



HAL
open science

La pertinence en sciences de l'information : des modèles, une théorie ?

Brigitte Simonnot

► To cite this version:

Brigitte Simonnot. La pertinence en sciences de l'information : des modèles, une théorie?. Problématiques émergentes dans les Sciences de l'Information, Hermès Lavoisier, pp.161-182, 2008. sic_00496291

HAL Id: sic_00496291

https://archivesic.ccsd.cnrs.fr/sic_00496291v1

Submitted on 30 Jun 2010

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.

Citer ce texte :

Simonnot B., 2008, « La pertinence en sciences de l'information : des modèles, une théorie ? », pp. 161-182 ch 6 in Papy F. (dir.), *Problématiques émergentes dans les Sciences de l'Information*, Paris : Hermes-Lavoisier.

La pertinence en sciences de l'information : des modèles, une théorie ?

Brigitte Simonnot

.1. Introduction

En sciences de l'information, le concept de pertinence est depuis longtemps une préoccupation des chercheurs. Bien qu'ayant fait l'objet de nombreuses publications, il reste encore une notion ambiguë et est périodiquement réinterrogé : s'il est au centre de la problématique de la recherche d'information, mieux le définir permettrait certainement de mieux comprendre la façon dont les individus se comportent lorsqu'ils cherchent de l'information. Depuis la fin des années 1990, ce concept a été proposé comme un candidat possible à l'élaboration d'une théorie permettant d'unifier dans un cadre commun les travaux sur le comportement de recherche des usagers de l'information (en sciences humaines et sociales) et ceux qui cherchent à concevoir des systèmes informatisés opérationnels pour les aider dans leurs quêtes (en informatique et en intelligence artificielle). Les scientifiques s'accordent en effet sur ce point : le but d'un système de recherche d'information (SRI) est de permettre à un individu de (re)trouver l'information pertinente, censée être présente dans une collection de documents.

La notion de pertinence est intimement liée à la notion de besoin d'information. Les besoins d'information des individus sont de différentes natures et sont plus ou moins explicites. Ainsi, un besoin d'information peut être viscéral pour reprendre le terme de Taylor, c'est-à-dire ressenti confusément, ou bien conscient c'est-à-dire que sa nature est mieux définie par celui qui le ressent. Il est exprimé par une demande, lorsque l'utilisateur s'adresse à tiers dans le langage habituel, ou par une requête lorsque l'utilisateur utilise un SRI : la requête est une traduction de la demande dans un langage que le SRI peut interpréter : c'est l'étape de formulation. La formulation provoque des déformations par rapport au besoin (c'est l'effet *label* mis en évidence par Taylor) : d'une part, l'individu n'exprime que ce qu'il pense important pour avoir une réponse et d'autre part, il a tendance à en dire le moins possible sur son besoin informationnel. L'interrogation par mots-clés oblige aussi l'utilisateur à évacuer une partie du sens de sa demande. La formulation est donc affectée par le vocabulaire employé, la diversité et l'ambiguïté du langage naturel : ce sont les problèmes d'homonymie, de synonymie mais aussi de spécificité ou de généralité du vocabulaire bien connus des documentalistes.

Pour répondre à la requête, le SRI présente à l'utilisateur un ensemble de résultats. Ces résultats sont en général d'abord affichés sous une forme abrégée. Il peut s'agir d'une partie de la notice du document dans un catalogue de bibliothèque ou, dans les pages de résultats des banques de publications, d'une courte suite d'éléments descriptifs (souvent tirés de la référence bibliographique). Ce peut être aussi un *snippet* comme dans les moteurs de recherche où l'on retrouve le titre, l'adresse du document et parfois la date de publication ainsi qu'un bref résumé ou quelques bribes du document contenant les termes de la requête.

Les travaux en recherche d'information (*information retrieval*) se sont longtemps focalisés sur les meilleures façons de représenter les documents et sur les algorithmes d'appariement entre un ensemble de documents et une requête donnée. Nous n'évoquerons pas ici les différents modèles algorithmiques élaborés ni la manière de représenter les documents qui ont beaucoup évolué pour passer des catalogues structurés des bibliothèques à la recherche en texte intégral dans d'énormes collections de documents. Pour une synthèse historique de ces approches, le lecteur pourra se reporter à [CHI 07]. Les recherches en sciences humaines ont discuté et remis en

cause l'objectif des SRI depuis les années 1990 : si les premiers systèmes se fixaient pour but de répondre à une requête ponctuelle, il est apparu important de les faire évoluer pour qu'ils tentent de répondre au besoin d'information de l'utilisateur. Avec le développement des SRI dédiés au grand public – dont les moteurs de recherche en ligne sont les derniers avatars –, cette question s'est posée avec davantage d'acuité : les premiers logiciels étaient conçus pour être utilisés par des personnes formées aux langages documentaires, censées être capables de formuler des requêtes appropriées. Les SRI actuels, dédiés au grand public, ne peuvent pas compter sur l'expertise de l'utilisateur en matière de formulation.

Satisfaire un usager de l'information, c'est lui proposer des documents pertinents. Mais qu'est-ce au juste que la pertinence ? Approfondir ce concept entraîne dans une longue quête où la maîtrise du vocabulaire n'est pas le moindre obstacle. Si, de manière intuitive, chacun comprend bien ce dont il s'agit, la question se complique lorsqu'on cherche à le modéliser. La langue anglaise distingue deux termes : *relevance* et *pertinence*¹. Tous les chercheurs ne s'accordent pas exactement sur la définition précise de ces deux termes. Le terme anglais *relevance*, par son étymologie, est apparenté au verbe « relever » au sens de « être du ressort de » mais aussi « mettre en relief ». La *relevance* serait ce qui, pour une question donnée, permet de distinguer un objet informationnel des autres. Le terme *pertinence*, lui, vient du latin *pertinere*, composé à partir du verbe *tenere* qui signifie « se souvenir de quelque chose » mais aussi « comprendre, conserver dans son esprit » et du préfixe *per* qui signifie « à travers, au moyen de ». Il désigne étymologiquement l'action ou le fait de comprendre grâce à quelque chose. Du même verbe *tenere* est d'ailleurs dérivé le terme « attention » (tendre son esprit vers). Dans la suite de ce chapitre, nous considérerons que la *relevance* concerne l'adéquation d'un document ou d'un objet informationnel à un besoin d'information ou à une demande d'information en général. La *pertinence* sera considérée comme l'adéquation d'un document ou d'un objet informationnel à la demande d'un individu donné : pour qu'un document *relevant* soit pertinent pour cet individu, il doit être compréhensible par lui et lié aux connaissances dont il dispose sur le sujet.

6.2. La nature de la pertinence

Cerner la nature de la pertinence, tel a été l'objet de nombreux travaux de recherche en sciences humaines et sociales pour prôner une meilleure prise en compte de l'utilisateur dans la conception des SRI.

