

+++++

INTRODUCTION

+++++

Savez-vous quelles ont été les conséquences majeures de l'encéphalite spongiforme bovine sur la filière agroalimentaire française ? Quel est le meilleur lycée de la ville de Strasbourg ? Quel est le moyen le moins onéreux pour aller d'Amiens à Lille par les transports en commun ? Quels risques présente l'amniocentèse ? Ces questions vous semblent probablement difficiles, et leur réponse paraît requérir des connaissances assez spécifiques. Il est probable que peu d'êtres humains connaissent la réponse à ces quatre questions. Pourtant, en quelques minutes, vous pouvez obtenir les réponses. Il vous faut pour cela rechercher de l'information. Cela ne devrait pas être très facile. Il y a vingt ans, ces réponses étaient à peu près inaccessibles directement : il vous aurait fallu interroger plusieurs autres êtres humains, qui connaissaient une des réponses, ou qui sachent utiliser des outils spécifiques de recherche d'information.

Ainsi, les êtres humains se retrouvent souvent dans des situations où ils manquent de connaissances pour comprendre ou pour agir. Ils ont alors quatre grandes catégories de solutions :

- ils renoncent à comprendre ou à atteindre leur but, parfois sans avoir pris conscience de leur manque de connaissances ;
- ils essaient, tâtonnent, font des erreurs, essaient encore, pour finir par comprendre ou agir par eux-mêmes : ils résolvent le problème ;
- ils demandent de l'aide à une autre personne, une experte, un professeur, une collègue : ils dialoguent ;
- ils recherchent de l'information, dans un document, sur le Web, dans une base de données.

La recherche d'information est donc une activité extrêmement commune pour les êtres humains, notamment ceux qui vivent dans les riches sociétés de l'information. Chacune d'entre nous, plusieurs fois par jour, recherche des informations, que ce soit dans le cadre académique, professionnel, ou dans celui de la vie quotidienne, culture, loisir, santé, tourisme, achats, etc. Nous recherchons un numéro de téléphone, un horaire de train, la défini-

tion d'un mot, une recette de cuisine. La plupart du temps, nous trouvons ce que nous cherchons, et ne prêtons pas plus d'attention que cela à cette activité.

Pourtant, depuis une vingtaine d'années, un bouleversement très profond a été amorcé : la mise à la disposition du grand public d'outils de recherche d'information extrêmement puissants. Ces outils étaient autrefois réservés à des spécialistes, des professionnels ayant suivi une formation (les documentalistes, les bibliothécaires). Puis, pendant que ces outils se sophistiquaient de façon extraordinaire, ils ont été largement diffusés. Si bien qu'aujourd'hui, la source documentaire probablement la plus volumineuse¹ de l'histoire de l'humanité, les documents Web, est à la portée de chacun, du jeune élève, de l'adolescent, de l'étudiante, de l'adulte ou de la personne âgée. Mais la plupart de ces honnêtes gens n'ont pas été formés, ou si peu, à la recherche d'information. Comment font-ils ? Quelles sont leurs démarches, leurs difficultés ? Pourquoi procèdent-ils de la sorte ? De leur côté, que savent faire les spécialistes de la recherche d'information, les documentalistes, les bibliothécaires ? Pourquoi y arrivent-ils mieux que les autres ? Leurs compétences peuvent-elles être aisément analysées et enseignées ? Existerait-il d'autres moyens de permettre aux honnêtes gens de chercher et de trouver les informations qu'ils souhaitent ?

Toutes ces questions se posent aujourd'hui aux chercheurs qui consacrent leurs travaux à l'activité de recherche d'information. Ainsi, la recherche dans le domaine des documents et des systèmes d'information a d'abord été à l'origine des progrès technologiques qui ont permis la révolution documentaire de la fin du xx^e siècle. Cette révolution a été tellement importante, elle a tellement bouleversé les pratiques informationnelles et les organisations sociales et professionnelles de l'information-documentation, qu'aujourd'hui un courant de recherche se développe considérablement : celui de l'étude de l'activité humaine de recherche d'information, plus connu sous l'expression de « paradigme usager ».

L'objectif de cet ouvrage est de présenter ce domaine de recherche, de recenser les résultats les plus significatifs, de décrire les méthodes utilisées, en rendant compte de la diversité des disciplines qui, au sein des sciences humaines, traitent de ces questions.

Ce domaine est tellement vaste que, avant de le présenter, nous proposons au lecteur francophone quelques définitions de la recherche d'in-

1. Environ 80 millions de sites actifs en février 2010 selon la société Netcraft < news.netcraft.com >. Sites pouvant contenir un nombre considérable de documents...

formation. Nous nous efforçons de clarifier la façon dont les termes anglo-saxons du domaine ont été traduits en français. De fait, les premières délimitations terminologiques, et donc les premières constructions théoriques de l'activité de recherche d'information, proviennent des travaux en langue anglaise conduits au sein de la *Library and Information Science (LIS)*. Après ces remarques préliminaires, nous donnons l'indispensable description des besoins d'information et tâches de recherche d'information qui, non seulement, sont les points d'ancrage communs aux deux disciplines fournissant l'essentiel des théorisations sur la recherche d'information, psychologie et LIS, mais surtout traitent les deux questions fondamentales qui doivent être posées en début : pourquoi et que rechercher ?

QU'EST-CE QUE LA RECHERCHE D'INFORMATION ? QUELQUES DÉFINITIONS

La réponse à cette question peut paraître aujourd'hui évidente à la chercheuse d'information : saisir des mots-clés dans un moteur de recherche du Web et trouver des documents correspondant à sa demande. Elle n'est pas si éloignée de la façon dont se présentait la question de la recherche d'information dans les travaux scientifiques portant sur les premiers systèmes d'information informatisés. Une simple demande pour un seul usage, *one query, one use*, selon la formulation lapidaire de [Bates 05a] qui résume la réduction de la question opérée jusqu'aux années soixante-dix. La complexité réside alors essentiellement dans le fonctionnement d'un système capable de fournir des résultats pertinents, et non dans les pratiques des usagers de ces systèmes. Les chercheurs ont rencontré dans les décennies suivantes d'autres questionnements propres aux interactions des usagers avec les systèmes et l'information, découvrant une activité informationnelle composée d'une multiplicité de traits susceptibles de s'agencer diversement. Les questionnements successifs ont fait émerger un nombre important de représentations théoriques de l'activité de recherche d'information, ce qui a conduit récemment les chercheurs à recenser les dénominations de l'activité, particulièrement dans les pays anglo-saxons qui disposent d'une riche terminologie pour qualifier l'activité, et à les fixer en leur adjoignant des définitions permettant de formuler rigoureusement les liens entre diverses dimensions de l'activité informationnelle. Avant d'en venir aux définitions scientifiques de la recherche d'information énoncées dans la sphère linguistique anglaise,

nous présentons les définitions données dans le domaine francophone, issues de travaux en grande partie professionnels.

Dénomination et définitions techniques

Les deux versants de la recherche d'information, activité du chercheur d'information vs fonctionnement du système d'information, ne sont pas toujours aisés à distinguer, tant il est difficile d'abstraire l'activité du chercheur d'information de celle des systèmes d'information. Les premières définitions, que nous qualifierons de techniques pour les distinguer des définitions émanant de travaux scientifiques, se confrontent à cette difficulté. Les définitions techniques font l'objet de normes élaborées par des experts, souvent issus d'associations de professionnels, au sein de l'Association française de normalisation (Afnor). Deux définitions sont données, la recherche documentaire et la recherche d'information, distinguées à partir du type de documents recherchés [Afnor 87].

