

**Stéphane Chaudiron**

*Université de Paris X-Nanterre*

*Centre de recherche en information spécialisée et en médiation des savoirs, Cris-Series*

**Madjid Ihadjadene**

*Université de Paris X-Nanterre*

*Centre de recherche en information spécialisée et en médiation des savoirs, Cris-Series*

# **ÉVALUER LES SYSTÈMES DE RECHERCHE D'INFORMATION**

## **Nouveaux modèles de l'utilisateur**

### **Introduction**

L'accès aux systèmes de recherche d'information sur Internet (bibliothèques numériques, moteurs de recherche...) reste problématique pour les non-spécialistes du processus de recherche. (Ihadjadene, 1999) et (Spink, 2001) ont montré que les ressources sont sous-utilisées dans les catalogues en ligne, les bibliothèques numériques ou les moteurs de recherche. L'une des premières difficultés pour les usagers est d'identifier les sources pertinentes et d'avoir une vision claire des contenus et de leur structure. Par ailleurs, bien que les moteurs de recherche proposent des fonctionnalités avancées comme la troncature ou les opérateurs booléens, les usagers les emploient rarement. Plusieurs études ont également montré que le choix des termes ne se fait pas selon un plan de recherche méthodique mais résulte au contraire d'une démarche d'essai erreur, aggravée par le fait que les usagers n'ont souvent qu'une vague idée de ce

qu'ils cherchent. Qu'ils préfèrent changer le contenu de la requête plutôt que de modifier sa formulation interroge la pertinence des outils logico analytiques mis à leur disposition et nécessite de définir un cadre d'évaluation des SRI qui prennent en compte les véritables comportements informationnels.

Dépassant l'approche techniciste fondée sur la notion de performance intrinsèque du système, l'approche « usager » de l'évaluation oriente l'attention vers la réalité et la complexité des comportements, dans ses dimensions sociales, organisationnelles et cognitives. L'approche évaluative orientée usager porte sur la pratique, au sens de processus d'identification, de recherche et d'accès à l'information médié par un dispositif technique.

En dépassant l'approche techniciste et ses limites (Chaudiron, 2001 ; Chaudiron, Ihadjadene, 2002), cet article présente l'apport des sciences cognitives à l'évaluation des SRI. Dans la première partie, nous présentons l'intérêt de l'approche cognitive par rapport à l'approche système tout en soulignant ses limites théoriques et méthodologiques ; dans la deuxième partie, l'extension de l'approche cognitive, qualifiée d'holistique.

## L'approche cognitive

L'approche cognitive en RI trouve ses origines à la fois dans le développement des sciences cognitives de la fin des années 1970 (De Mey, 1977 et 1980), et dans le rejet de l'approche strictement algorithmique en recherche d'information. Ce champ de recherche a connu un essor particulier à la suite des travaux de Marc De Mey présentés lors du séminaire *International Workshop on the Cognitive Viewpoint* (1977), souvent considéré comme le point de départ de l'approche cognitive en sciences de l'information<sup>1</sup>. Elle se fonde en particulier sur la thèse selon laquelle : « any processing of information, whether perceptual or symbolic, is mediated by a system of categories or concepts which, for the information processing device, are a model of his (its) world »<sup>2</sup>.

En ce sens, la tâche d'un SRI peut être décrite comme l'appariement des structures cognitives des producteurs d'informations, des concepteurs de SRI, des indexeurs avec celles des utilisateurs (usager final et intermédiaire de l'information). Ainsi, pour Nicholas Belkin, le véritable problème n'est pas celui de la représentation des documents mais celui des caractéristiques cognitives et situationnelles qui déclenchent le processus de recherche d'information et l'utilisation d'un SRI. Il propose la notion de *Anomalous States of Knowledge* pour rendre compte des besoins individuels d'information (Belkin *et al.*, 1982).