6.2.1 La pertinence-sujet ou thématique (*topicality*)

Les premiers SRI se contentaient d'apparier les termes de la requête à ceux décrivant – ou, dans le cas de la recherche en texte intégral, contenus dans – les documents de la collection selon l'hypothèse suivante : si un terme ou un ensemble de termes se retrouvent à la fois dans une requête et dans un document (ou sa description), ce document est pertinent vis-à-vis de la requête. Au fur et à mesure que les collections sont devenues plus vastes – et avec l'application de la RI au web –, il est devenu important de classer les documents candidats pour proposer à l'utilisateur un ensemble ordonné présentant les « meilleurs » documents en tête. Il ne suffisait plus de « récupérer » les documents susceptibles d'être pertinents mais de trouver le moyen de les distinguer, d'en faire émerger certains par rapport aux autres, de les classer. Cette pertinence-sujet ou thématique (*topicality*) a été longtemps la seule prise en compte dans les algorithmes. Il faut dire que cela correspondait à l'époque où les documents étaient indexés manuellement par des documentalistes qui cherchaient à en déterminer le topique, le sujet (ce dont on dit quelque chose).

Dans les années 1990, un certain nombre de scientifiques, à partir d'études du comportement informationnel des individus, ont montré que la pertinence-sujet ou thématique n'était pas le seul critère à prendre en compte pour répondre au besoin informationnel de l'utilisateur. Harter [HAR 92] soulignait que le sens commun de la pertinence ne correspondait pas à la pertinence-sujet ou thématique : en effet, un document qui ne traite pas du même sujet que celui exprimé dans la requête de l'utilisateur peut aussi être pertinent pour lui. La pertinence-sujet n'est qu'un critère parmi d'autres, même s'il reste très important dans la plupart des situations de recherche [GRE 95]. Pour tenter d'identifier les autres variables qui interviennent dans les jugements de pertinence, de nombreuses études empiriques ont été menées sur les utilisateurs, en laboratoire ou en contexte réel.

¹ L'italique signale les termes en langue étrangère dans ce chapitre pour les distinguer des termes en français.

6.2.2 La nature dynamique de la pertinence

Le processus de recherche d'information a d'abord été modélisé sous forme linéaire par une suite d'étapes : formulation de la requête – appariement par le SRI avec la collection de documents – affichage des résultats. Rapidement, la nature itérative du processus s'est imposée. En *information retrieval*, Rocchio avait introduit dès 1966 le principe de rétroaction de pertinence (*relevance feedback*) qui permettait d'implanter dans les SRI une fonction de reformulation automatique de la requête initiale en s'appuyant sur le choix des documents pertinents/non pertinents faits par l'utilisateur. Markey et Atherthon (1978) avaient aussi remarqué que les chercheurs d'information s'inspirent des termes trouvés dans les premiers documents retrouvés pour étendre, poursuivre ou affiner leur recherche. La rétroaction de pertinence peut porter sur les documents entiers ou sur les termes d'indexation des documents. Il s'agissait principalement, dans ces tentatives, d'aider l'utilisateur dans sa démarche d'interrogation en travaillant sur le vocabulaire utilisé pour interroger le SRI.

Mais lorsqu'un individu cherche de l'information, les documents qu'il trouve peuvent modifier non seulement la connaissance du vocabulaire approprié pour interroger le système mais aussi modifier son besoin, et ce, d'autant plus que ce dernier est flou et mal défini. Si le SRI doit répondre au besoin et non seulement à la requête telle qu'elle est formulée à un moment donné, il doit prendre en compte cette évolution. Bates [BAT 89] a fortement remis en cause la vision d'un processus de recherche où le besoin de l'utilisateur reste toujours le même. Dans la vie réelle, les individus procèdent par étapes successives, et l'information qu'ils collectent à chaque étape modifie ou amène de nouvelles questions. Elle nomme ce comportement de recherche le *berry-picking* : c'est par cueillette de bribes d'informations successives que les individus résolvent leur problème. Comme dans la forêt, les baies ne se cueillent pas par branches entières : le cueilleur prélève çà et là quelques baies pour remplir son panier en essayant de choisir les meilleures. Harter [HAR 92] a également développé l'analyse de l'aspect dynamique de la pertinence : s'appuyant sur la théorie de la pertinence de Sperber et Wilson [SPE 85], il développe la vision d'un processus de recherche d'information durant lequel l'individu passe d'un contexte initial, avec une certaine représentation de son besoin d'information, à d'autres contextes qu'il construit peu à peu au fur et à mesure de sa démarche. Prendre en compte la session complète de recherche – et non simplement la soumission d'une requête ponctuelle – est donc important pour concevoir des SRI qui répondent à un besoin d'information (et pas seulement à une requête ponctuelle). Au cours d'une session de recherche, les documents pertinents sont à la fois ceux qui aident l'individu à mieux cerner son problème, à poser les questions de manière différente, et ceux qui apportent l'information nécessaire pour y répondre.

6.2.3 La nature situationnelle de la pertinence

C'est Patrick Wilson qui semble être le premier à avoir souligné l'importance de prendre en compte la situation du chercheur d'information pour l'aider dans sa recherche. La pertinence situationnelle fait référence à la façon dont les individus utilisent l'information et aux effets réels de l'information sur leur changement de vision du problème. Elle recouvre l'utilité, la valeur pragmatique d'un document pour une tâche ou un problème à résoudre [XU 06]. Pour une même requête (c'est-à-dire formulée avec les mêmes termes), ce n'est pas forcément le même ensemble de documents qui sera pertinent pour des individus qui ont des tâches différentes à résoudre ou même pour le même individu dans deux situations différentes. Mais comment définir et caractériser le contexte d'une recherche d'information ? Si les documentalistes peuvent interroger l'usager sur ce qu'il veut faire avec l'information trouvée pour mieux l'aider, implanter une telle fonctionnalité dans un SRI pose de nombreux problèmes. En psychologie cognitive, le contexte a été principalement défini à partir de la tâche que l'individu a à faire. Du point de vue bibliométrique, le contexte peut comprendre les documents qui entourent ou sont liés à un document pertinent, par exemple par le moyen des citations. Du point de vue technologique, le contexte prend en compte l'environnement technique et matériel de l'individu pour lui présenter des documents dans des formats adaptés (par exemple, qu'un téléphone portable peut afficher). Du point de vue géographique, la notion de contexte peut donner lieu à des recherches locales : si un individu cherche un service particulier, il peut préférer que ce service soit à proximité de l'emplacement où il se trouve. Bref, le notion de contexte en recherche d'information est multiple et ses acceptions divergent.

6.3. Les modèles multidimensionnels de la pertinence

Pour caractériser les différents aspects de la pertinence, les chercheurs ont tout d'abord essayé de décomposer le concept en sous-catégories. Il est en effet rapidement apparu que la pertinence se manifestait à différents niveaux durant un processus de recherche et qu'à chaque niveau, différents éléments jouaient sur les jugements

de pertinence des utilisateurs. Tefko Saracevic s'est attaché à analyser les manifestations de la pertinence pour en distinguer les différents attributs. Son modèle a été revisité par Cosign et Ingwersen. Mizzaro a proposé un modèle élégant intégrant différentes relations de pertinence. Enfin Wang et Soergel se sont focalisés sur le processus de prise de décision des utilisateurs lorsqu'ils choisissent ou rejettent un document proposé par un SRI. Nous revenons plus en détail sur ces différents efforts de modélisation où la pertinence joue un rôle central.

