



La Recherche floue

Gabriel Gallezot, Michel Roland, Jacques Araszkieviev

► To cite this version:

Gabriel Gallezot, Michel Roland, Jacques Araszkieviev. La Recherche floue. Document numérique et Société, Paris, CNAM, 17-18 novembre 2008, Nov 2008, France. pp.411-429. sic__00340835

HAL Id: sic__00340835

https://archivesic.ccsd.cnrs.fr/sic__00340835v1

Submitted on 22 Nov 2008

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

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

Version auteur du texte publié p.411-430 de

Traitements et pratiques documentaires : vers un changement de paradigme ?

Actes de la deuxième conférence Document numérique et Société, Paris, CNAM, 17-18 novembre 2008

Sous la direction d'Evelyne Broudoux et Ghislaine Chartron

Les éditions de l'Adbs

Collection : Sciences et techniques de l'information

456 page(s),

ISBN 978-2-84365-116-8

La Recherche floue

- GALLEZOT, Gabriel, Université de Nice Sophia Antipolis, URFIST, I3M, gallezot@unice.fr
- ROLAND, Michel, Université de Nice Sophia Antipolis, URFIST, roland@unice.fr
- ARASZKIEWIEZ, Jacques, Université de Nice Sophia Antipolis, IUT Nice Côte d'Azur, Département Information-Communication, I3M, araszkie@iutsoph.unice.fr

Résumé: Les enquêtes sur les pratiques informationnelles en milieu universitaire tendent à reposer sur le modèle de la fracture numérique générationnelle, où le milieu universitaire est impacté de l'extérieur par les usages des *digital natives*, et à évaluer les pratiques des étudiants selon leur écart avec ce que les médiateurs estiment les "bonnes pratiques". Un mode d'évaluation non normatif mais basé sur l'efficacité des stratégies de recherche permettrait de rendre mieux compte de leur diversité voire de leur inventivité. Une telle évaluation suppose une nouvelle compréhension de ces stratégies. A travers trois axes de questionnement (connaissance des ressources, appropriation critique des interfaces, réassomption de la subjectivité des pratiques) nous esquissons un nouveau paradigme d'une "recherche floue" qui associe la maîtrise des outils selon le critère de la pertinence à la sérendipité en tant qu'elle suppose l'aptitude à accueillir l'inattendu.

Abstract: The many surveys about information usages in academic environment tend to rest on the model of the generational digital divide, according to which the academic world is impacted from outside by the "digital natives" usages. As a consequence, they tend to evaluate students information practices according to the distance from the "good practices" as conceived by the mediators. A non-normative evaluation mode, based on efficiency, could give a better view on the diversity and inventivity of the strategies. Such an evaluation implies a new understanding of those information retrieval. Through three lines of questioning (knowledge of resources, critical appropriation of interfaces, reassumption of subjectivity) we give a sketch for a new paradigm, that of a "fuzzy search" wich associates mastering the tools according to the relevance criteria with serendipity as implying the trained aptitude to greet the unexpected.

Mots-clés: relevance, pertinence, sérendipité, recherche floue, pratiques informationnelles

Keywords: searching, digital natives, serendipity, fuzzy search, information literacy

Contexte / Introduction

Si l'on a pu croire, dans un premier temps, que le domaine universitaire profiterait de la "révolution numérique" sans que soient remis en cause ses équilibres fondamentaux, force est de constater que la situation est plus problématique. Dans les outils de la recherche scientifique sont peu à peu importés des techniques et des pratiques originaires du web 2.0. Le caractère spontané et désordonné de cette importation entraîne une situation de crise où les acteurs sont désorientés par la déconstruction/reconstruction de leur univers documentaire. Ils ont tendance à objectiver indistinctement n'importe quel résultat et les formateurs à l'IST ont l'impression de mener un combat perdu d'avance contre Google.