De nombreux travaux s'inscrivent dans cette approche. Ceux de Marcia Bates proposent différents modèles décrivant le processus de recherche d'information et dont le point commun est d'insister sur la notion d'interactivité entre l'usager et le système. En 1981, elle propose un modèle qui décrit les tactiques de recherche (*search tactics*) mises en œuvre par les usagers des SRI en ligne. De leur côté, (Borgman, 1989) et (Fidel, 1991) proposent différentes analyses de l'interaction et du processus de *feedback* par les utilisateurs.

Les tenants de cette approche tentent d'expliquer comment les usagers organisent leur pensée et leur activité. Pour eux, les caractéristiques de l'utilisateur (styles cognitifs) représentent des mesures primaires à étudier et celles-ci restent constantes durant l'interaction avec le système. Un ensemble d'auteurs dont (Chen et Dhar, 1990) considère que la démarche de recherche des utilisateurs dépend des représentations qu'ils ont du monde qui les entoure. Quand elles sont fausses, ces représentations sont à l'origine de l'improductivité des recherches. Pour les optimiser, les usagers ont à réactualiser ces représentations. Parmi ces idées fausses, on peut citer :

- les représentations issues d'une connaissance inexacte du domaine informationnel et qui se traduisent par des problèmes dans la formulation des requêtes ;
- celles concernant les éventuels outils classificatoires (vocabulaires contrôlés, thésaurus, ontologies,...) et le niveau de connaissance du système d'indexation ;
- celles, enfin, qui concernent le fonctionnement du système qui entraîne une mauvaise interprétation des commandes du système et des fonctionnalités qu'il propose.

D'autres auteurs comme (Kerr, 1990) et (Shaw, 1991) soutiennent que les performances d'un système sont aussi influencées par l'interface et par les perceptions visuelles de l'utilisateur. Par exemple, le choix des résultats d'une requête est affecté par les marqueurs typographiques (taille, couleur, clignotement, icônes, etc.). Les travaux de (Belkin, 1990) visent à améliorer les interfaces des SRI en se fondant sur l'identification de cadres cognitifs et stéréotypes à partir de l'analyse des échanges lors des interactions.

À inclure dans cette catégorie, les travaux qui reposent sur l'hypothèse que les écarts dans les performances d'un SRI sont liés aux différences individuelles entre les usagers : expériences, objectifs, habileté verbale et spatiale, dépendance au champ, etc. (Borgman, 1989) et (Chen *et al.*, 2000).

L'approche cognitive stricte fait l'objet de trois types de critiques : le manque de réalisme de certains travaux, la méconnaissance d'apports extérieurs et le manque de vision globale.

Le manque de réalisme concerne essentiellement le fait que la plupart des études visaient à modéliser les structures cognitives des différents acteurs du processus informationnel sans tenir compte des contraintes techniques des SRI. Il semble que peu de chercheurs ayant adopté cette approche connaissent suffisamment en détail les différentes technologies (booléennes, vectorielles, ensemblistes, etc.) pour améliorer véritablement le fonctionnement des systèmes. Il est également lié au fait que beaucoup d'études se font en laboratoire et non dans des conditions réelles.

Le deuxième type de critique concerne la méconnaissance, par les tenants de l'approche cognitive individuelle, d'autres paramètres tels que le contexte socio organisationnel, l'affectif, la cognition située, etc. Ainsi ont été peu pris en compte les travaux de Dervin et Nilan, proposant la théorie de « Construction du sens » (*Sense-making theory*) (Dervin et Nilan, 1986) fondée sur une théorie de la communication, ceux de (Winograd et Flores, 1986) sur la conception cognitive des systèmes ou ceux de (Kuhlthau, 1993) sur la dimension affective de la RI.

Enfin, l'approche cognitive montre que si les différentes variables et facteurs (métacognition, différences interindividuelles, style cognitif, apprentissage, expertise, stratégies,...) jouent un rôle dans le processus de recherche d'information, elle reste néanmoins limitée par le fait qu'elle ne prend pas en compte l'interaction de ces différentes variables.