6.3.1 Le modèle d'interaction stratifié de Saracevic (1996)

Saracevic a beaucoup travaillé sur le concept de pertinence depuis 1975. S'inspirant des travaux de Schutz en philosophie et de ceux de Sperber et Wilson en pragmatique linguistique, il souligne que la notion cognitive de pertinence implique la création interactive et dynamique de relations par inférence², l'utilisateur étant doté d'intentions et situé dans un certain contexte. Selon lui, la pertinence peut être caractérisée par un certain nombre d'attributs : la relation, l'intention et la motivation, le contexte, l'inférence – appréhendée à travers des jugements le plus souvent gradués sur l'efficacité ou le degré de maximisation d'une relation donnée–, et l'interaction. Il propose de modéliser l'interaction entre l'utilisateur et le SRI comme une suite de processus se produisant dans un système de strates [SAR 96]. Son modèle révisé [SAR 97] distingue différents types de manifestations de la pertinence (*relevance*) :

- la pertinence système ou algorithmique, c'est-à-dire l'évaluation par le système de l'adéquation entre documents et requête ;
- la pertinence-sujet ou thématique (*topicality*), relation entre le sujet de la requête et celui des documents ;
- la pertinence cognitive (*pertinence*), relation entre l'état des connaissances et le besoin d'information de l'utilisateur ; elle est inférée à partir de la correspondance avec ses connaissances (l'information doit être adaptée au niveau de compréhension de l'utilisateur), de l'informativité, de la nouveauté, de la qualité de l'information et des goûts de l'utilisateur ;
- la pertinence situationnelle ou utilité (*utility*), relation entre la tâche ou le problème à résoudre et les textes retrouvés ; elle est inférée à partir de leur valeur d'usage pour la prise de décision, de l'adéquation de l'information pour résoudre le problème, de la réduction d'incertitude et des goûts de l'utilisateur ;
- la pertinence affective ou motivationnelle, relations entre les intentions, les buts et les motivations de l'utilisateur et les textes ; elle est inférée à partir de la satisfaction, du succès, de la réussite et des goûts de l'individu.

Saracevic souligne l'interdépendance dans ce système de pertinences³ – au sens où un type de pertinence affecte les autres – et plaide pour qu'aucune ne soit ignorée dans la conception des systèmes.

Le modèle stratifié permet de distinguer différents niveaux qui interviennent dans les jugements de pertinence des utilisateurs. L'auteur le présente comme un point de départ pour mieux modéliser l'utilisateur d'un SRI et améliorer les retours du système suite à son interrogation.

Néanmoins, une telle représentation linéaire ne rend pas compte de toutes les possibilités de relations qui s'établissent entre les différentes strates. Saracevic revient sur ce modèle dans une double publication récente [SAR 07A ; 07B] et affine sur un certain nombre de points les enseignements apportés par des recherches plus récentes.

² L'inférence en psychologie se distingue du raisonnement en ce que les éléments qui la sous-tendent ne sont pas forcément explicites.

³ Saracevic a emprunté à Schutz cette expression de « systèmes de pertinence ».

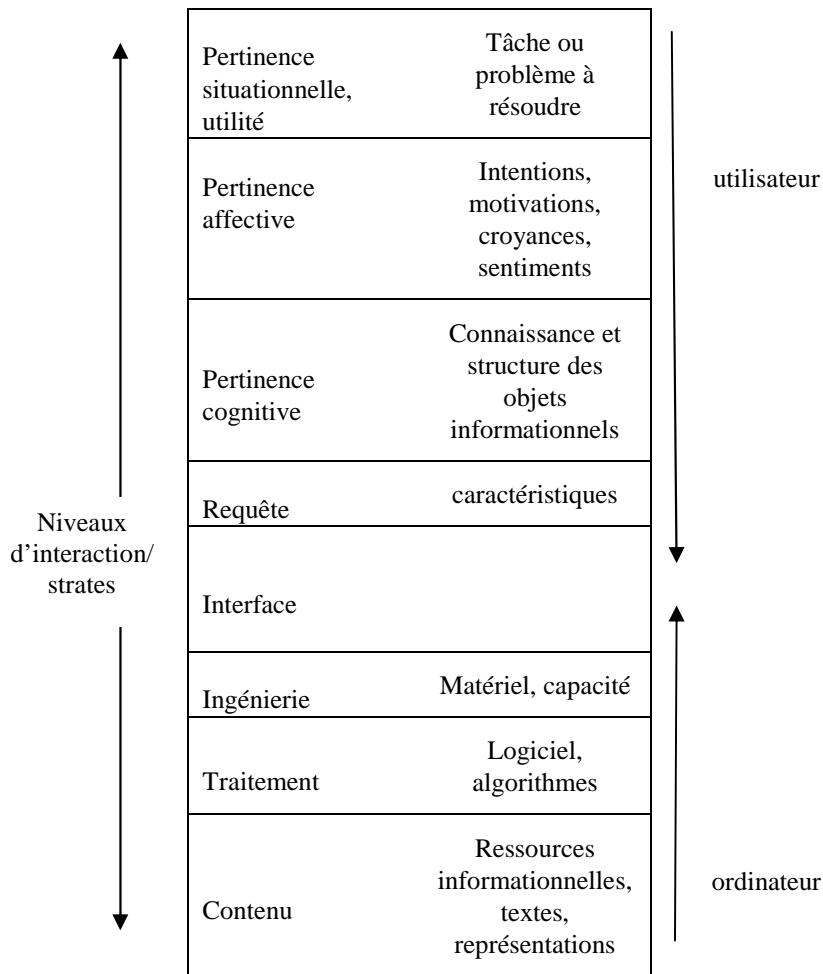


Figure 6.1 Le système de strates de Saracevic (1997)

6.3.2 Le modèle revisité par Cosijn et Ingwersen

Cosijn et Ingwersen [COS 00] tentent une synthèse des différentes dimensions du concept de pertinence. A partir du modèle stratifié de Saracevic, ils proposent le terme de pertinence sociocognitive pour désigner la relation établie entre une situation, une tâche à effectuer ou un problème à résoudre dans un contexte socioculturel donné et les objets informationnels, tels qu'ils sont perçus par un ou plusieurs agents cognitifs. Cette vision de la pertinence englobe à la fois le système, un groupe d'utilisateurs ou agents individuels et leur environnement socio-organisationnel. L'exemple de pertinence sociocognitive qu'ils donnent est le processus de relecture par des pairs dans le domaine des publications scientifiques, qui aboutit à un classement des publications sur lequel s'accordent les relecteurs. Dans ce processus, différents facteurs culturels peuvent influencer les évaluateurs. Les auteurs classent la pertinence sociocognitive comme un type de pertinence subjective déterminée par un acteur en interaction avec d'autres acteurs dans une communauté. Cette intersubjectivité peut montrer des caractéristiques statistiquement objectives et mesurables. Selon eux, la pertinence situationnelle est différente de la pertinence sociocognitive parce qu'elle est purement subjective. Ces deux types de pertinence diffèrent principalement vis-à-vis des attributs que sont l'intention, l'inférence et l'interaction. Leur article discute largement le caractère subjectif ou non des différents types de pertinence.