La « recherche documentaire » est définie comme « Action, méthodes et procédures ayant pour objet de retrouver dans des fonds documentaires les références des documents pertinents ». La recherche documentaire est donc une recherche de documents secondaires. La façon de les rechercher est décrite dans des termes très généraux, action, méthodes et procédures qui ne permettent pas de décider si la définition envisage le fonctionnement du système documentaire ou l'action humaine de rechercher.

La « recherche d'information » est présentée comme « Action, méthodes et procédures ayant pour objet d'extraire d'un ensemble de documents les informations voulues. Toute opération (ou ensemble d'opérations) ayant pour objet la recherche, la collecte et l'exploitation d'informations en réponse à une question sur un sujet précis ». La recherche d'information est représentée sommairement en trois étapes. Elle s'effectue au sein de documents primaires. La même ambiguïté dans la formulation de la définition ne permet pas d'identifier s'il s'agit de la description d'un système ou d'une pratique.

Ces définitions techniques sont antérieures à l'avènement d'Internet. Elles correspondent aux modalités traditionnelles de recherche, interrogations distinctes d'un système bibliographique (la recherche documentaire) et d'un système de recherche d'information, banque de données de documents en texte intégral (la recherche d'information). Elles paraissent désormais inappropriées, compte tenu de l'évolution des systèmes d'information, regroupant à l'instar des moteurs du Web les deux types de recherche. Plus

fondamentalement, elles reflètent un état de la question qui méconnaît le processus de recherche tel que conduit par le chercheur d'information. Plus récemment, un groupe de travail réunissant membres de l'Afnor et de l'Institut national des techniques documentaires (INTD) [Boulogne 04] s'est proposé d'actualiser ces définitions. Toutefois, ces derniers travaux ne sont pas homologués. Les définitions proposées dans ce cadre s'efforcent de prendre en compte les évolutions des environnements documentaires, et de caractériser plus précisément la recherche d'information. Sur ce plan, les deux définitions, recherche documentaire et recherche d'information, sont maintenues. Cependant, malgré l'insertion de l'utilisateur dans la seconde formulation, il reste peu aisé d'identifier ce qui appartient à une description des systèmes ou à celle des pratiques.

La « recherche documentaire » consiste en un « Ensemble des méthodes, procédures et techniques ayant pour objet de retrouver des références de documents pertinents (répondant à une demande d'information) et les documents eux-mêmes ». La définition inclut désormais recherche de documents primaires et/ou secondaires et tertiaires.

La définition de la « Recherche d'information » fait référence à tout type d'environnements documentaires et aux phases du processus de recherche : « Ensemble des méthodes, procédures et techniques permettant, en fonction de critères de recherche propres à l'utilisateur, de sélectionner l'information dans un ou plusieurs fonds de documents plus ou moins structurés ». Toute recherche d'information suppose trois phases successives :

- (a) une recherche bibliographique des références de documents pertinents ;
- (b) une recherche documentaire, c'est-à-dire une recherche bibliographique complétée par la recherche (l'acquisition) des documents eux-mêmes ;
- (c) et enfin le repérage de l'information dans les documents sélectionnés (recherche de l'information).

Cette définition décrit donc une recherche d'information, réalisée par un utilisateur sélectionnant l'information à partir de critères subjectifs de pertinence, dans des fonds documentaires plus ou moins structurés. Cette dernière qualification fait référence aux différents types de traitements attachés aux indexations humaines ou automatiques. Malgré les distinctions opérées par les révisions des deux définitions, les dénominations « recherche documentaire » et « recherche d'information » restent peu différenciables. L'usage consacre d'ailleurs la synonymie des deux expressions.

Les définitions techniques – et la complexité de l'interaction des usagers avec l'information qu'elles peinent à décrire – imposent de mettre à jour les présuppositions théoriques qu'elles contiennent implicitement.

Définitions théoriques et usage scientifique des dénominations

+++++

Dans le domaine scientifique francophone, et particulièrement dans les travaux portant sur l'ingénierie des systèmes d'information, les expressions « recherche d'information », et bien plus rarement « recherche documentaire », sont utilisées, le lexique français relatif à l'action de « rechercher », ou de « chercher » offrant peu le loisir d'en forger de nouvelles. Lorsqu'il s'agit de définir la recherche d'information, renvoi est fait aux travaux anglo-saxons (voir par exemple la présentation faite par [Ihadjadène 04a]). En effet, c'est dans le domaine de la *Library and Information Science*² que des expressions et des sens théoriques en plus grand nombre viennent spécifier la recherche d'information. La langue anglaise offre plusieurs verbes pour définir l'action de chercher par les humains, et trois expressions désignent la recherche d'information : *Information Retrieval*, *Information Searching* et *Information Seeking*. Les nuances que chaque verbe apporte ont permis à la fois de produire des définitions témoignant de l'élargissement de la compréhension de la recherche d'information et d'utiliser ces dénominations pour désigner des courants distincts de la *Library and Information Science*.

2. La discipline *Library and Information Science* (LIS) se retrouve dans les pays de langue anglaise, les pays nordiques mais aussi, de plus en plus, asiatiques. Ses objets de recherche correspondent, à l'exception notable des travaux sur l'activité de recherche d'information, bien plus développés, aux sciences de l'information françaises. En France, les sciences de l'information forment avec la communication une seule discipline institutionnelle, spécificité très largement soutenue par les chercheurs en sciences de l'information et de la communication (SIC). Le singulier « Science de l'information », repris de l'*Information science* de langue anglaise, reste donc en France une dénomination exotique. La dénomination institutionnelle *Library and Information Science* s'est généralisée dans le cours des années soixante-dix. Cette appellation peut être rencontrée sous la forme d'« *Information science* », utilisée par les chercheurs se centrant plus particulièrement sur l'usage des technologies de l'information, et notamment ceux de l'*Information Retrieval*. En revanche, les chercheurs investiguant les problématiques des bibliothèques emploient de préférence *Library science* [Hjørland 00]. À l'intérieur de la LIS, les études sur l'activité de recherche d'information sont aujourd'hui désignées par *Information seeking studies*, ou encore par *Information needs and uses studies*, dénomination plus ancienne [Wilson 07].

Issue du champ professionnel de l'information-documentation, la LIS a construit ses objets de recherche à partir de préoccupations des professionnels, ce qui explique la fragmentation de la discipline en divers objets d'études, bibliothéconomie, bibliométrie, études orientées systèmes (*Information Retrieval*), études orientées usagers (*Information Seeking*). Mais ce sont l'*Information Retrieval* et l'*Information Seeking* qui jouent un rôle essentiel dans la discipline [Ellis 99]. Ces deux approches centrales examinent la recherche d'information.

L'Information Retrieval : retrouver l'information

Le verbe *to retrieve* signifie « trouver à nouveau », avec le sens de localiser et de récupérer. On y entend le « retrouver » français, construit à partir du latin *tropare*³. Littéralement, l'expression *Information Retrieval* pourrait être traduite par « retrouvage de l'information », en empruntant à [Maniez 94] sa réflexion sur une traduction possible du mot *retrieval*. Calvin N. Mooers a choisi ce mot *retrieve*⁴ pour forger en 1948 l'expression *Information Retrieval*. Ce chercheur est à l'origine de la conception des premiers systèmes d'indexation. Selon [Saracevic 99], Mooers l'a proposée pour définir l'objet scientifique d'une science émergente (*Information Science*) étudiant les aspects conceptuels de la description de l'information, et les spécifications qui en résultent pour la conception de systèmes, qu'il s'agisse de techniques ou de machines utilisées pour effectuer l'action de trouver. L'expression *Information Retrieval* désigne ainsi les travaux sur les modélisations des systèmes de recherche d'information. Les études francophones similaires sont le plus souvent désignées par « informatique documentaire » ou « ingénierie documentaire », ou plus largement appartenant au domaine de l'« organisation des connaissances ».