Cependant de récents articles américains (HEAD, 2007 ; GUESS, 2008) ont pointé que l'utilisation majoritaire de Google par les étudiants n'était pas exclusive de l'utilisation d'autres outils. S'il existe ainsi d'autres *patterns* d'organisation des procédures de recherche, encore faut-il les recenser et en évaluer l'importance et l'effectivité (MARCUS 2007). Une étude sur les pratiques informationnelles des étudiants avancés et des enseignants chercheurs des pôles universitaires de la région Paca, pilotée par l'Urfist-Paca,

est programmée pour 2008-2009. Elle a pour ambition de collecter des données permettant de reconnaître, de décrire et d'évaluer les procédures élaborées dans ce nouveau contexte, procédures que nous regroupons sous l'intitulé de "recherche floue"¹, qui désigne à la fois le changement de repères ou flou sémiotique (ARASZKIEWIEZ, à paraître), une modification de l'intelligibilité ou flou cognitif et les éventuelles remédiations à un effet de "flouté" des résultats.

Nous souhaitons, par le biais d'une observation participante au sein de stages réalisés à l'Urfist et le défrichage de la littérature sur le sujet, proposer, en amont de l'enquête, des pistes d'observation, des éléments à quantifier ou à vérifier, des faits à expliciter. Nous considérerons ainsi des pratiques informationnelles comprises dans une palette très large de situations, entre une vision amoindrie de la réalité documentaire et une recherche aléatoire contrôlée, en y associant la notion et la fonction heuristique de la sérendipité qui nous permettra d'esquisser un nouveau paradigme. Et au-delà nous tenterons d'appréhender le cadre conceptuel nécessaire à la compréhension des pratiques observées comme de leur effectivité.

Notre propos sera structuré autour de trois questionnements:

1. Les universitaires connaissent-ils les ressources disponibles en ligne (sur leur campus ou sur le web), et maîtrisent-ils les procédures d'interrogation des sources?
2. La fausse évidence des interfaces d'interrogation des ressources en ligne induit-elle une perte des repères cognitifs généralement convoqués dans l'activité de recherche d'information?
3. La réintroduction d'une subjectivité (d'un flou contrôlé) par des pratiques informationnelles avancées permet-elle une véritable intelligibilité des résultats de la recherche d'information?

I- Connaissance et maîtrise des sources

Il s'agit ici de mettre en avant le flou des modèles cognitifs auxquels ont recours les publics concernés (étudiants avancés et enseignants chercheurs). Les universitaires connaissent-ils les sources disponibles en ligne (sur leur campus ou sur le web) et maîtrisent-ils les procédures d'interrogation des sources ?

Observation participante

Dans le cadre de la formation à l'IST au sein de l'Urfist Paca-c ou dans le cadre d'interventions externes (cursus universitaire ou autres lieux de formation), nous sommes en contact avec des étudiants avancés, des enseignants chercheurs et des bibliothécaires-documentalistes des pôles universitaires de la région. Nous avons donc pu observer lors des formations dispensées à ces publics quels étaient leurs comportements et interrogations. Outre le caractère empirique de la méthode, nous devons souligner certains biais : le stagiaire présent à nos formations est en phase d'apprentissage de manière volontaire ou non volontaire. Il y aura une différence entre celui qui sait qu'il ne sait pas et souhaite se former et celui qui reçoit une formation sans l'avoir souhaité. Cette évidence conduit à distinguer deux situations. Dans la première, les lacunes sont identifiées, les attentes sont importantes, mais le stagiaire connaît déjà le cadre et le contexte. Dans la seconde, se trouvent souvent des stagiaires qui croient connaître l'essentiel nécessaire à leur activité. De même, de par sa composition professionnelle et estudiantine et de par les disciplines représentées, la population observée est loin d'être homogène. Malgré ces biais, nous pouvons relever quelques traits saillants des pratiques informationnelles des populations avec lesquelles nous interagissons.