Cosijn et Ingwersen caractérisent la pertinence cognitive par la nouveauté, le degré d'informativité, les préférences et la qualité des objets informationnels relatifs au besoin d'information de l'utilisateur à un moment donné. Ils souhaitent différencier la pertinence affective et la pertinence motivationnelle. En s'appuyant sur Searle (1984), ils distinguent deux niveaux d'intentionnalité. Le premier niveau serait un phénomène intentionnel intrinsèque, constitué des croyances et, pourrait-on dire, des représentations internes des individus. Le second niveau correspondrait en une intentionnalité dérivée à partir du niveau précédent et qui comprendrait les objectifs, les rôles et les attentes de l'utilisateur.

Pour ces auteurs, la pertinence affective est personnelle et donc très subjective. Elle est liée au point de vue de l'individu et concerne ses réponses émotionnelles à n'importe quel aspect du document. Selon eux cependant, même si elle a des effets sur les autres types de pertinence, elle ne peut pas être classée de façon claire comme une catégorie à part car elle agit de manière totalement différente. Chaque type de pertinence peut ne pas être totalement mesurable : on peut disposer d'une variable opérationnelle mais ne pas savoir réellement quelle est son association avec la variable théorique sous-jacente. Par exemple, écrivent-ils, il peut être difficile de distinguer expérimentalement la pertinence cognitive de la pertinence situationnelle, sauf à tenter de mesurer l'effet d'apprentissage obtenu durant une session de recherche.

Tout en reconnaissant l'importance du contexte pour la compréhension du concept de pertinence, Cosijn et Ingwersen ne définissent pas vraiment la notion mais se contentent d'en donner quelques exemples. Ils évoquent le contexte de test des algorithmes de RI, par exemple celui des conférences TREC. Pour ces auteurs, le contexte contient la situation donnée et/ou la stratégie organisationnelle sous-jacente à l'environnement de l'utilisateur mais n'est pas un attribut suffisant pour distinguer les différents types de pertinence.

6.3.3 Le modèle de Mizzaro (1998)

En 1998, Mizzaro [MIZ 98], à partir d'une revue conséquente de littérature sur la question, proposait un modèle comportant quatre dimensions de la RI : les ressources informationnelles, la représentation du besoin d'information de l'utilisateur, le temps et les « composants » contextuels ou situationnels. Il est utile de revenir en détail sur ce modèle qui a clarifié un certain nombre de points et a été affiné par la suite [GAB 99].

Les ressources informationnelles auxquelles peut accéder l'utilisateur d'un SRI sont de différents types :

- l'ensemble de documents (SD) qui, ensemble, pourraient satisfaire son besoin ;
- le document lui-même (D), c'est-à-dire l'entité physique retrouvée par l'utilisateur, ou un objet informationnel complet auquel l'utilisateur peut accéder grâce à sa recherche ;
- les métadonnées (MD), des données structurées sur le document qui concernent aussi bien sa description bibliographique que des indications ou suggestions faites par des humains (comme dans le cas de LOM, qui spécifie le niveau d'étude auquel l'objet se rapporte) ;
- le représentant ou substitut du document ($S = \textit{surrogate}$), c'est-à-dire par exemple la notice du catalogue ou les fragments (*snippet*) présentés dans les pages de résultat des moteurs de recherche.

Mizzaro propose de classer ces différents objets informationnels dans l'ordre partiel suivant : $S \leq MD \leq D \leq SD$ par rapport à la capacité à répondre au besoin informationnel de l'utilisateur. Par exemple, la notice d'un document peut comporter un résumé qui laisse l'utilisateur penser que le document peut être intéressant pour répondre à sa demande, mais seule la lecture du document proprement dit apportera véritablement ou non l'information désirée.

La seconde dimension est la représentation du problème informationnel de l'utilisateur : Elle comporte :

- le besoin d'information réel (RIN) - Taylor parlait de besoin « viscéral » au sens où il n'est pas réellement conscient dans l'esprit de l'utilisateur ; c'est une abstraction d'un besoin idéal qui inclut l'ensemble complet d'information que la personne devrait avoir sur sa situation et sur son besoin d'information pour que l'information puisse la résoudre de manière appropriée. Inclure le besoin d'information réel dans le modèle est important : un expert, par exemple, pourrait aider l'utilisateur à comprendre que le besoin d'information qu'il perçoit ne correspond pas à la réalité ;
- le besoin perçu (PIN) ou conscient, de manière plus ou moins correcte et souvent imprécise ; c'est une représentation implicite de la situation problème et des différents moyens possibles pour satisfaire le besoin ;
- la demande (EIN), c'est-à-dire l'expression du besoin en langage « naturel » ;
- et la requête (FIN), soit la demande exprimée sous une forme que le SRI puisse traiter ;

Ces différents éléments peuvent être ordonnés ainsi : $FIN \leq EIN \leq PIN \leq RIN$, selon la complétude de leur description et leur capacité à exprimer le besoin réel. Par exemple, un individu peut avoir une représentation erronée de ce qu'il doit chercher. Il peut être conscient d'avoir besoin d'information sur tel aspect du sujet mais pas sur un autre qui est pourtant important pour bien le traiter. Par conséquent, son besoin d'information tel qu'il le perçoit ne correspond pas au besoin d'information réel. S'il ne maîtrise pas le vocabulaire lié au sujet ni le langage d'interrogation du SRI auquel il recourt, ses requêtes seront pauvrement formulées. Selon la nature du

besoin et le niveau d'expertise de l'utilisateur, à la fois sur le sujet et dans le domaine de la recherche d'information, l'expression en est une représentation déformée plus ou moins approximative.

La troisième dimension du modèle est le temps : durant le processus de recherche d'information, l'utilisateur apprend et son besoin d'information se modifie. Un document qui lui semblait non pertinent au début de ses recherches peut lui paraître très pertinent lorsqu'il a pris connaissance des éléments qui lui permettent d'en apprécier la teneur ou de le comprendre. Certains besoins d'information sont dynamiques, ils évoluent au fur et à mesure que l'utilisateur apprend de nouvelles choses sur le sujet. Cela se traduit souvent par des reformulations successives de la demande et de la requête. Mizzaro supprimera ensuite cette dimension du modèle parce qu'elle ne peut pas être structurée comme les autres dimensions (mais aussi, selon l'auteur, parce qu'elle complique la représentation graphique du modèle). Néanmoins, certains chercheurs étudieront cette dimension en analysant le temps non pas comme une variable continue, mais comme des étapes ou des intervalles de temps discontinus.. C'est par exemple ce qu'ont fait dans une étude exploratoire Taylor et al. [TAY 07] en modélisant le temps selon les étapes du processus de recherche d'information ISP de Carol Kuhlthau [KUH 93].

La dernière dimension proposée par Mizzaro est plus complexe et hétérogène, il la nomme les « composants » [du problème] et y regroupe

- le sujet thématique (To) qui intéresse l'utilisateur (dont une partie peut rester implicite) et sur lequel porte sa recherche,

- la tâche (To+Ta), c'est-à-dire l'activité pour laquelle l'utilisateur a besoin de cette information, et

- les attributs liés à l'utilisateur (To+Ta+UA), qui se substituent à la notion de contexte proposée dans la première version du modèle. Dans la première version du modèle, le contexte était défini de manière très floue, regroupant tout ce qui n'est lié ni au sujet ni à la tâche mais influençait la manière dont l'individu cherche et évalue les résultats. Mizzaro classait dans cette dimension par exemple les documents que l'utilisateur connaît déjà, ceux qu'il n'est pas capable de comprendre, ou encore le temps et l'argent dont il dispose pour sa recherche. Dans la deuxième version du modèle, les attributs liés à l'utilisateur comprennent aussi son degré d'expertise dans le domaine et ses préférences.