Information Searching et Information Seeking : chercher et rechercher

Les deux expressions sont traduites par « recherche d'information » ou « recherche documentaire ». L'expression *Information Searching* est parfois donnée dans le domaine anglo-saxon comme synonyme d'*Information Retrieval*. Elle définit alors le fonctionnement des systèmes de recherche d'information pour répondre à la demande et afficher l'information [Marchionini 95]. Mais elle connaît une deuxième acception, plus fréquente, qui la rapproche de l'expression *Information Seeking*. En langue anglaise, la distinction entre les deux verbes n'est pas évidente. Les emplois scientifiques de *search* et *seek* restent relativement mouvants, et entraînent des entreprises définitives assez régulières. L'usage commun des verbes *search* (d'origine latine) et *seek* (d'origine germanique) les donne comme synonymes. Cependant, les usages relevés par le dictionnaire rendent compte

3. Selon la définition et l'étymologie données par le Merriam-Webster.
[En ligne] < <http://www.merriam-webster.com/dictionary/retrieve> >.

4. Dans le langage commun, *find* est le plus souvent utilisé pour exprimer l'idée de trouver. *Retrieve* reste rare et employé pour des circonstances spécifiques. On l'utilisera par exemple pour désigner l'action de sauvetage effectué par un Saint-Bernard. Ainsi, Mooers rend-il compte de cette double tâche de localisation et d'extraction de l'information qu'il avait assigné aux systèmes d'information de première génération, tout en satisfaisant aux exigences de l'anglais scientifique pour l'usage de mots d'origine latine.

pour *seek* d'une idée de processus, « se mettre à chercher », tentative pour découvrir quelque chose ou but, chercher à atteindre. Cette nuance semble exploitée par les chercheurs, qui caractérisent de préférence à l'aide de l'expression *Information Seeking* le processus entier de recherche d'information, et utilisent *Information Searching* pour désigner les seules interactions avec les systèmes de recherche d'information (formuler des requêtes, par exemple) [Spink 06 ; Wilson 00]. Nous voudrions conclure ce premier point en retournant vers l'étymologie de deux verbes français « rechercher » et « chercher », qui pourrait être mise à contribution pour la traduction de *search* et de *seek*. Chercher vient de *cercer*, « parcourir en tous sens, fouiller », et rechercher ajoute avec *recercier*, plus tardif, l'idée d'« être à la recherche, ou en quête de quelque chose ». Le premier, « chercher », constituerait donc une traduction correcte de *search*, et le second, « rechercher », de *seek*.

Émergence du paradigme usager et stabilisation des définitions

Les trois expressions en langue anglaise renvoient à deux approches théoriques qui coexistent au sein de la *Library and Information Science* et qui s'intéressent, pour l'une, l'*Information Retrieval*, aux problèmes des systèmes de recherche d'information, et pour l'autre, l'*Information Seeking* ou *Information Searching*, à ceux des humains quand ils recherchent de l'information. Cette deuxième approche, centrée sur les humains, reste peu explorée en France. Palermi et Polity [Palermi 02], dans leur étude comparative des thèses et articles des années soixante-dix à quatre-vingt-dix, notent le « peu de recouvrement » entre la *Library and Information Science* et les sciences de l'information françaises, courant des sciences de l'information et de la communication. Elles soulignent l'importance des thèmes d'informatique documentaire et la faible présence du traitement des aspects « humains, sociaux et éthiques du processus d'information », et relèvent dans les années quatre-vingt, au début de l'informatisation, un « cantonnement » en grande partie technique de la recherche française en sciences de l'information. La situation évolue peu. Un constat similaire est dressé par [Chaudiron 04a], qui regrette « l'intérêt très limité » que suscitent en France les investigations inscrites dans l'approche orientée usager.

Ainsi, après la constitution du champ de l'*Information Retrieval*, un nouvel ensemble d'hypothèses et de méthodes est proposé au sein de la *Library and Information Science*. Qualifié par la suite de changement de paradigme, ce nouveau programme de recherche émerge dans les années soixante-dix

en réaction à la vision appauvrie de l'activité de recherche d'information proposée par l'*Information Retrieval* [Kuhlthau 91]. Sont particulièrement critiquées les visions trop réductrices du besoin d'information et du processus d'évaluation de l'information conduit par l'utilisateur [Chaudiron 04 b]. À partir de cette proposition d'une meilleure intégration de l'activité de l'utilisateur dans la conception des systèmes de recherche d'information, les travaux ont peu à peu structuré un champ de recherche examinant les dimensions cognitive, sociale, voire anthropologique, de la recherche d'information. Ce glissement théorique a reconfiguré la discipline et a abouti à la quasi-autonomisation de deux champs dénommés respectivement « approche orientée système » et « approche orientée usager »⁵, fréquemment abrégés en IS (pour *Information Seeking*) et IR (pour *Information Retrieval*). Pour être tout à fait précis, il faudrait ajouter une autre partition, au sein même de l'*Information Retrieval*. En effet, la division issue de l'émergence du paradigme usager n'a pas seulement conduit à autonomiser le courant de l'*Information Seeking*. Deux approches coexistent également au sein de l'*Information Retrieval*. L'une d'entre elles, l'*Interactive Information Retrieval* (IRR), est centrée sur l'utilisateur [Vakkari 05]. La finalité des travaux contribue à les distinguer : les travaux de l'*Interactive Information Retrieval* ont toujours pour but d'améliorer les systèmes d'information (à partir de l'observation des utilisateurs) ; ceux de l'*Information Seeking* restent tournés vers la compréhension de l'activité de recherche uniquement⁶. Les chercheurs de l'*Interactive Information Retrieval* sont souvent présentés comme des « passeurs », favorisant les communications entre les deux courants, *Information Retrieval* et *Information Seeking*. On peut citer ici quelques noms de chercheurs importants de chacune des deux approches : N. J. Belkin, M. Bates, P. Ingwersen, K. Jarvelin, G. Marchionini, T. Saracevic, A. Spink, P. Vakkari, T. D. Wilson, etc. pour la seconde approche de l'*Information Retrieval*, et C. Kuhlthau, B. Hjørland, R. Savolainen, etc. pour l'approche de l'*Information Seeking*.

Il nous reste à aborder d'autres désignations des études ou de l'activité de recherche d'information pour compléter un tableau bien chargé en dénominations. Les études issues de l'approche usager ont pu être antérieurement regroupées sous l'expression *Information needs and uses*, révélant leur filiation avec les préoccupations professionnelles liées aux

5. Pour le contenu théorique et méthodologique détaillé de la *Library and Information Science*, voir le chapitre II.

6. Les finalités différentes ont des conséquences méthodologiques que nous présentons dans le chapitre II.

exigences des métiers de service dans les centres de ressources documentaires. Cet intitulé générique continue à qualifier les études, bien que la dénomination d'*Information Behavior* (activité informationnelle) vienne récemment la concurrencer [Julien 00]. Elle s'ajoute aux expressions précédemment citées d'*Information Seeking* et *Information Searching*. La présence d'expressions différentes pour qualifier un même champ d'étude, référant à diverses modalités de l'action de s'informer, a engagé les chercheurs à établir des distinctions. [Wilson 00] a proposé de les classer et de les définir. Ce classement est accepté par une grande partie de la communauté de chercheurs du domaine selon [Fischer 05].

L'*Information Behavior* est défini comme l'ensemble des activités humaines en relation avec les sources et canaux d'information, et inclut recherche « active » (recherche d'information dans des fonds documentaires, demandes à des proches ou des médiateurs) et « passive »⁷ d'information (télévision par exemple) et son usage. Cette activité informationnelle est ainsi considérée tout autant dans le cadre des communications interpersonnelles que dans celui des interactions avec les systèmes de recherche d'information (SRI). Ces dernières, les interactions avec les SRI, restent cependant la configuration la plus étudiée par la *Library and Information Science*.

