne : Presses de l'Enssib, 2005. 288 p. (Les Cahiers de l'Enssib ; 3) ISBN 2-910227-60(X)

LARDY, Jean-Pierre. *Recherche d'information sur internet : outils et méthodes*. 7<sup>è</sup>m éd. Paris : ADBS, 2001. 123 p. (Sciences de l'information. Série Recherches et documents) ISBN 2-84365-052-6

LAWSON-MATHIEU, Eléonore, BARSANTI, Emily. *Guide des sources d'information : 2006-2007*. Paris : Victoires Editions, 2006. 622 p. (Métier Journaliste) ISBN 2-35113-015-4

LAZAR, Judith. *100 mots pour introduire aux théories de la communication*. Paris : Les Empêcheurs de penser en rond, 2004. 259 p. ISBN 2-84671-086-4

LECLERC-REYNAUD, Sylvie. *Pour une documentation créative : l'apport de la philosophie de Raymond Ruyer*. Paris : ADBS, 2006. 193 p. (Sciences et techniques de l'information) ISBN 2-84365-085-2

LE COADIC, Yves François. *Le besoin d'information : formulation, négociation, diagnostic*. 2<sup>è</sup>m éd. Paris : ADBS, 2007. 204 p. (Collection Sciences de l'information.).ISBN 978-284365-090-0

LE COADIC, Yves François. *La science de l'information*. Paris Presses universitaires de France, 1994. 126 p. (Que sais-je ?2873) ISBN 2-13-046381-9

LE COADIC, Yves F. *Usages et usagers de l'information*. Paris : ADBS : Nathan université, 1997.127 p. (collection 128) ISBN 2-09-190366-3

LEFEVRE, Philippe. *La recherche d'informations : du texte intégral au thésaurus*. Paris : Hermès science Publication, 2000. 253 p. ISBN 2-7462-0173-9

LENORMAND, Patrick. *Internet : technique de recherche pour les professionnels*. Saint Herblain (44 800) : ENI, 2007. 305 p. (Objectif solutions) ISBN 978-2-7460-3498-3

LEVY, Pierre. *Les technologies de l'intelligence : l'avenir de la pensée à l'ère informatique*. Paris : Seuil, 1993. 233 p. (Points sciences) ISBN 2-02-013091-2

LEVY-LEBLOND, Jean-Marc. *Impasciences*. Paris : Seuil, 2003. 236 p. (Points Sciences) ISBN 2-02-054137-8

LIBERA, Antoine de. *La philosophie médiévale*. Paris : Presses universitaires de France, 1992, 127 p. (Que sais-je ; 1044) ISBN 2-13-044212-9

LINE, Maurice B. Les bibliothécaires et la connaissance : gardiens, fournisseurs, intermédiaires, guides, ressources ? In *Bibliothécaire, quel métier ?* Paris : Electre-Éditions du Cercle de la Librairie, 2004, p. 41-51

MALCLES, Louise-Noëlle, LHERITIER, Andrée. La bibliographie. 5<sup>ème</sup> éd. Paris : Presses universitaires de France, 1989. 127 p. (Que sais-je ? 708) ISBN 2-13-042292-6

MARC, Edmond, PICARD, Dominique. *L'école de Palo Alto : un nouveau regard sur les relations humaines*. Paris : Retz, 2000. 224 p. (Psychologie dynamique) ISBN 2-7256-1456-2

MARESCA, Bruno. *Les bibliothèques municipales en France après le tournant Internet : attractivité, fréquentation et devenir*. Paris : Bibliothèque publique d'information : Centre Pompidou, 2007. 283 p. (Études et recherche) ISBN 978-2-84246-103-4

MARTINET, André. *Éléments de linguistique générale*. 3<sup>ème</sup> éd. Paris : Armand Colin, 1991. 221 p. (U Prisme. Langages) ISBN 2-200-32208-9

MATHELOT, Pierre. *L'informatique*. 10<sup>ème</sup> éd. Paris : Presses universitaires de France, 1998. 127 p. (Que sais-je ? 1371) ISBN 2-13-047169-2

MATTELART, Armand. *Histoire de la société de l'information*. 3<sup>e</sup> éd. Paris : La Découverte, 2006. 122 p. (Repères) ISBN 2-7071-4866-0

MERLEAU-PONTY, Maurice. *Phénoménologie de la perception*. Paris : Gallimard, 1945. p

MESGUICH, Véronique, THOMAS, Armelle. *Net recherche : le guide pratique pour mieux trouver l'information utile*. Paris : ADBS, 2006. 148 p. (Sciences et techniques de l'information) ISBN 2-84365-083-6