Ces éléments peuvent être partiellement ordonnés ainsi : $To \leq To+Ta \leq To+Ta+UA$ selon le nombre de variables prises en compte pour faciliter la recherche d'information en contexte. La notion de contexte dans le modèle qu'il est plutôt une non catégorie qui permet d'intégrer tous les paramètres ne relevant pas des autres et de raisonner dans l'hypothèse d'un monde fermé.

Ce modèle considère la recherche d'information comme un processus de résolution de problème où l'on peut distinguer la pertinence liée aux objets informationnels, la pertinence liée à la représentation du problème de l'utilisateur et la pertinence relative aux connaissances de et sur l'utilisateur en contexte. Il souligne aussi l'évolution du besoin informationnel dans le temps. Enfin, il s'avère très intéressant pour explorer et caractériser les différents courants de recherche sur les SRI.

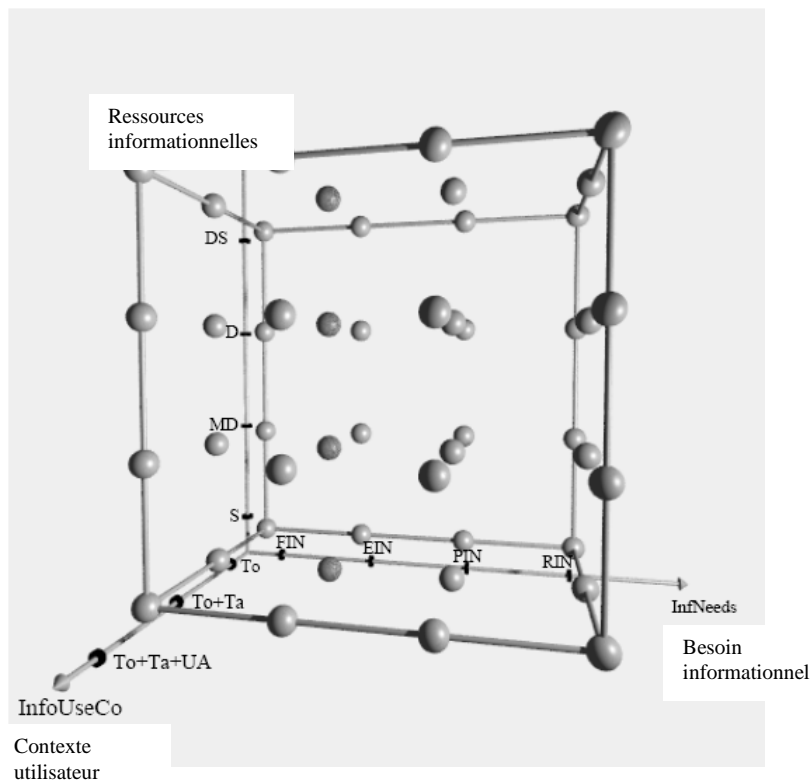


Figure 1.2 Le modèle tri-dimensionnel de Mizzaro, d'après [GAB 99]

6.3.4 Les cinq dimensions du modèle cognitif de Wang et Soergel

S'appuyant sur les travaux de recherche en psychologie traitant des processus de prise de décision, Wang et Soergel [WAN 98] interrogent la nature de la valeur attribuée aux documents par les individus en situation de recherche documentaire. À partir d'une étude qualitative longitudinale portant sur un échantillon de 25 participants (11 professeurs et 14 étudiants de second cycle), ils analysent la façon dont ces universitaires décident d'utiliser ou non les documents dans leur projet de recherche. Ils discernent cinq catégories de valeur attribuée aux documents retrouvés :

- la valeur épistémologique, qui concerne le domaine de connaissance sur lequel portent documents et demandes ; selon eux, elle conditionne toutes les autres mais n'explique pas seule le choix d'un document ;
- la valeur fonctionnelle du document pour la tâche en cours ; le document peut servir de base théorique, d'aide empirique, de guide méthodologique ou simplement permettre une bonne citation ;
- la valeur conditionnelle, c'est-à-dire le fait que le document pourrait avoir une valeur épistémologique et fonctionnelle, mais pas forcément pour la tâche en cours ;
- la valeur sociale, attribuée en fonction de la notoriété et de l'autorité de l'auteur, de la revue, ou de l'organisme pour qui l'auteur travaille dans le champ ;
- la valeur émotionnelle, c'est-à-dire la capacité du document à aiguïser les sentiments de façon positive ou négative. Par exemple, certains participants de l'étude manifestent des jugements affectifs vis-à-vis de certains auteurs. Ces sentiments les incitent ou non à exploiter leurs travaux, ou à en différer l'exploitation.

Les auteurs analysent ensuite les critères mis en œuvre par les personnes observées pour sélectionner les documents lors d'une recherche. Parmi les critères qu'ils relèvent, la pertinence-sujet (*topicality*) est la plus fréquemment évoquée dans l'explicitation des jugements de pertinence. Elle peut correspondre à une valeur épistémologique (le sujet du texte correspond au sujet traité), fonctionnelle (la méthodologie employée dans le texte correspond à ce que cherche l'individu) ou conditionnelle. Le second critère le plus cité est l'orientation (théorique/empirique) et le niveau d'expertise attendu de l'audience du texte (universitaire ou scolaire, académique ou professionnel). La discipline de rattachement de l'auteur d'un texte joue également un rôle, peut être parce qu'elle est prédictive de la capacité du lecteur à comprendre le contenu. L'article met en évidence d'autres critères sociocognitifs qui interviennent dans les jugements de pertinence : l'autorité de l'auteur, sa

réputation dans le domaine mais aussi la prise en compte des instances d'édition ou des organismes qui ont supporté la recherche.

Les auteurs distinguent donc différents types de connaissances qui influent sur les jugements de pertinence des participants : leur expertise dans le domaine du sujet, leur connaissance des personnes, des publications et des organisations du champ. Ils proposent une liste d'indices que le SRI pourrait fournir sur les documents pour aider les utilisateurs dans leur processus de choix.

L'intérêt de cette publication est d'adopter une approche originale pour analyser les jugements de pertinence des usagers : plutôt que de se focaliser sur la résolution du problème, elle décale le point de vue sur la prise de décision. Elle montre aussi l'importance des facteurs sociocognitifs dans les jugements de pertinence des utilisateurs, au moins chez un public particulier, celui des universitaires. Enfin, elle note aussi la valeur émotionnelle qui s'exprime par rapport aux documents : les individus ne fondent pas seulement leurs jugements des documents qui leur sont proposés sur des critères rationnels.