## **L'approche holistique**

Par approche holistique<sup>3</sup>, nous entendons un ensemble de travaux menés en sciences de l'information qui se caractérisent par la volonté de fournir un modèle plus global et plus cohérent des différents éléments du SRI, prenant en compte aussi bien l'analyse de l'usager du SRI et son environnement socio-organisationnel que les objets informationnels (bases de données, sources d'information, etc.), les différents composants du dispositif. L'approche holistique se distingue donc de l'approche cognitive individuelle par la volonté de construire une théorie unifiée de la recherche d'information. Cette approche se construit sur un certain nombre de considérations nouvelles insistant sur le rôle du contexte, de l'émotion, de l'incertitude et des interactions dans l'élaboration des processus cognitifs.

Plusieurs auteurs ont proposé des modèles s'inspirant de l'approche holistique parmi lesquels (Kuhlthau, 1993), (Ingwersen, 1996), (Saracevic, 1996) et (Wilson, 1999). Nous ne prétendons pas être exhaustifs ni présenter ces modèles en détail mais souhaitons donner des exemples significatifs de la diversité du paradigme usager en soulignant les éléments qui les caractérisent. Pour une présentation et une discussion plus détaillées de ces modèles comportementaux (cf. Chaudiron, 2001).

### ***Le modèle de Carol Kuhlthau***

Pour Carol Kuhlthau, qui s'intéresse essentiellement à la recherche d'information médiatisée par un expert, un usager qui recherche de l'information s'inscrit dans un processus d'apprentissage dont le succès résulte de ses stratégies cognitives et de l'aide reçue. Influencée par les travaux de constructivistes tels que Bruner, Kelly, Dewey et Taylor, (Kuhlthau, 1993) a établi un modèle selon lequel la recherche d'information est un processus de *construction du sens* à partir de l'information trouvée. Son modèle segmente le processus de recherche en six étapes (initiation, sélection d'un thème, identification, exploration, formulation, collection, présentation). Elle y incorpore trois dimensions qui sont communes aux six étapes : la dimension affective, cognitive et physique.

Le processus de recherche d'information y est enclenché par un état d'incertitude dû à un manque de compréhension, à un sens inexpliqué, à une structure incomplète. Il s'agit d'un état de nature cognitive qui provoque des symptômes affectifs comme l'anxiété, le manque de confiance ou le contentement. En soulignant le rôle majeur de la dimension émotionnelle dans le processus de recherche, elle montre que les usagers ont des attentes différentes mais surtout que leur comportement, leur perception de la difficulté de recherche et leur état évoluent au cours du processus. À chaque étape, l'usager passe de l'incertitude à la satisfaction ou à l'insatisfaction. Elle montre ainsi que ces états affectifs influencent la manière dont les usagers recherchent et utilisent l'information. Elle suggère donc aux bibliothécaires de fournir une aide appropriée selon l'état affectif de l'usager.

Si ce modèle apporte de nouveaux éclairages dans le cas d'une recherche médiatisé par un expert à travers la prise en compte des états cognitifs et affectifs des usagers, il offre peu de pistes dans le cas d'une interaction usager-SRI.

### ***Le modèle de Peter Ingwersen***

L'intérêt de ce modèle est de proposer une vision globale des différents éléments impliqués dans la recherche d'information, au-delà du seul SRI, et de montrer que le processus est dynamique. Il propose une nouvelle approche de la pertinence qui s'appuie sur le constat qu'une même source d'information est perçue différemment par les usagers, en fonction de leurs structures cognitives propres. Cette variabilité de la représentation de l'objet est appelée par P. Ingwersen la « polyreprésentation des objets informationnels » (Ingwersen, 1996). Un même document par exemple est perçu et représenté différemment par le créateur du document, le concepteur du thésaurus utilisé pour l'indexer, l'indexeur, le concepteur du SRI et l'utilisateur du système.

Un SRI n'est donc plus simplement le lieu où s'apparient différentes structures cognitives dans le cadre d'interactions (entre l'utilisateur et le système essentiellement). Il est considéré comme un système plus global dans lequel entrent en jeu de multiples variables : l'espace cognitif des différents acteurs, les caractéristiques contextuelles psychologiques, sociales et organisationnelles, l'évolutivité de la représentation des sources d'information, le changement des besoins d'information et les types de médiation proposés par les systèmes techniques.