L'*Information Seeking Behavior* est défini comme l'activité de recherche intentionnelle (*purposive seeking*) de l'information, comme conséquence d'un besoin à satisfaire. À cette fin, les individus interagissent avec des systèmes d'information manuels et/ou informatisés (comme les livres ou revues dans leurs différents formats, les banques de données, les catalogues ou encore suivant le sens large de système (ensemble cohérent), les bibliothèques, le Web ou encore les autres humains, une bibliothécaire par exemple)

L'*Information Searching Behavior* est le « niveau micro » de l'interaction avec l'ensemble des systèmes d'information : sont concernées toutes les actions physiques (les clics de souris, les choix d'hyperliens, etc.), mais aussi les opérations cognitives comme les jugements de pertinence.

L'*Information Use Behavior* consiste dans les actions physiques (comme les annotations) et mentales associées (comparaison de la nouvelle connais-

7. Qualificatif discutable qui semble sorti de la dichotomie entre publics (des médias) actifs et passifs (*i.e.*, critiques ou non vis-à-vis des productions médiatiques façonnées par les intérêts de la classe sociale dominante), que les problématiques sur la réception rendent aujourd'hui plus complexe. On peut considérer ici qu'il s'agit de mettre en avant l'interrogation ou non d'un système de recherche d'information.

sance avec celles existantes) utiles à l'appropriation de l'information trouvée. [Spink 06] souligne que l'usage de l'information constitue une des dimensions les moins étudiées.

Ces deux dernières activités (*Information Searching* et *Information Use*) sont à considérer comme deux sous-ensembles de la recherche d'information (*Information Seeking*), elle-même s'intégrant dans le cadre plus large de l'activité informationnelle (*Information Behavior*). D'un strict point de vue définitionnel, les activités informationnelles ainsi décrites pourraient être imaginées comme des cercles concentriques de rayon croissant. Cette présentation des dénominations et des définitions pourrait laisser penser que celles-ci sont désormais fermement établies. Pourtant, de nouvelles dénominations commencent déjà à émerger, comme celle d'*Human-Information Interaction* (interaction humain-information) bâtie sur le modèle de l'interaction humain-machine, ou encore *Exploratory search* (recherche exploratoire) [Marchionini 06], plus à même de rendre compte des phénomènes informationnels observés aujourd'hui.

POURQUOI RECHERCHER DE L'INFORMATION ? LE BESOIN D'INFORMATION

La notion de besoin d'information est centrale dans le domaine de la recherche d'information dans les documents, puisque cette dernière est définie comme une interaction entre « un individu qui a besoin d'information » et « un document qui contient ou non la réponse à ce besoin » [Mizzaro 98]. La notion de pertinence est définie par ce même auteur comme relative à la réponse au besoin d'information : est pertinent ce qui comble, au moins partiellement, mon besoin d'information. Ce dernier a été défini, au cours de l'histoire des sciences de l'information, comme :

- rechercher une réponse à sa question [Taylor 67 et 68]⁸ ;
- réduire l'incertitude [Atkin 73] ou un état de connaissance insatisfaisant [Belkin 76] ;
- donner du sens [Dervin 83].

8. [Taylor 68] est la publication dans la revue *College & Research Libraries* (référence la plus fréquente, mais plus difficile à localiser aujourd'hui) de travaux publiés un an plus tôt dans un rapport intitulé *Question negotiation and information seeking in libraries*, aisément accessible à < <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=AD659468&Location=U2&doc=GetTRDoc.pdf> >. En outre, les annexes de ce rapport fournissent une vue remarquable et rare d'activités de recherche dans une bibliothèque physique telles qu'elles étaient conduites par des étudiants dans la décennie soixante.

Rechercher une réponse à sa question

+++++

Taylor s'efforce de restituer les différents états du besoin d'information. Historien et bibliothécaire, il s'interroge sur les difficultés éprouvées par les usagers pour exprimer leur besoin d'information. Il produit un modèle en quatre étapes, considéré comme une des premières conceptualisations adoptant un point de vue cognitif du besoin d'information. Le besoin d'information est décrit comme un processus cognitif comprenant quatre niveaux :

- le besoin réel mais inexprimable (niveau 1) que Taylor nomme *visceral need*, besoin « intuitif », vaguement conscient ;
- le besoin conscient, *conscious need*, tel qu'il est perçu par la personne, mais toujours inexprimable ;
- le besoin pouvant être exprimé en langage naturel, *formalized need*, le besoin formalisé ; le besoin a donc pris la forme d'une question ;
- le besoin accommodé, adapté, *compromised need* : la question est présentée au système d'information dans un langage de compromis (mots-clés). À ce stade, la question est formulée en anticipant ce que le système pourra restituer.

Ce modèle, fondement des travaux sur le processus de recherche d'information, a été très tôt discuté d'un point de vue méthodologique, parce qu'il prenait appui sur des observations d'un nombre jugé trop réduit. Cependant, dans le rapport de 1967, nous avons recensé 8 professionnels interviewés durant une heure à une heure trente, et 4 observations d'étudiants, ce qui peut être considéré comme un nombre acceptable dans une méthodologie de nature qualitative. De plus, les travaux de Taylor sont aujourd'hui théoriquement critiqués pour l'absence de prise en compte des éléments contextuels, comme la tâche à effectuer, à l'origine du besoin d'information, ainsi que rappelé par [Ingwersen 05]. Ces dimensions n'ont cependant pas été entièrement ignorées par Taylor qui les reprend dans la liste de « filtres » qu'il propose aux professionnels d'employer lors de leurs communications avec les usagers.

Le modèle de Taylor a le mérite de mettre en avant la difficile prise de conscience du besoin d'information, mais les modalités des passages d'un niveau à l'autre, ce qui rend possibles les changements de ces niveaux, ne

sont pas envisagées. La proposition de Taylor, souvent associée à celle de Dervin, constitue le point de départ de nombreuses théorisations ultérieures.

Manque de connaissances ou besoin de réduction d'incertitude

Belkin travaille plus particulièrement le premier niveau du modèle de Taylor. Sa définition du besoin d'information, élaborée dès 1976 [Belkin 05], s'appuie sur les niveaux de conscience de Taylor puisqu'il se propose de partir d'une idée approachante, celle d'un état incomplet de connaissances de l'utilisateur. Le modèle est dénommé *Anomalous State Knowledge* (ASK). Les « anomalies » dans les connaissances, le manque de connaissances sur un thème, les doutes sur certaines connaissances, sont à l'origine du besoin d'information. Cette façon de présenter le problème posé par le besoin d'information pourrait sembler banale, reprenant la signification générale d'un besoin comme situation de manque. L'intérêt est ailleurs. D'une part, il introduit l'idée de doute, d'incertitude, et d'autre part, il précise avec ses collègues Oddy et Brooks en 1982 [Belkin 82] que si l'utilisateur est capable de reconnaître ces anomalies, en revanche, il ne peut caractériser avec précision ce qui serait nécessaire pour les résoudre. Il est donc peu approprié de demander à un utilisateur d'exprimer ce qu'il ne sait pas. À partir des travaux sur la mémoire associative de la psychologie cognitive [Belkin 05], la modélisation du besoin d'information ASK va servir à la conception d'un système de recherche d'information capable de construire une représentation de l'état de connaissance de l'utilisateur, et la mettre en relation avec les représentations des informations contenues dans le système. Belkin et ses collègues confient donc au système le soin de faire évoluer l'état de connaissance de l'utilisateur.