6.3.5 Synthèse

On le voit, le concept de pertinence a été approfondi à partir d'inspirations de différentes théories : Saracevic s'appuie sur la phénoménologie de Schutz et la théorie de la pertinence de Sperber et Wilson en pragmatique linguistique. Wang et Soergel s'inspirent de la théorie de la décision pour analyser la rationalité des choix des utilisateurs. Ces auteurs en tirent des propositions pour affiner la modélisation des utilisateurs de SRI. Mizzaro, quant à lui, propose un modèle qui synthétise les travaux en recherche d'information (*information retrieval*) et tente de spécifier le modèle d'un SRI générique.

Malgré toutes ces recherches, de nombreuses questions restent encore ouvertes et discutées dans le champ. Parmi celles-ci, la pertinence est-elle un phénomène totalement objectivable ou certains de ses aspects resteront-ils à jamais du domaine de la subjectivité ? Quels paramètres le SRI doit-il prendre en compte pour satisfaire au mieux l'utilisateur et dans quel but ?

Si la pertinence est le phénomène qui permet de faire sens, de comprendre et d'apprendre à partir d'informations, les SRI doivent-ils être conçus comme des systèmes d'apprentissage ? Doivent-ils chercher à réduire le plus rapidement possible l'incertitude de l'individu qui les interroge sur un sujet ou lui permettre d'approfondir son besoin, voire diagnostiquer les questions mal posées, à la manière dont certains identifient actuellement les erreurs orthographiques ? Ou s'agit-il simplement de concevoir des outils auxquels on recourt ponctuellement pour satisfaire un besoin immédiat, comme on consulte un annuaire pour trouver un numéro de téléphone ? On le voit, l'espace est ouvert à différentes familles de systèmes selon la conception que l'on se fait de l'utilisateur de l'information.

6.4. Les critères qui interviennent dans les jugements de pertinence

Analyser la façon dont se forment les jugements de pertinence des individus, a été une autre façon d'approcher le concept par les scientifiques. Ces études ont montré que les critères employés par les individus sont très nombreux et, de plus, changent durant une session de recherche. Barry et Schamber [BAR 98] ont procédé toutes deux à des études empiriques sur cette question. Tentant de comparer les résultats obtenus dans leurs études respectives, réalisées dans des contextes différents, elles trouvent des catégories de critères plutôt cohérentes. Les principales sont :

- l'étendue, le degré d'approfondissement, la spécificité de l'information ;
- sa justesse, sa validité ;
- sa clarté ;
- son degré d'actualité, le fait que l'information soit récente ;
- son caractère tangible, si l'information comporte des données ou des faits concrets ;
- la qualité ou la réputation de la source ;
- l'accessibilité, l'utilisabilité ;
- la disponibilité de l'information ou de la source d'information ;

- la vérification, le fait que l'information peut être recoupée par plusieurs documents ou sources ;
- l'émotion induite par l'information.

D'autres critères restent spécifiques à l'une ou l'autre étude, montrant bien l'effet du contexte et de la situation de recherche sur les jugements que les usagers portent sur l'information.

Spink, Greisdorf et Bateman [SPI 98] ont cherché à déterminer différentes régions de pertinence. Ils proposent un modèle à trois dimensions des jugements de pertinence : type de manifestation de pertinence selon celles déterminées par Saracevic, niveau de pertinence estimé (ils choisissent quatre niveaux : très pertinent, partiellement pertinent, partiellement non pertinent, non pertinent) et le temps. Ils montrent que les utilisateurs ont plus de mal au début de leur recherche à juger les documents pertinents, retenant alors plus de documents partiellement pertinents. Plus les usagers ont de connaissances et se focalisent sur un problème, mieux ils sont à même de prononcer des jugements tranchés. Vakkari [VAK 00A; VAK 00B] montre que les utilisateurs identifient plus de documents comme pertinents au début de la recherche et moins à la fin. Ils expliquent ce phénomène par l'acquisition de connaissance de la part des utilisateurs. Pour Greisdorf [GRE 03], les jugements de pertinence des individus sur l'information trouvée durant la recherche évoluent par niveaux, depuis la pertinence-sujet en passant par la pertinence cognitive, le sens qu'ils peuvent lui attribuer, jusqu'à l'utilité.

Knight et Burn [KNI 05] recensent les critères nécessaires à l'évaluation de l'information en ligne. S'appuyant sur d'autres études portant sur la qualité informationnelle, elles identifient de nombreux critères dont bon nombre recourent les observations de Barry et Schamber. Parmi les critères qu'elles retiennent, un plus grand nombre d'entre eux concernent l'utilité de l'information. Elles soulignent aussi l'influence du SRI sur les jugements des utilisateurs.

Rieh [RIE 02], à propos du comportement des chercheurs d'information en ligne, distingue les jugements prédictifs, faits à partir des éléments d'information proposés par le SRI pour représenter les documents dans les résultats de recherche, et les jugements évaluatifs qui portent sur le document lui-même. Elle note que les individus n'accordent pas la même importance aux différents critères à ces deux étapes, l'importance relative attribuée à chacun variant selon la situation. Les critères sur lesquels s'appuient les jugements de pertinence des individus semblent dépendre d'un certain nombre de facteurs notamment l'expérience de l'individu vis-à-vis du SRI et sa connaissance du domaine sur lequel porte l'interrogation. Selon Wang et Soergel [WAN 98], les critères tendent à devenir de plus en plus exigeants au fur et à mesure de l'avancement de la session de recherche. L'interface du système joue aussi notamment quant à l'ordre dans lequel sont présentés les résultats [SPI 98]. Enfin, même si certains documents sont jugés pertinents par un chercheur d'information, ils ne seront pas tous exploités ou utilisés réellement.

Toutes ces études enrichissent la connaissance des critères qui jouent sur les jugements de pertinence et surtout, la façon dont ils interviennent durant le processus de recherche. Elles suggèrent aussi des fonctionnalités dont les SRI devraient être dotés pour aider les usagers dans leurs quêtes informationnelles. Enfin, certaines d'entre elles engagent à faire preuve de prudence lorsqu'on établit des panels pour juger l'information dans des sessions de type Cranfield dans le but d'évaluer l'efficacité et les performances des SRI.

6.5. Communication et pertinence

6.5.1 Impact de la théorie de Sperber et Wilson

La théorie de la pertinence de Sperber et Wilson [SPE 85], développée en pragmatique linguistique, a inspiré de nombreux chercheurs qui ont tenté d'en tirer des conclusions en sciences de l'information. En considérant une recherche d'information comme un dialogue entre un individu et un système qui est susceptible de la lui fournir, cette théorie apporte en effet des éléments intéressants. Récemment, White a proposé une nouvelle interprétation de cette théorie [WHI 07a ; 07b]. Il relève que de nombreuses études caractérisent le comportement informationnel des utilisateurs comme guidé par la loi du moindre effort. Or, Sperber et Wilson attribuent un rôle important à l'effort dans la pertinence. Deux axiomes de leur théorie servent de base à ce parallèle :

- une hypothèse est pertinente dans un contexte dans la mesure où elle produit des effets cognitifs positifs dans ce contexte ;

– une hypothèse est pertinente dans un contexte dans la mesure où l'effort nécessaire pour la traiter dans ce contexte est faible.