MIEGE, Bernard. *L'information- communication : objet de connaissance*. Bruxelles : De Boeck ; Bry sur Marne : INA, 2004. 248 p. (Médias Recherches). ISBN 2-8041-4668-5

MILON, Alain. *La valeur de l'information : entre dette et don, critique de l'économie de l'information*. Paris : Presses universitaires de France, 1999. 232 p. (Sociologie d'aujourd'hui) ISBN 2-13-049740-3

MKADMI, Abderrazak, SALEH, Imad. *Bibliothèque numérique et recherche d'informations*. Paris : Hermès-sciences publications : Lavoisier, 2008. 281 p. (Collection information, hypermédia et communication) ISBN 978-2-7462-1820-8

MORIZIO, Claude. *La recherche d'information*. Paris : ADBS, 2002. 126 p. (Information documentation; 267). ISBN 2-09-191244-1

MOUREN, Raphaële, PEIGNET, Dominique. Histoire du livre et des bibliothèques. In *Le Métier de bibliothécaire*. 12<sup>ème</sup> éd. Paris : Electre-Éditions du Cercle de la Librairie, 2007 . p. 17-36

MUSIL, Robert. *L'homme sans qualités : tome 1*. Paris : éditions du Seuil, 1995. 637 p. (Points) ISBN 2-02-023815-2

NASSE-KOLMAYEUR, Elisabeth. *Contribution à l'analyse des processus cognitifs mis en jeu dans l'interrogation d'une base de données documentaire*. Thèse de doctorat en psychologie. Université René Descartes-Paris V, 1997.

NGUYEN, Claire. *Mettre en place et développer un service de référence virtuelle dans une perspective d'intégration à un réseau collaboratif : analyses comparées et propositions à partir de l'exemple de l'Université de Montréal*. Villeurbanne : Enssib, Diplôme de conservateur de bibliothèque. 120 p. Mémoire d'étude : Enssib, 2006

NORDON, Didier. *Deux et deux font-ils quatre ? Sur la fragilité des mathématiques*. Paris : Pour la science, 1999. 111 p. ISBN 2-84245-011-6

PEDAUQUE, Roger T. *La redocumentarisation du monde*. Toulouse : Cepaduès-éditions, 2007. 213 p. ISBN 978-2-85428-728-8

PERRIN, Georges (dir.). *Développer et exploiter un fonds spécialisé*. Villeurbanne : Enssib, 1999. 158 p. (La boîte à outils) ISBN 2-910227-25-1

PIOLLAT, Annie. *La recherche documentaire : manuel à l'usage des étudiants, doctorants et jeunes chercheurs*. Marseille : Solal éditeur, 2002. 150 p. (Psychologie. Théories, méthodes, pratiques). ISBN 2-914513-30-5

PLATON. *Ménon*. Paris : Flammarion, 1991. 350 p. (Gf Flammarion) ISBN 2-08-070491-5

POCHET, Bernard. *Méthodologie documentaire : comment accéder à la littérature scientifique à l'heure d'Internet*. Bruxelles : De Boeck université, 2002. 141 p. (Sciences et méthodes) ISBN 2-8041-4142-X

*Pourquoi les manchots n'ont pas froid aux pieds ? et 111 questions stupides et passionnantes*. Paris : Seuil, 2007, 195 p. (Science ouverte) ISBN 2-02-092599-0

PRAX, Jean-Yves, LARCHER, Simon. *La gestion électronique documentaire*. 3<sup>ème</sup> Ed. Paris : Dunod, 2004. 341p. ISBN 2-10-007891-7

PREVOTEAU, Marie-Hélène, UTARD, Jean-Claude. *Manuel de bibliographie générale*. Nouv. éd. Paris : Electre-Ed. du Cercle de la librairie, 2005. 524 p. (Bibliothèques). ISBN 2-7654-0912-9

RICHTER, Noé. *Les langages documentaires encyclopédiques : guide pratique d'indexation...* Marigné (72 220) : les éditions de la queue du chat, 1990. 172 p.