Si, comme [Atkin 73] et plus tard [Belkin 76], on conçoit le besoin d'information comme un besoin de réduire l'incertitude, alors l'incertitude doit être définie non pas comme un manque de connaissances mais comme la prise de conscience d'un manque de connaissances. Le biais de confirmation, étudié en psychologie du raisonnement, illustre ce phénomène. Après avoir pris une décision ou émis un jugement, les individus ne recherchent pas d'information qui pourrait contredire cette décision. Ils ont même tendance à ne pas traiter ces informations quand elles leur sont présentées. Jonas et ses collaborateurs [Jonas 01] montrent que ce biais est particulièrement fort quand l'information est présentée de façon séquentielle (alors qu'il est classiquement attesté dans des présentations simultanées). Pour ces auteurs, ce biais provient d'une focalisation de plus en plus importante

des individus sur leur décision. Plus on tient à une décision, à une opinion, à une position, moins on a besoin d'informations (potentiellement contradictoires) à son propos. Réciproquement, plus on est soumis à des informations contradictoires que l'on refuse et plus on tient à sa décision.

Il est nécessaire d'avoir des connaissances pour prendre conscience qu'on manque de connaissances, pour avoir de l'incertitude. Pour autant, avec ou sans connaissances, l'individu peut avoir l'« illusion de savoir ». Des recherches montrent en effet que les individus experts d'un domaine peuvent commettre des erreurs dues à leur connaissance d'une situation très fréquente. Par exemple, Besnard et Bastien-Toniazzo [Besnard 99] ont proposé une tâche de diagnostic de panne d'un dispositif d'alarme électronique à des électroniciens experts (professionnels, 20 ans d'expérience) et novices (étudiants, 2^e année d'IUT). La panne avait été introduite artificiellement : il s'agissait d'un court-circuit sur un condensateur placé sous le dispositif, donc invisible. Or, une panne est extrêmement fréquente avec ce type d'alarme électronique, mettant en cause un circuit intégré. Les résultats montrent que les experts testent en moyenne 18 fois le circuit intégré qui n'est pas défaillant ; les novices 2 fois seulement. Les experts éprouvent beaucoup plus de difficultés que les novices pour trouver la panne, parce qu'ils n'arrivent pas à prendre conscience que leur connaissance de la panne fréquente n'est pas la bonne. Ainsi, le besoin d'information correspond à un manque de connaissance d'un individu particulier dans une situation particulière. Cela n'a rien de spécifiquement documentaire. Prendre conscience de ce besoin implique que l'individu ait des connaissances, mais ces connaissances ne garantissent pas cette prise de conscience. La recherche d'information dans les documents est une solution parmi d'autres.

Pour résumer, selon [Tricot 04a] : si j'ai de l'incertitude alors j'ai besoin d'information ; si je n'ai pas de connaissances alors je n'ai pas d'incertitude ; si j'ai de la certitude alors je n'ai pas besoin d'information. Avoir besoin d'information implique que l'on ait de l'incertitude, et donc des connaissances.

Donner du sens

 [Dervin 83] propose une réflexion sur le besoin d'information radicalement différente des conceptualisations psychologiques que nous venons de décrire. Son intention est de clarifier comment les individus, dans le processus de construction de leur réalité, construisent à un moment donné leurs besoins

et usages de l'information, les interrogations des systèmes de recherche d'information pouvant ou non participer à cette construction personnelle. Dervin résume sa proposition : l'information n'est pas une donnée qu'il suffit d'extraire, mais une création personnelle de sens, ancrée dans un espace-temps donné. Dervin utilise la métaphore du marcheur qui fait face à une faille, une brèche (*gap*). Les êtres humains sont sans cesse confrontés à des discontinuités. Ces failles, présentées comme changements de l'entourage, de l'environnement ou du savoir, prennent la forme de question, d'énigme, de confusion, d'anxiété métaphysique. Elles doivent être comblées (*bridge*) par des idées, des attitudes, des croyances, des valeurs, des sentiments, des intuitions, des récits. La recherche d'information dans les systèmes d'information n'est qu'une des façons de combler les failles [Dervin 05]. « Les choses changent constamment », ce qui oblige les êtres humains à créer du sens continuellement. Dervin nomme sa « métathéorie », *sense-making*, faire sens ou fabrique du sens.

Pour attribuer du sens, les individus utilisent aussi bien les observations des autres que les leurs. On reconnaît, dans cette réalité chaotique que l'individu ordonne, l'inspiration phénoménologique. Replacée dans la problématique du besoin d'information, la conceptualisation est donc la suivante :

- le besoin est une discontinuité dans ce mouvement ici et maintenant. Les personnes se perçoivent alors comme étant dans un moment problématique, comme devant enjamber un fossé d'une façon ou d'une autre ;
- le besoin d'information est défini comme une combinaison de trois éléments : les questions et confusions, la situation (expérience, histoire tout autant présente que passée) et utilisation d'aides ;
- un système d'information peut servir ce besoin, le pont pouvant être par exemple l'information stockée. Les phénomènes internes (idées, émotions, sentiments, intuition) contribuent à construire le sens autant que les phénomènes externes. Ainsi, un pont est-il aussi bien un fait retrouvé dans une base de données qu'un changement dans le comportement de quelqu'un d'autre.

Dervin assortit la proposition théorique et une méthodologie basée sur des entretiens chronologiques (*Time-line interview*). Les questions portent sur

le contexte de vie des personnes étudiées, leurs ambitions, leurs satisfactions, les expériences avec la technologie [étudiée par le chercheur], et prennent en compte leur déroulement dans le temps. La grille d'analyse s'appuie sur le repérage de la prise de conscience d'un changement nécessaire, les moments pour rééquilibrer la situation [Frenette 05]. La métathéorie de Dervin est considérée comme la base générale de l'ensemble des théorisations du paradigme usager et continue de nos jours d'être très fortement citée dans les études empiriques. La métathéorie *sense-making* de Dervin peut se lire comme une modélisation de l'activité informationnelle, partant d'un schéma de base qui n'est pas une nouveauté pour arriver à la formulation d'une proposition inédite. En effet, la théorie sous-jacente - une situation problématique - ne constitue pas une proposition théorique originale. Nous avons noté l'inspiration phénoménologique, influence philosophique s'exerçant de façon renouvelée dans les sciences humaines et sociales. [Dervin 05] situe son travail à l'intersection de nouveaux travaux américains et européens dans les champs de la philosophie, de la sociologie, de la psychologie, de l'éducation, de la communication, les *cultural, feminist et postmodern studies*. Ceux-ci critiquent les explications attribuant aux contraintes externes, normes sociales, place dans la hiérarchie sociale, un rôle déterminant dans la conduite d'une action individuelle, et considèrent l'action comme une élaboration par un acteur dans le temps et l'espace sans contraintes antécédentes. La dénomination même de la théorie, *sense-making*, banale dans le sens de commune à plusieurs champs, témoigne simplement de l'ancrage théorique choisi. Cependant, appliquée à une activité non prise en compte ailleurs, la recherche d'information, la suggestion théorique de Dervin au début des années quatre-vingt, forme une réponse nouvelle en formulant l'idée que la situation problématique peut être résolue par une recherche d'information. Elle offre de la sorte la possibilité de révéler la nature de la situation problématique, l'importance prise par l'information pour résoudre le problème (franchir le fossé), et enfin identifier à partir de l'usage de l'information les conséquences de l'activité [Wilson 99]. La proposition de Dervin, bien que située dans un cadre général, contribue à la réflexion sur la formation et le développement du besoin d'information dans le cours de l'activité, et plus encore dans les épisodes multiples et successifs de recherche d'information. Cependant, peut-être du fait de sa généralité, ce sont surtout les conséquences méthodologiques de la proposition qui sont aujourd'hui considérées comme présentant le plus d'intérêt [Wilson 99]. L'auteure elle-même [Dervin 05] désigne désormais sa proposition de *sense-making methodo-*

logy, désignant une méthode capable de repérer les déterminations de l'action, le temps étant une donnée importante de la méthode, selon les phases du processus de recherche d'information.