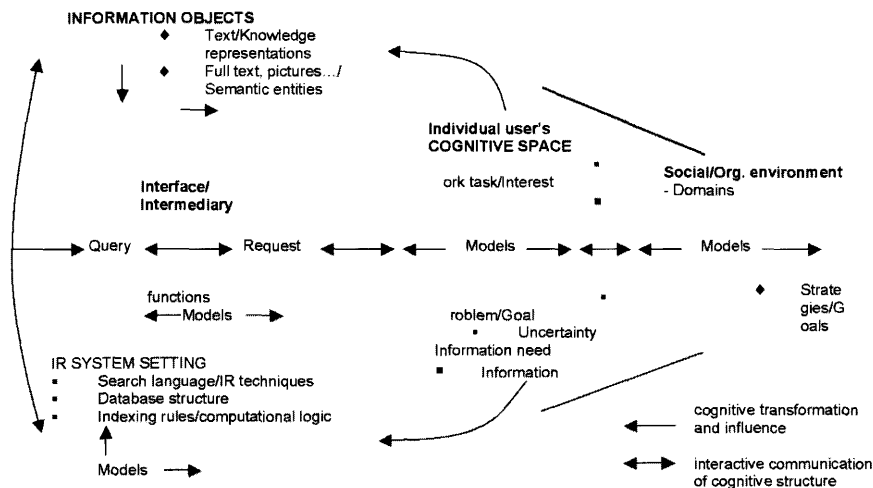


Figure 1. Le modèle de P. Ingwersen.

### ***Le modèle de Tefko Saracevic***

Le modèle de Saracevic propose également une vision globale du processus de recherche. Il considère en particulier (Saracevic, 1997b) qu'une information doit être considérée comme un message traité d'un point de vue cognitif par un usager en fonction d'un contexte donné.

(Saracevic, 1997a) envisage donc le contexte réel de l'activité des usagers dans sa dynamique et sa complexité mais sans faire l'impasse sur l'impact des opérations de traitement de l'information (indexation, classification) et des modes d'appariement utilisés. La pertinence d'un document ne s'apprécie pas selon une seule dimension mais se caractérise au contraire de façon dynamique par l'interrelation entre différentes composantes :

- contextuelle (informations liées aux contextes d'usages : tâches, définition du problème) ;
- intentionnelle (analyse des croyances et motivations des usagers) ;
- cognitive (représentation des connaissances de l'usager) ;
- requête (caractéristiques des questions de l'usager) ;
- interface homme-machine ;
- système du logiciel de recherche (puissance, mémoire, etc.) ;
- logiciel et algorithmique (composants du SRI) ;
- fonds documentaire (structure et caractéristiques de la base documentaire).

### ***Le modèle de Tom Wilson***

T. Wilson a proposé en 1996 un modèle (Wilson, 1996) général du comportement informationnel qui s'appuie sur trois éléments :

- l'origine du besoin informationnel, c'est-à-dire les facteurs qui ont conduit à la perception du besoin par l'individu ;
- les facteurs qui déterminent la réponse de l'individu en réaction à la perception du besoin ;
- les processus ou les actions impliqués par cette réponse.

Ce modèle enrichit celui proposé en 1981 en introduisant d'autres variables (psychologiques, démographiques, environnementales, etc.) qui permettent d'expliquer les facteurs qui facilitent ou au contraire freinent l'accès à l'information.

Wilson s'appuie sur des travaux issus d'autres disciplines pour rendre compte de certains aspects particuliers du comportement. Il introduit d'autres éléments dans le modèle expliquant comment associer

une stratégie de recherche à un besoin d'information (*stress/coping theory*), pourquoi certaines sources d'informations sont utilisées et d'autres non (*risk/reward theory*) et comment l'utilisateur procède par apprentissage pour adapter son comportement à celui qui est requis pour arriver au résultat escompté (*social learning theory*).

Ce modèle présente l'intérêt d'appréhender l'utilisateur en situation de recherche d'information de manière beaucoup plus globale que dans le modèle initial et pose en particulier les questions de la décision de lancer une stratégie de recherche et du choix des sources d'information qui sont en effet des questions essentielles «en amont» de la procédure de recherche elle-même.