RIVIER, Alexis. *Aide-mémoire d'informatique documentaire*. Paris : Ed. du Cercle de la Librairie, 2007. 156 p. (collection Bibliothèques). ISBN 978-2-7654-0953-3

SARACEVIC, T. The concept of "relevance" in Information Science: a historical review, *Introduction to Information Science*, Londres: Bowker, 1970

SAVOY, Jacques. La toile et ses moteurs de recherche. In *Indices, index, indexation : actes du colloque international organisé les 3 et 4 novembre 2005 à l'université Lille 3 par les laboratoires CERSATES et GERICO*, p. 49-59

SUCH, Marie-France, PEROL, Dominique. *Initiation à la bibliographie scientifique*. Paris : Promodis : Cercle de la Librairie, 1987. 302 p. ISBN 2-903181-4

TAMBA-MECZ, Irène. *La sémantique*. Paris : Presses universitaires de France, 1991. 127 p. (Que sais-je ? 655) ISBN 2-13-043316-2

THIVANT, Eric. Les pratiques de recherche d'information pour la conception de bons d'option. In *Recherches récentes en sciences de l'information : convergences et dynamiques : actes du colloque international organisé les 21 et 22 mars 2002 à Toulouse par l'équipe Médiations en information spécialisées (MICS) du Laboratoire d'études et de recherches appliquées en sciences sociales (LERASS) de l'Université Toulouse 3*, p.57-72

TRAVERSO, Véronique. L'analyse des conversations. Paris : Nathan, 1999. 128 p. (128. Linguistique) ISBN 2-09-191048-1

TRICOT, André ROUET, Jean-François. Avant propos In *Les hypermédias : approches cognitives et ergonomiques*, Paris : Hermès, 1998, p .7-11

TRICOT, André ROUET, Jean-François. Chercher de l'information dans un hypertexte : vers un modèle de processus cognitifs. In *Les hypermédias : approches cognitives et ergonomiques*, Paris : Hermès, 1998, p .57-74

WALLER, Suzanne, MASSE, Claudine. *L'analyse documentaire : une approche méthodologique*. Paris : ADBS, 1999. 319 p. (Sciences de l'information Série études et techniques) ISBN 2-84365-030-5

WITTGENSTEIN, Ludwig. *Le cahier bleu et le cahier brun*. Paris : Gallimard, 1996. 313 p. (Bibliothèque de philosophie) ISBN 2-07074018-8

WITTGENSTEIN, Ludwig. *Tractacus logico-philosophicus*. Paris : Gallimard, 2001. 121 p. (Tel) ISBN 2-07075864-8

WRIGHT, Patricia. L'utilisation de documents dans la prise de décision. In *Les hypermédias : approches cognitives et ergonomiques*, Paris : Hermès, 1998, p .77-93

ZOURABICHVILI, François. Le vocabulaire de Deleuze. Paris : Ellipses, 2004. 95 p. (Vocabulaire de...) ISBN 2-7298-1291-1

## Reuves

Bertrand, Annick. Compréhension et catégorisation dans l'indexation des documents. *Cahiers du LCPE*. Linguistique et ergonomie.1994, n°1, pp 53-71.

Eboli, Gilles. Bibliothèques hybrides. *Bibliothèques(s)*, déc.2007, n°36, p. 9

Di Pietro, Christelle, Calenge, Bertrand. Le guichet du Savoir : répondre aux demandes de contenu. *BBF, Bulletin des bibliothèques de France*, 2005, n°04, p. 38-42

Heurtematte, Véronique. Journée d'études ABF : comment prend la greffe. *Livres hebdo*, déc. 2007, n°714, p.59

Papy, Fabrice, Chauvin Sophie. Au-delà de la transfiguration du catalogue : le Visual...Catalog. *BBF, Bulletin des bibliothèques de France*, 2005, n°04, p.5-12

Sciences et techniques. *Bibliothèque(s)*, oct, 2007, n°34-35, p.10

## Ressources électroniques

**N.B.** : le référencement des ressources électronique est établi selon la norme **ISO 690-2**

**Bruit** : TLF : *Trésor de la langue française*. 2008. [en ligne]. [réf. du 21-03-2008] Disponible sur le World Wide web : <<http://atilf.atilf.fr/dendien/scripts/tlfiv5/advanced.exe?8;s=1153135500> >

**Enigme** : TLF : *Trésor de la langue française*. 2008. [en ligne]. [réf. du 31-03-2008] Disponible sur le World wide web : <<http://atilf.atilf.fr/dendien/scripts/tlfiv5/search.exe?23;s=3760148655;cat=0;m=%82nigme> >

**Lamy**, Jean-Philippe. L'entretien de référence : Une notion à introduire dans un enseignement renoué », *BBF*, 1998, n° 6, p. 52-58 [en ligne] [réf. du 14 avril 2008]. Disponible sur le World wide web :< <<http://bbf.enssib.fr>>

**Réponse** : *TLF : Trésor de la langue française*. 2008. [en ligne]. [réf. du 21-03-2008]. Disponible sur le World wide web :< <http://atilf.atilf.fr/dendien/scripts/tlfiv5/visusel.exe?28;s=1153135500;r=2:nat=;sol=2>>