La négation du besoin d'information

Si les différentes conceptions du besoin d'information se situent à des niveaux d'analyse différents, toutes indiquent que ce qui est dénommé « besoin d'information » est une condition pour s'engager dans une activité de recherche d'information, et que ce préalable peut ne pas avoir lieu. Plus encore, un besoin d'information peut être conscient, identifié par le chercheur d'information, et ne pas entraîner *ipso facto* l'engagement dans une démarche de recherche d'information. De fait, des « négations » du besoin d'information ont été observées. [Marchionini 95] note que le besoin peut ne pas être accepté, voire être supprimé en fonction des connaissances antérieures de l'utilisateur, de ses habiletés à utiliser un système de recherche d'information, des circonstances externes et de son estimation de la situation (coûts et bénéfices). L'auteur propose donc une explication de nature cognitive prenant appui sur la difficulté à résoudre le problème informationnel. D'autres explications ont été fournies qui éclairent le phénomène de la négation du besoin informationnel. [Chatman 91 et 96] a particulièrement travaillé cette question. S'inscrivant dans le courant des usages et gratifications, elle propose la notion de « pauvreté informationnelle », à partir d'études empiriques sur les besoins d'information des populations marginalisées, le plus souvent économiquement pauvres, étudiant les besoins d'informations de concierges, mais aussi du fait de leur isolement, les femmes âgées dans des maisons de retraite. Chatman fait le constat que la pauvreté informationnelle n'est pas systématiquement corrélée à la pauvreté économique. Elle définit un monde informationnel appauvri comme un monde dans lequel une personne est incapable de résoudre une inquiétude, une préoccupation majeure, à cause de normes sociales du groupe d'appartenance qui définissent ce qui peut être cherché et les informations qui peuvent être partagées ou non. Des besoins informationnels peuvent être cachés pour témoigner de comportements socialement acceptables, et des informations peuvent être rejetées.

Le concept de besoin d'information, en grande partie caractérisé dans l'approche cognitive, reçoit ainsi un éclairage social.

Critiques de la notion de besoin d'information

+++++

Centrale dans la description de la recherche d'information, la notion de besoin d'information est discutée. L'examen de ses fondations théoriques a conduit à lui adjoindre une notion supplémentaire, à l'agréger avec celle de tâche complexe, mais il s'est également achevé dans des critiques plus radicales.

Besoin ou tâche ?

En décrivant la recherche d'information principalement autour de la notion de besoin d'information, en faisant référence à une résolution de problème mal défini ou situation problématique, en termes de manque de connaissances et de moyens à mettre en œuvre, de prise de conscience du besoin d'information, de construction de sens, on laisse dans l'ombre trop d'événements se produisant dans l'ensemble du processus. D'autres travaux existent sur la recherche d'information, se centrant sur d'autres dimensions du processus de recherche, et il faudrait connecter l'ensemble de ces travaux pour comprendre comment se développe le besoin d'information dans le cours de l'activité. C'est la critique adressée par [Vakkari 99]. Celui-ci relève dans les conceptualisations du besoin d'information et le manque de liens avec les autres phases du processus informationnel. Il note l'importance d'examiner les évolutions du besoin d'information en reliant les variations de l'état des connaissances telles que présentées par Belkin et celles des stratégies décrites par Bates (cf. chapitre II). Pour ce faire, il propose de partir de la notion de « tâche complexe », c'est-à-dire la façon dont le chercheur d'information perçoit la complexité de la tâche. Partir de la notion de tâche renvoie à l'idée de définition par le chercheur d'information de l'ensemble des éléments de la tâche, c'est-à-dire à la possibilité pour celui-ci de déterminer la structure du problème (le besoin d'information), la manière d'accomplir la tâche, et les résultats qu'il en attend. La notion de tâche a en outre l'avantage d'ouvrir aux problématiques du contexte, une tâche d'information (*search task*) ayant souvent lieu dans un milieu professionnel spécifique constituant la tâche de travail (*work task*). La notion de tâche dans la recherche d'information a été par la suite utilement appliquée à un contexte non professionnel, celui de la vie de tous les jours (*daily life*) particulièrement examiné par les chercheurs scandinaves.

On trouve chez [Ingwersen 05a] une synthèse des travaux sur la complexité de la tâche et ses liens avec le besoin d'information. La perception de la complexité de la tâche de travail ou de la vie quotidienne à laquelle

s'ajoute celle de la tâche de recherche d'information est classée dans trois catégories :

- (1) les tâches automatisées, les routines (*automatic/routine task*) de traitement de l'information (simples, rapides, souvent factuelles, le chercheur d'information disposant des connaissances procédurales et déclaratives) ;
- (2) les tâches régulières de traitement de l'information ou de prise de décision, habituelles (*normal task*), pour lesquelles la façon de résoudre le problème posé par la tâche tout comme le domaine sont familiers, mais pas le contenu précis de la tâche elle-même, qui demande des recherches d'information. Dans ces deux tâches, les moins complexes, le besoin d'information peut être énoncé, formulé. Le chercheur d'information peut dire ce qu'il sait déjà sur le thème de recherche, et pourquoi il cherche de l'information ;
- (3) les tâches authentiques, sérieuses (*genuine tasks*), des prises de décisions qui peuvent être partiellement connues ou totalement inconnues, ce qui renforce la complexité de la tâche. Elles sont exigeantes en temps pour l'acquisition d'informations du domaine, mais également d'informations pour résoudre la tâche.

Dans cette tâche complexe, le besoin d'information peut être très vague, et le chercheur d'information n'a pas idée de l'information qui pourrait résoudre le problème, alors même que ce type de tâche perçue comme complexe est la plus exigeante en information. La notion de tâche complexe suppose donc que l'on s'interroge sur l'utilité des connaissances dont dispose le chercheur d'information (ce qu'il sait des objets d'information utiles pour résoudre la tâche) mais aussi sur l'intention sous-jacente à la tâche de recherche (les buts du chercheur, chercher des contenus, des données structurées).

Le degré de complexité perçue de la tâche de travail ou de la vie quotidienne refléterait l'état du besoin d'information. Ainsi, le concept de besoin d'information continue à être régulièrement caractérisé, et le questionnaire se déplace, passant de la notion de situation problème, de prise de conscience, de connaissances, à celle de tâche complexe. Une définition ne

chasse pas l'autre. Ces définitions s'agrègent souvent dans les arrière-plans théoriques des études empiriques.

À partir d'une critique identique, le courant dit de « l'approche rationnelle » en psychologie cognitive aboutit à des propositions qui, plutôt que de s'interroger sur ce qu'est le besoin d'information, tentent de décrire ce qui entraîne une activité de recherche d'information (par ex. [Pirolli 99]). Tricot et Raufaste [Tricot 07a] ont proposé une synthèse de ces approches (figure 1) : ce qui entraîne une activité de recherche d'information, c'est le fait d'espérer qu'une information potentielle a une certaine utilité (une certaine valeur), supérieure au coût estimé de l'activité de recherche. L'estimation de l'utilité espérée de l'information potentielle est réalisée en confrontant la valeur espérée de l'information, la valeur espérée du *statu quo* (c'est-à-dire de la non-recherche) et l'existence de marques de pertinence de cette information (c'est-à-dire de la mise en exergue d'indices de sa valeur). À son tour, l'évaluation du *statu quo* prend en compte la sensibilité de l'individu à l'incohérence, son anxiété, sa prudence et son engagement dans la situation antérieure (plus il est engagé dans la situation antérieure plus il risque de surestimer la valeur du *statu quo*). Les estimations sont d'autant plus justes que l'individu est expert dans le domaine de connaissance concerné. L'anxiété et la prudence de l'individu sont liées à son expérience.