¹ L'expression « recherche floue » est généralement employée pour désigner un « mode de recherche d'information tolérante vis-à-vis d'approximations ou d'erreurs (fautes d'orthographe par exemple) sur les termes de la question et/ou des documents, utilisant des techniques statistiques, phonétiques ou de reconnaissance de forme » (cf vocabulaire de la Doc, Site de l'ADBS, . Cette expression est la traduction de *fuzzy search* (ou *fuzzy matching*). Elle fait référence à la *fuzzy logic* (la logique floue) où, à la différence de la logique booléenne qui ne connaît d'autres états que vrai ou faux, la « recherche floue » traite les états avec une précision relative. Dans ce texte nous élargissons donc cette définition de « recherche floue »

Connaissance des ressources

Les ressources de la bibliothèque sont peu ou mal connues et en conséquence peu utilisées. Ce constat serait à valider formellement par un examen des statistiques collectées par les services communs de documentation (SCD). Deux éléments récurrents peuvent néanmoins apparaître d'ores et déjà comme des indices de sa justesse : la méconnaissance des listes des ressources sur le site du SCD et la satisfaction lors de découverte d'une base de données du domaine étudié. Plus généralement, la distinction entre une base bibliographique, une base de périodiques, un catalogue, etc. n'est pas connue. Cette méconnaissance entraîne nécessairement des insatisfactions dues aux écarts entre les attentes et les résultats obtenus. De manière générale, il convient de noter que les ressources en bibliothèque ne sont pas les premières à être consultées. Les outils de recherche du web sont devenus prédominants. Pourtant là aussi peu d'outils sont connus et utilisés, et là encore la typologie des outils et leur champ d'application est méconnue.

Soulignons en outre que les facteurs « temps », « efficacité », « simplicité » sont des arguments régulièrement avancés pour ne pas se confronter à un outil. Qu'il s'agisse d'une ressource du SCD ou du web, il est souvent réclamé un apprentissage très prosaïque, proche d'une aide en ligne pourtant peu explicite et rarement lue. Les aspects historiques, méthodologiques, conceptuels sont souvent perçus comme inutiles. L'environnement documentaire est ainsi refusé : quelques mots entrés dans la fenêtre de recherche de Google suffisent à produire des réponses « pertinentes ».

Maîtrise des outils

Le maître-mot serait donc l'efficacité. Sous cet angle, le moteur de recherche semble parfaitement convenir. Son usage est considéré comme facile au regard des ressources de la bibliothèque : « Je tape, j'obtiens quelque chose qui me convient. Les bases de données ? Je tape... et je n'obtiens pas forcément quelque chose ». En effet, pour bien utiliser les ressources de la bibliothèque, il convient de connaître, ou du moins d'appréhender *a minima* les « codes documentaires » : la terminologie, la typologie, la structuration, l'indexation et les standards des ressources et des documents. Cet apprentissage supplémentaire est perçu comme obsolète au regard de la simplicité d'utilisation des moteurs de recherche. Parallèlement, peu d'intérêt est manifesté concernant la manière, au demeurant complexe, dont les moteurs de recherche délivrent leurs résultats. Les problèmes d'un Neil Montcrief, ou de la société American Blinds² (BATTELLE, 2006) et *a fortiori* la censure consentie par Google sur son moteur de recherches en Chine sont des faits connus. Ils auraient dû attirer l'attention sur la qualité des résultats fournis. Paradoxalement, du fait de leur caractère exceptionnel, ils ont été perçus comme autant de preuves de la qualité de ces résultats en temps normal, c'est-à-dire hors des pressions politiques et économiques.

La tactique gagnante

« Il savait qu'il disposait par ailleurs d'un outil plus simple pour faire ses repérages. Il ne pensait donc pas qu'il lui était utile d'acquérir une meilleure compréhension des modes d'interrogation du catalogue, ni une démarche de recherche experte. Il avait une stratégie qu'il considérait comme efficace et qui était construite dans l'interaction entre une phase exploratoire sur Internet et une phase de confirmation à la bibliothèque. » (DESPRES-LONNET & COURTECUISSÉ, 2006)

Cet extrait montre les habiletés développées par les universitaires pour contourner la difficulté d'usage des outils de la bibliothèque tout en les conservant comme sources d'information validée. Peu savent d'ailleurs que la plupart des ressources de la bibliothèque sont consultables de chez eux par le biais de l'ENT. Plus généralement, pour nombre de stagiaires, toutes les ressources et outils présentant dans leurs interfaces un champ dans lequel on peut insérer des caractères sont des moteurs de recherche. On perçoit alors que non seulement les notions qui gravitent autour de la représentation d'un corpus ne sont pas acquises mais encore que les utilisateurs croient en des tactiques gagnantes. Du mode « recherche flou »,

on passe alors au mode « résultats floutés ». Ce constat est évidemment d'autant plus accablant qu'il s'accompagne de l'autosatisfaction aveuglante liée à une découverte par soi-même effectuée.