White s'inspire de la proposition d'Andrew Goatley (dans *The language of metaphors*, 1997) qui définit la pertinence par le ratio effets cognitifs/effort de traitement. Dans le domaine de la bibliométrie, il interprète la fameuse formule $tf*idf$ comme une application de ce principe. Dans cette formule à l'origine, le paramètre tf (*term frequency*) représente la fréquence du terme cherché dans le document. Le paramètre idf (*inverse document frequency*) représente la fréquence inverse d'occurrence du terme dans la collection et donc la spécificité informationnelle du terme par rapport à cette collection. White assimile le coefficient tf à l'effet cognitif positif (plus le terme est fréquent dans le document, plus le document aura d'effets cognitifs positifs) et le coefficient idf à l'effort de traitement : plus le terme est fréquent dans une collection, plus l'effort de traitement sera grand. S'appuyant sur d'autres formules développées en bibliométrie, il montre qu'en interprétant les diagrammes sagittaux obtenus en projetant les valeurs $tf*idf$ pour classer les documents d'une collection par rapport à un terme permet d'envisager des présentations des résultats à valeur ajoutée.

Il développe aussi la notion de pertinence créative : les relations de pertinence créées par les individus ne portent pas seulement sur le fait que documents et requêtes comportent des termes communs (pertinence-sujet). Il dénombre :

- la pertinence analogique lorsque les individus mettent en œuvre des raisonnements par analogie, en faisant le parallèle entre deux concepts différents ;
- la pertinence causale, lorsqu'ils identifient une relation causale entre deux concepts, reprenant un exemple développé dans le domaine médical (telle maladie provoque telle déficience sanguine ; tel produit joue sur la même déficience sanguine donc ce produit pourrait agir sur la maladie) ;
- la pertinence méthodologique lorsque l'individu est intéressé par une méthodologie plutôt que par le sujet abordé dans le document ;
- la pertinence probante, lorsque des exemples (même si exemple n'a pas force de preuve) ou des données permettent de renforcer ou d'affaiblir une assertion ;
- la pertinence poétique, lorsqu'une métaphore ou une allusion aide à faire le point sur une question ;
- la pertinence terminologique, lorsque des écrits commentent de manière métalinguistique la manière dont le langage est utilisé dans d'autres écrits.

Il revient sur le caractère subjectif ou objectif de la pertinence, arguant que la subjectivité supposée des jugements de pertinence n'est due qu'au fait que les individus n'ont pas l'occasion d'explicitement les raisons pour lesquelles ils jugent tel document pertinent. Selon lui, il n'y aurait pas de pertinence subjective, au sens où la pertinence est toujours inférée à partir d'éléments objectivables. Le caractère subjectif de la pertinence est lié à la créativité des individus qui ne raisonnent pas forcément de manière rationnelle. Cependant, il peut exister des connaissances initiales ou des modes de raisonnement faux ou incorrects chez les utilisateurs.

De nombreuses études menées en situation réelle ont montré l'importance du critère de nouveauté de l'information pour qu'elle soit considérée comme pertinente par l'individu. La théorie de Sperber et Wilson permet d'interpréter aussi ce phénomène : un document déjà connu par un individu ne nécessite pas d'effort de traitement, mais ses « effets cognitifs positifs » sont également nuls, sauf si l'individu l'interprète d'une nouvelle manière, en le mettant en relation avec des éléments qu'il ne connaissait pas lors de sa lecture précédente.

Même si certains contestent l'appellation de théorie pour la thèse développée par Sperber et Wilson, parce qu'elle est proposée sous formes d'axiomes essentiellement appuyés sur des exemples de dialogue plus ou moins imaginaires, nul doute qu'elle a fait avancer les réflexions et la compréhension du phénomène de la pertinence.

6.5.2 Modélisations émergentes

Ottaviani [OTT 94] s'était aussi appuyé sur la théorie de la pertinence de Sperber et Wilson ainsi que sur le concept de portée significative (*significance*) chez Minski 1986 : un document est pertinent s'il fait une différence, s'il a un effet discernable sur la pensée et le comportement de l'individu. Il a comparé le processus de recherche d'information à un phénomène chaotique, au sens de la théorie du chaos de Mandelbrot, et proposé un modèle fractal qui permette de définir des clusters de documents pertinents. Sa proposition n'a pas, à notre connaissance, donné suite à de nombreux développements. Elle avait surtout l'intérêt de stipuler que, pour une

même requête, il existe dans une collection plusieurs clusters de documents pertinents,. Mais rien ne permet d'affirmer que le motif qui permet les obtenir soit pour autant régulier.

Brouard et Nie [BRU 04], dans leurs travaux sur le filtrage d'information, ont proposé un nouveau concept, celui de résonance. Ils s'appuient sur la théorie de la résonance adaptative de Grossberg (1976), issue de la cognition située (Clancey, 1997) qui a été employée pour modéliser différents processus cognitifs, de la perception à la catégorisation. Cette théorie repose sur le principe suivant : lorsqu'un système cognitif interagit avec son environnement, il capte les signaux qui en proviennent mais projette aussi sur certains d'entre eux des signaux (on peut faire l'analogie avec l'attention que l'individu porte aux objets à un moment donné) qui dépendent de ses représentations internes. Seuls les signaux qui sont aussi activés en retour par ceux des représentations internes se propagent à nouveau, en étant amplifiés ou prolongés, à la manière d'un écho. Le système entre ainsi dans un état de résonance. C'est ce processus qui permettrait la sélection de l'information. Le stimulus entré en résonance sera pris en compte pour l'apprentissage. Les autres ne seront pas pris en compte ou seront traités différemment. Il faut donc que les connaissances internes soient appliquées aux nouveaux stimuli pour que ceux-ci soient pris en compte. Et un nouveau stimulus ne sera pris en compte que s'il correspond bien à une représentation interne. Le concept de résonance semble bien rendre compte de la manière dont se produisent les interactions de pertinence. Ainsi, les neurosciences peuvent aussi participer à mieux comprendre ces mécanismes et aider à déterminer les éléments qui sous-tendent le phénomène de résonance.

Le processus de pertinence peut être assimilé dans une certaine mesure à un processus d'acquisition et de synthèse de connaissances : une information pertinente serait une information qui nous permet d'améliorer nos connaissances mais plus largement de mieux comprendre une question. A ce titre, une information est pertinente même si elle ne porte pas directement sur le sujet qui nous préoccupe, mais apporte des éléments facilitant notre raisonnement. Le concept de pertinence peut aussi être lié au phénomène de l'attention : pour que nous jugions une information pertinente, certains éléments doivent retenir notre attention. L'attention que nous portons à l'information est la première des conditions pour que nous prenions le temps d'en évaluer la pertinence.

6.6. Conclusion

Les travaux sur la pertinence ont longtemps eu pour objectif d'améliorer le fonctionnement des SRI pour les utilisateurs. Une théorie de la pertinence en science de l'information pourrait finir de réconcilier les deux champs scientifiques distincts qui travaillent sur ces questions : ceux de l'*information retrieval* qui cherchent à améliorer les algorithmes et ceux sur le comportement informationnel en sciences humaines. Si de nombreux modèles descriptifs ont été proposés jusqu'ici, une théorie pourrait aller au-delà : être non seulement explicative mais aussi prédictive de ce que l'on peut ou non attendre de l'automatisation du traitement de l'information. Elle pourrait permettre de concevoir des SRI performants et adaptés aux besoins des usagers, mais aussi de mieux structurer les recherches en sciences humaines. Le caractère subjectif ou objectif de la pertinence a longtemps tenu une place importante dans les débats entre chercheurs. Plutôt que de se poser la question en ces termes, il vaudrait mieux continuer à chercher quels en sont les éléments objectivables.