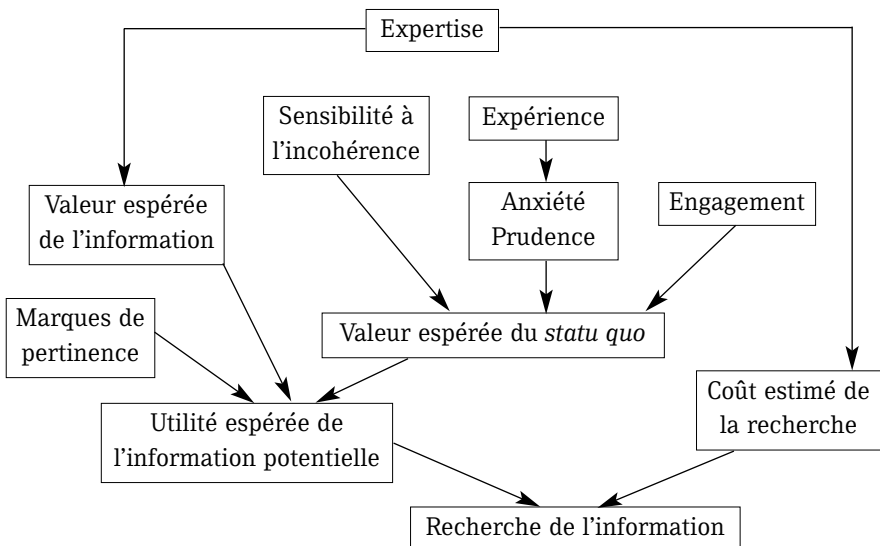


Figure 1 : le modèle de Tricot et Raufaste [Tricot 07 a]

Besoin ou désir d'information ?

Pour certains chercheurs, la notion de besoin semble avoir le double inconvénient de référer à un manque et aux sciences cognitives. Ils ont suggéré qu'il fallait se défaire de la notion de besoin d'information. [Case 02] remarque que toutes les recherches d'information ne sont pas liées à un problème à résoudre ou à une décision à prendre. Critiquant implicitement l'orientation cognitive des études, il note que l'on peut chercher de l'information pour se divertir, avoir plus d'assurance, ou bien moins d'ennui. Une étude francophone récente sur les pratiques « ordinaires » d'Internet par les jeunes étudiants pourrait venir conforter la critique de Case. Elle souligne l'aspect ludique des pratiques informationnelles juvéniles dans le Web. Ainsi les jeunes, des étudiants, disent-ils apprécier « échapper à un cadre strict » ; ils n'ont pas toujours de recherches précises à effectuer et surfent de liens en liens, alliant l'utile à l'agréable [Jouet 05]. Toutefois, quelle qu'en soit la nature, il semble difficile de considérer une action humaine sans but, sans motif, ce à quoi renvoie avantageusement la notion de besoin d'information. Autrement dit, il vaut mieux considérer qu'il existe des besoins d'information qui seraient sérieux, et d'autres futiles – si l'on traduit ainsi la partition de Case – faute de quoi l'on se retrouverait devant certaines recherches d'information face non plus à une activité (avec toute la complexité théorique à laquelle la notion renvoie, cf. chapitre II), mais à un comportement d'un faible intérêt théorique. Une deuxième critique est adressée par [Boullier 97] dans sa réflexion sur l'usage des objets techniques. Rejoignant Case, il conteste les notions de « besoin » et de satisfaction du besoin à cause de la référence au behaviorisme. Il lui préfère celle de désir et son explication psychanalytique, susceptibles de mieux rendre compte de l'activité du sujet. Le terme « besoin » gêne donc les chercheurs, s'inscrivant dans une perspective sociologique et communicationnelle. Si le terme « besoin » avait été abandonné au profit de celui d'« intention », pour reprendre le vocabulaire philosophique de l'action [Livet 05] – intention de s'informer – en mettant l'accent sur la finalité de l'activité, ces critiques se seraient-elles maintenues ? Assurément, pour l'activité de recherche d'information, il nous semble que le plus grand inconvénient du concept de besoin d'information réside précisément dans sa dénomination, renvoyant trop directement au behaviorisme. Cependant, les premières théorisations autour du besoin d'information que nous avons évoquées montrent à quel point le concept de besoin, lorsqu'il est associé à celui d'information, est éloigné des conceptions behavioristes, qu'il relève

de l'approche cognitive, critique elle-même du behaviorisme ou d'une perspective sociologique elle aussi fructueuse.

Le concept de besoin d'information ne rend pas seulement compte de ce qui déclenche l'activité de recherche, mais aussi de la co-transformation du besoin et de l'activité. Sur ce point, Case apporte une seconde critique plus convaincante, notant que lorsqu'il s'agit de définir la recherche d'information, les chercheurs se réfèrent surtout au besoin d'information. Il note l'emphase théorique du concept de besoin d'information, qui finit par servir à caractériser l'ensemble du processus informationnel.

Pour conclure... provisoirement

+++++

Le concept de besoin d'information est jugé utile pour comprendre ce qui motive et guide une recherche d'information. Dans l'inflexion phénoménologique donnée par Dervin, ou instruit au niveau cognitif de l'interaction avec les systèmes d'information par Taylor, Atkin, Belkin, le besoin est non seulement ce qui enclenche une recherche d'information mais aussi ce qui prend forme dans le processus même de recherche d'information, et qui en retour le façonne. Une définition plus récente, proposée par des chercheurs scandinaves, s'efforce de mettre en relation la notion de tâche complexe telle que perçue par l'utilisateur et celle de besoin d'information. Reconnu, le besoin peut cependant être rejeté pour des raisons cognitives ou sociales par le chercheur d'information. Les critiques dont il est l'objet laissent penser qu'il s'agit peut-être d'un concept mal nommé. Il est certainement trop large, contenant deux problèmes informationnels distincts, le commencement de l'activité et son accomplissement.

Enfin, on peut aussi proposer de catégoriser les situations dans lesquelles apparaissent les besoins d'information. Cette catégorisation est issue de nos travaux empiriques [Tricot 03] et de l'analyse de la littérature du domaine. Les trois premiers besoins d'information ont une source interne à l'individu, les trois derniers une source externe. Un besoin d'information peut apparaître dans une situation où l'on a :

- besoin d'une connaissance que l'on n'a pas ; besoin d'une confirmation d'une connaissance que l'on a ;
- besoin d'une connaissance plus complète que celle qu'on a, un exemple, une illustration, un contre-exemple, etc. ;
- besoin d'être conforme aux buts, aux contraintes, aux attentes de la situation ;

- besoin d'indications sur la forme de la connaissance à utiliser dans la situation (par exemple : qu'est-ce que « faire un exposé » ?) ;
- détecté un marqueur de pertinence dans la situation (ostension, mise en exergue visuelle, sonore, etc. Par exemple : un mot en gras dans le texte ; un enseignant qui me dit : « tu es bien sûr ? »).

Ainsi, les êtres humains ont des besoins d'information différents, ils recherchent pour des raisons différentes. Une deuxième dimension fondamentale est décrite maintenant : la tâche de recherche d'information.