Le modèle révisé de Wilson est le suivant :

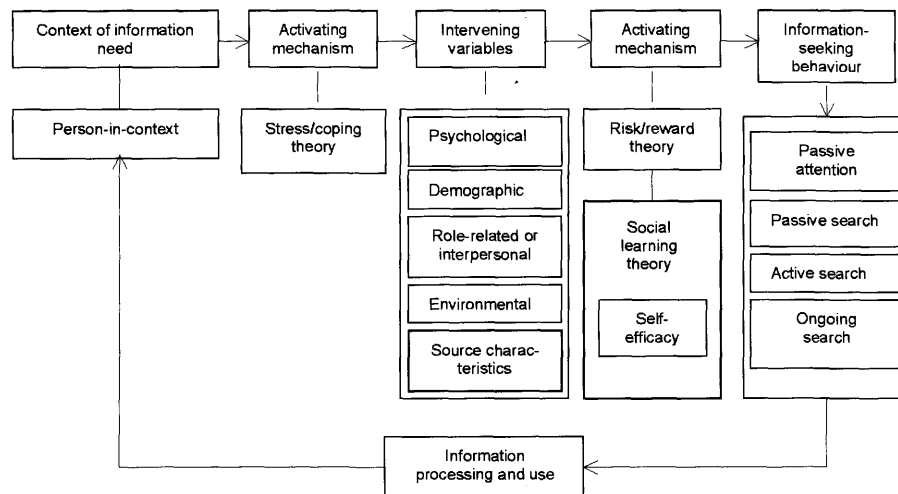


Figure 2. Le modèle de T. Wilson.

## Conclusion

L'approche cognitive constitue l'une des alternatives à l'approche système de l'évaluation des SRI. Néanmoins, cette approche reste limitée car les processus de repérage, de recherche et de navigation sont occasionnés par d'autres facteurs, notamment contextuels. D'un point de vue méthodologique, l'approche cognitive a la plupart du temps considéré l'utilisateur comme un individu isolé, en négligeant ses dimensions

sociale et culturelle. De plus, les études de l'approche cognitive stricte, souvent réalisées « en laboratoire » ne portent donc pas sur des tâches réelles.

Les approches holistiques présentées ici fournissent un cadre « conceptuel » permettant d'explicitier les logiques d'usage en situation de repérage et de recherche d'information. En ce sens, elles constituent un dépassement de l'approche cognitive stricte et offrent de nouvelles pistes de recherche, tant d'un point de vue théorique que méthodologique même si l'idée d'un modèle explicatif unique en matière de RI ne semble plus pertinent et si on peut légitimement s'interroger sur les difficultés méthodologiques de mise en œuvre d'une telle approche et sur ses conditions de validation.

## NOTES

1. Voir par exemple (Belkin, 1990).
2. De Mey, 1977, p. XVI-XVII.
3. Du grec holos = qui forme un tout. Conception de la société qui fait que le tout domine les parties, qui ne sont alors que des composantes secondaires, soudées les unes aux autres, interdépendantes, subordonnées à une fin qui leur préexiste.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- BATES, M., « Search Tactics », in *Annual Review of Information Science and Technology*, M. WILLIAMS (dir.), vol. 16, 1981, ASIS, p. 139-169.
- BELKIN, N., « Cognitive Viewpoint in Information Science », in *Journal of Information Science*, 1990, vol. 16, n° 1, p. 11-15.
- BORGMAN, C., « All Users of Information Retrieval Systems are not Created Equal: an Exploration into Individual Differences », in *Information Processing and Management*, 1989, vol. 25, n° 3, p. 237-252.
- CHEN, H, DHAR, V., « User Misconceptions of Information Retrieval Systems », in *International Journal of Man-Machine Studies*, 1990, vol. 32, n° 6, p. 673-692.
- CHEN, C., CZERWINSKI, M., MACREDIE, R., « Individual Differences in Virtual Environments: Introduction and Overview », in *Journal of the American Society for Information Science*, 2000, vol. 51, n° 6, p. 499-507.
- CHAUDIRON, S., *L'Évaluation des systèmes de traitement de l'information textuelle: vers un changement de paradigme*, mémoire pour HDR, université de Paris X, novembre 2001, 300 p.
- CHAUDIRON, S., IHADJADENE, M., « Quelle place pour l'utilisateur dans l'évaluation des SRI? », in *Recherches récentes en sciences de l'information: convergences et dynamiques*, actes du colloque MICS-LERASS, 21-22 mars 2002, Toulouse, Paris, ADBS éd., p. 211-213.
- DE MEY, M., « The Cognitive Viewpoint: its Developments and its Scope », in *International Workshop on the Cognitive Viewpoint*, sous la direction de M. De Mey, R. Pinxten, M. Poriau, F. Vandamme, 24-26 mars 1977, Gand (Belgique), p. XVI-XXXI.