Comme le note Saracevic [SAR 07b], la plupart des recherches empiriques sur le comportement des utilisateurs de SRI se font sur des échantillons d'étudiants, qui ne sont pas forcément des juges qualifiés ou représentatifs de l'ensemble de la population des usagers. Faute de moyens financiers, les chercheurs sont souvent contraints de mener leurs études sur ces seules populations. Bien sûr, du fait de la généralisation du recours aux SRI par le grand public⁴, on ne peut pas supposer que les utilisateurs sont suffisamment « qualifiés » pour déterminer quelle est la « bonne » information parmi celles qui leurs sont proposées. Néanmoins, le besoin existe, comme en atteste le succès des moteurs de recherche commerciaux. Prospérant grâce au marketing et à la publicité, ces sociétés n'ont pas pour objectif premier de fournir la meilleure information à un moment donné mais de distribuer leur publicité de la manière la moins « dérangeante » possible pour en faciliter l'acceptation par l'internaute en la rendant « pertinente ». Les scientifiques devraient mobiliser davantage les acteurs culturels, ceux qui considèrent l'information comme un bien commun, pour faire des propositions alternatives et participer à des travaux d'envergure dans ce domaine.

⁴ Par grand public, nous voulons désigner les personnes qui n'ont pas reçu de formation conséquente à la recherche documentaire ou informationnelle.

6.7. Bibliographie

- [BAR 94] BARRY C. L., « User-Defined Relevance Criteria: An Exploratory Study », *Journal of the American Society for Information Science*, vol.45 (3), p. 149-159, 1994
- [BAR 98] BARRY C. L., SCHAMBER L., « Users' criterias for relevance evaluation : a cross study comparison », *Information Processing and Management*, vol.24 (2/3), p. 219-236, 1998.
- [BAT 89] BATES M. J., « The design of browsing and berry-picking techniques for the online search interface », *Online Review*, vol.13, p. 407-424, 1989.
- [BRO 04] BROUARD C., NIE J.-Y., « Relevance as Resonance: a new theoretical perspective and a practical utilization in information filtering », *Information Processing and Management*, vol.40, p. 1-19, 2004.
- [CHI 07] CHIARAMELLA Y., MULHEM, P., « La recherche d'information. De la documentation automatique à la recherche d'information en contexte », *Document numérique*, vol.10 (1), p. 11-38, 2007.
- [COS 00] COSIJN E., INGWERSEN P., « Dimensions of relevance », *Information Processing and Management*, vol.36 (4), p. 533-550, 2000.
- [GAB 99] GABRIELLI S., MIZZARO S., « Negotiating a multidimensinnal framework for relevance space », In Draper et al.(eds), *Proceedings of Mira 99: Evaluating Interactive Information Retrieval*, Glasgow, Scotland, UK, avril 1999.
- [GRE 95] GREEN R., « Topical Relevance Relationships. I Why topic matching fails », *Journal of the American Society for Information Science*, vol.46 (9), p. 646-653, 1995.
- [GRE 03] GREISDORF H., « Relevance thresholds: a multi-stage predictive model of how users evaluate information », *Information Processing and Management*, vol.39, p. 403-423, 2003.
- [HAR 92] HARTER S. P., « Psychological Relevance and Information Science », *Journal of the American Society for Information Science*, vol.43 (9), p. 602-615, 1992.
- [KNI 05] KNIGHT S-A., BURN J., 2005, « Developing a framework for assessing information quality on the World Wide Web », *Informing Science Journal*, vol 8, [en ligne] <http://www.inform.nu/Articles/Vol8/v8p159-172Knig.pdf>
- [KUL 93] KUHALTHAU C. C., *Seeking Meaning : A Process Approach to Library and Information Services*, Norwood, N.J.: Ablex Publishing Corp., 1993.
- [MIZ 98] MIZZARO S., « How many relevances in information retrieval ? », *Interacting with Computers*, vol.10, p. 303-320, 1998.
- [OTT 94] OTTAVIANI J., « The fractal dimension of relevance: a hypothesis », *Journal of the American Society for Information Science*, vol.45 (4), p. 263-272, 1994.
- [RIE 02] RIEH S., « Judgement of Information Quality and Cognitive Authority in the Web », *Journal of the American Society for Information Science and Technology*, vol.53 (2), p. 145-161, 2002.
- [SAR 96] SARACEVIC T., « Relevance reconsidered - Information Science: Integration in perspective », In *Proceedings of the second conference on Conception of Library and Information Science*, Copenhagen, Danemark, p. 201-218, 1996.
- [SAR 97] SARACEVIC T., « The stratified model of information retrieval interaction: Extension and Applications », In *Proceedings of the American Society for Information Science meeting*, Vol. 34, p. 313-327, 1997.
- [SAR 07a] SARACEVIC T., « Relevance: a review of the literature and a framework for thinking on the notion in Information Science. Part II:Nature and manifestations of relevance », *Journal of the American Society for Information Science and Technology*, vol.58 (13), p. 1915-1933, 2007.
- [SAR 07b] SARACEVIC T., « Relevance: a review of the literature and a framework for thinking on the notion in Information Science. Part III: Behavior and Effects of Relevance », *Journal of the American Society for Information Science and Technology*, vol.58 (13), p. 2126-2144, 2007.
- [SPE 85] SPERBER D., WILSON D., *La Pertinence : communication et cognition*. (Trad. de l'anglais par A. Gerschenfeld et D. Sperber), Paris: Editions de Minuit, 1985/1989.
- [SPI 98] SPINK A., GREISDORF H., BATEMAN J., « From highly relevant to not relevant: examining different regions of relevance », *Information Processing & Management*, vol.34 (5), p. 599-621, 1998.
- [TAY 07] TAYLOR A. R., COOL C., BELKIN N. J., AMADIO W. J., « Relationships between categories of relevance criteria and stage in task completion », *Information Processing and Management*, vol. 43, p. 1071-1084, 2007.
- [WAN 98] WANG P., SOERGEL D., « A Cognitive Model of Document Use during a research project. Study I. Document Selection », *Journal of the American Society for Information Science*, vol.49 (2), p. 115-153, 1998.
- [WHI 07A] WHITE H. D., « Combining Bibliometrics, Information Retrieval, and Relevance Theory, Part 1: First examples of a Synthesis », *Journal of the American Society for Information Science and Technology*, vol.58 (4), p. 536-559, 2007.

[WHI 07B] WHITE H. D., « Combining Bibliometrics, Information Retrieval, and Relevance Theory, Part 2: Some implications for Information Science », *Journal of the American Society for Information Science and Technology*, vol.58 (4), p. 583-605, 2007.

[XU 06] XU Y., CHEN Z., « Relevance Judgment: What do Information Users Consider beyond topicality ? », *Journal of the American Society for Information Science and Technology*, vol.57 (7), p. 961-973, 2006.