QUE RECHERCHER ? LES TÂCHES DE RECHERCHE D'INFORMATION

Après « Pourquoi rechercher ? », la seconde grande question du domaine d'étude de la recherche d'information est donc « Que rechercher ? ». Pour répondre à cette question générale, deux voies complémentaires ont été suivies, l'une et l'autre avec de nombreuses embûches : analyser les tâches de recherche d'information, c'est-à-dire ce qu'un individu a « à faire » pour satisfaire un besoin d'information, et catégoriser ces tâches. Nous présentons brièvement ces deux voies.

Analyse et catégorisation des tâches

L'analyse d'une tâche de recherche d'information doit probablement d'abord prendre en compte le besoin d'information et la façon dont l'individu transforme ce besoin en but informationnel [Tricot 98]. Cette représentation mentale du but a une composante conceptuelle (quel contenu sémantique est recherché ?) et procédurale (quelle suite d'actions va permettre d'atteindre le but ?), voire informationnelle (quel besoin d'information doit être satisfait ?). Cette représentation du but est donc influencée par les connaissances de l'individu. En second lieu, la description de la tâche de recherche d'information doit probablement prendre en compte des caractéristiques de la cible, c'est-à-dire la localisation et le nombre de cibles dans le(s) système(s) d'information, les procédures à utiliser pour atteindre ces cibles, la structure générale du (des) système(s) et des interfaces de recherche utilisées.

En croisant ces deux premières dimensions, la représentation du but et les caractéristiques de la cible, on décrit donc la tâche. On peut aussi établir une catégorisation. [Tricot 93] a par exemple proposé une catégorisation sommaire en considérant que la représentation mentale du but pouvait être précise ou floue, et que la localisation de la cible dans le système pouvait être unique et localisée, ou multiple et diffuse. Ce croisement permet de catégoriser quatre tâches (cf. tableau ci-dessous).

Tableau : quatre tâches de recherche d'information [Tricot 93]

		Représentation mentale du but	
		précise	floue
Localisation de la cible	unique et localisée	extraire	explorer
	multiple et diffuse	collecter	butiner

[Bernstein 93] a proposé une catégorisation des usages des systèmes d'information qui a fait date. Il distingue trois types d'usages : l'extraction, l'exploitation et le jardinage. Pour une recherche d'extraction d'information, l'information pertinente est une ressource de valeur qui doit être extraite efficacement et raffinée. L'exploitation serait une recherche permettant la conception ou l'élaboration d'un document. Ce type de recherche d'information conçoit l'acquisition, le raffinement, l'assemblage et la maintenance d'une information comme une entreprise continue. Le jardinage conçoit la culture de l'information comme une activité continue et collaborative, conduite par des groupes de personnes travaillant ensemble à la réalisation de buts changeants, individuels et communs. Bernstein fait remarquer que les critères d'évaluation de ces trois activités sont radicalement différents, et qu'à partir du moment où l'on veut faire une activité dans un système qui n'est pas prévu à cet effet, la démarche est alors vouée à l'échec. Une tâche avec représentation précise de la tâche principale mais distribution des cibles est appelée « tâche de cristallisation des connaissances » par Pirolli et Card [Pirolli 99], sans qu'un essai de standardisation de la description des tâches ou de leur effet ne soit évoqué par les auteurs. On trouve des descriptions de tâches, mais peu abouties, chez [Marchionini 95] ou [Wright 90]. Belkin fait une différence intéressante entre le but de la tâche (où il oppose les buts d'apprentissage aux buts de

sélection) et la méthode (où il oppose le *scanning* à la recherche). Mais, là encore, l'auteur ne fournit pas de recension des effets des tâches sur l'activité. Pourtant [Belkin 85] est sans doute le premier auteur à avoir mis en exergue le rôle primordial de la description des tâches de recherche d'information dans la conception des systèmes d'information.

Tricot et Golanski [Tricot 02] ont essayé de recenser les effets de caractéristiques objectives des cibles sur l'activité de navigation. Ils ont mis en évidence empiriquement quelques effets, et ont aussi rapporté quelques effets robustes. Leur approche empirique a consisté à mesurer les performances d'individus réalisant des tâches qui variaient, successivement, selon l'une des variables présentées ci-dessous. Voici les caractéristiques de la tâche qui semblent les plus pertinentes : la répétitivité de la tâche (le nombre de fois où une même cible est recherchée par un même individu), le degré d'explicitation des cibles, le fait que la cible soit localisée ou distribuée, la complexité de la procédure à mettre en œuvre (c'est-à-dire le nombre de décisions différentes à prendre entre le début et la fin de l'activité). Ce point est repris plus bas dans l'ouvrage (p. 95).

Ce domaine de recherche doit être développé car il comporte encore beaucoup d'inconnues. Par exemple, le fait que la cible soit définie *a priori*, pendant la recherche d'information, ou découverte *a posteriori* a-t-il une influence sur la difficulté de la tâche ? Le type de modalité sensorielle (audition/vision) impliquée dans la recherche d'information, par exemple dans le cadre de comparaison entre serveurs vocaux et base de données textuelles classiques, a-t-il une influence sur l'efficacité de la recherche d'information [Le Bigot 04] ? Le poids des données en termes de débit, de stockage ou d'affichage a-t-il un effet sur les performances des utilisateurs. Par exemple, la recherche d'information sur les *Personal Digital Assistant* (PDA) est-elle aussi efficace que sur les ordinateurs communs avec écrans 17 pouces ? Ces recherches doivent aussi être développées car elles ont des implications ergonomiques importantes. Les résultats recensés sont autant de connaissances ergonomiques utilisables par les concepteurs, qui peuvent définir des contraintes de conception en fonction de contraintes de la situation d'utilisation prévue.

Enfin, la description de la tâche doit sans doute aussi décrire le contexte de l'activité et la raison pour laquelle un individu choisit d'utiliser un tel système pour chercher telle information dans le cadre de la réalisation d'une tâche principale.

Tâches en contexte

+++++

De plus en plus de travaux, notamment ceux des chercheurs scandinaves évoqués dans la partie précédente à propos de tâche complexe, décrivent les tâches de recherche d'information en contexte : ils tentent de décrire la tâche en tenant compte des éléments qui ont pu contribuer à faire naître le besoin d'information et de ceux qui vont permettre l'exploitation de l'information trouvée [voir par exemple Chaker 10 ; Cool 02 ; Ingwersen 05a ; Johnson 03 et plus avant dans cet ouvrage, chapitre III]. Li et Belkin proposent une approche à facettes de la conceptualisation des tâches de recherche d'information [Li 08]. Ces travaux ont pour objectif de réaliser une classification des tâches qui pourrait être applicable à tous les niveaux (tâches liées au contexte, *Information Seeking* et *Information Searching*) et ce à partir de points communs (facettes). Ils se sont fondés sur une étude assez exhaustive de la littérature et montrent que des modèles existants ont eu tendance à ne s'intéresser qu'à des aspects partiels des tâches, et qu'une approche à facettes peut fournir un modèle plus holistique reposant notamment sur la source de la tâche, l'utilisateur, le temps, le produit, le processus, le but, les caractéristiques de la tâche et la perception de la tâche qu'a l'utilisateur.

Nous avons commencé cet ouvrage, dans cette partie introductive, en décrivant l'arrière-plan théorique des dénominations de la recherche d'information et en restituant les questionnements sur le concept premier de besoin d'information et celui fortement associé de tâche de recherche. Ces deux concepts seront par la suite mis en relation avec d'autres concepts, stratégies et pertinence, et précisés à l'occasion de la présentation des développements théoriques et méthodologiques des disciplines examinant l'activité informationnelle. Pour l'heure, nous proposons un déplacement du regard vers les travaux empiriques portant sur la recherche d'information de publics spécifiques, autres champs d'études du domaine générant des problématiques de recherche spécifiques.