- DE MEY, M., «The Relevance of the Cognitive Paradigm for Information Science», in *Theory and Application of Information Research: Proceedings of the 2<sup>nd</sup> International Research Forum on Information Science*, 3-6 août 1977, Copenhagen, Royal School of Librarianship, 1980, p. 48-61.
- DERVIN, B., NILAN M., «Information Needs and Uses», in *Annual Review of Information Science and Technology*, sous la direction de M. Williams, vol. 21, 1986, p. 3-33.
- FIDEL, R., «Searcher's Selection of Search Keys (part I) : the Selection Routine», in *Journal of the American Society for Information Science*, 1991, vol. 42, n° 7, p. 490-500.
- FIDEL, R., «Searcher's Selection of Search Keys (part II) : Controlled Vocabulary or Free-Text Searching», in *Journal of the American Society for Information Science*, 1991, vol. 42, n° 7, p. 501-514.
- FIDEL R., «Searcher's Selection of Search Keys (part III) : Searching Styles», in *Journal of the American Society for Information Science*, 1991, vol. 42, n° 7, p. 515-527.
- IHADJADENE, M., *La Recherche et la navigation dans un SRI grand public*, thèse de doctorat en science de l'information, Université Lyon 1, 1999.
- INGWERSEN, P., «Cognitive Perspectives of Information Retrieval Interaction : Elements of a Cognitive IR Theory», in *Journal of Documentation*, vol. 52, n° 1, 1996, p. 3-50.
- KERR, S.-T., «Wayfinding in an Electronic Database: the Relative Importance of Navigational Cues vs. Mental Models», in *Information processing and management*, 1990, vol. 26, n° 4, p. 511-523.
- KUHLTHAU, C., *Seeking Meaning: a Process Approach to Library and Information Services*, Norwood, Ablex Pub. Corp., 1993.
- SARACEVIC, T., «The Stratified Model of Information Retrieval Interaction: Extension and Applications», in *Proceedings of ASIS'97*, 1997a, vol. 34, p. 313-327.
- SARACEVIC, T., «Users Lost: Reflections on the Past, Future, and Limits of Information Science», in *SIGIR Forum*, 1997b, vol. 31, n° 2, p. 16-27.
- SHAW, D., «The Human-Computer Interface for Information Retrieval», in *Annual Review of Information Science and Technology*, 1991, vol. 26, 155-195.
- SPINK, A., WOLFRAM, D., JANSEN, B. J., SARACEVIC, T., «Searching the Web: the Public and their Queries», in *Journal of the American Society for Information Science*, vol. 53, n° 2, 2001, p. 226-234.
- WILSON, T., «Information Behaviour: an Inter-Disciplinary Perspective», in *Proceedings of an International Conference on research in information needs, seeking and use in different contexts*, sous la direction de Vakkari Pertti, Savolainen Reijo et Dervin Brenda, Taylor Graham, 1996, p. 39-49.
- WILSON, T., «Models in Information Behaviour Research», in *Journal of Documentation*, 1999, vol. 55, n° 3, p. 249-270.
- WINOGRAD, T., FLORES F., *Understanding Computers and Cognition: a New Foundation for Design*, Norwood, Ablex, 1986.