Résultats d'enquêtes

Googlisation ?

Un certain nombre d'enquêtes en milieu universitaire (en 2003 aux Etats-Unis (MARCUM & GEORGE, 2003) et au Québec (MITTERMEYER & QUIRION, 2003), en 2005 sur deux universités parisiennes (MARESCA et al. 2005), en 2006 sur une université lilloise (DESPRES-LONNET & COURTECUISSE, 2006), en 2008 en Belgique francophone (THIRION, 2008) et sur les universités de Bretagne (HENRIET et al., 2008)...) semble confirmer les résultats de notre observation participante : utilisation réflexe des outils hégémoniques du web, méconnaissance de tout ou partie des ressources documentaires et de leur logique d'interrogation, appropriation accrue de l'informatique.

Cette constatation se laisse facilement thématiser selon un modèle qui correspond à une appréhension commune et spontanée des effets de la « révolution numérique », celui de la « fracture numérique générationnelle » qui opposerait les *digital natives*³, nés avec le numérique ou peu avant, spontanément « alphabétisés » dans l'univers numérique et dans l'internet, aux générations plus anciennes et aux compétences et procédures de ces dernières notamment en matière de recherche et de validation de l'information⁴. Caractéristique des *digital natives* serait une appropriation spontanée, prédiscursive (préconsciente) des outils, ie des interfaces du web. Cette fracture est appréciée de manière ambivalente : les *digital natives* s'ils sont à l'aise avec le monde numérique, manqueraient des outils de compréhension, d'analyse et de critique d'une information appréhendée sans distance et passivement.

Ainsi la plupart des enquêtes citées ci-dessus confrontent les populations étudiantes (en situation de *digital natives*) avec les « bonnes pratiques » de la recherche documentaire, telles que fixées, pour l'essentiel, dans la situation documentaire précédant l'irruption du web (usage de la bibliothèque et de son OPAC, interrogation des bases de données), « bonnes pratiques » envisagées si nécessaire dans l'environnement du web mais indemnes des usages spécifiques engendrés par celui-ci. Et avec plus ou moins de nuances, les constats confirment la fracture générationnelle : usage massif et non-critique des outils du web les plus populaires (Google et Wikipedia), délaissement voire ignorance des outils documentaires classiques.

Beyond Google ? Les bonnes pratiques du St Mary's College

Une enquête présentée dans un article de *First Monday* (HEAD, 2007) et portant sur les pratiques informationnelles des étudiants de St Mary's College, en Californie, semble contredire ces constats. Elle fait état de pratiques plus conformes à ce que les médiateurs tendent à considérer comme les « bonnes pratiques » : recherches en bibliothèque, outils spécialisés, recours aux médiateurs...⁵.

On peut relativiser ces surprenants résultats car ils sont issus d'une université catholique américaine dans le domaine des SHS et correspondent donc à une stratégie pédagogique délibérée. Ils suffisent cependant à indiquer que la situation n'est pas homogène et un retour, par-delà les synthèses, sur les résultats des enquêtes nous montre que l'apparente homogénéité des constats globaux cache des pratiques diverses et fait apparaître une contradiction moindre entre les constatations de St Mary's College et ces résultats.

3 ou *Google generation*: (cf. Jason Frand, The information mindset: Changes in students and implications for higher education, EDUCAUSE Review, March / April 2006, p.15.)

4 Voir en français: P. Lardellier, *Le pouce et la souris : Enquête sur la culture numérique des ados* (Fayard, 2006).

5 *We found students: (1) accessed convenient, vetted, and aggregated online resources from course readings and the campus library Web site, (2) to a lesser extent, used Internet sites, such as Yahoo!, Google, and Wikipedia, and (3) worked with professors or librarians one-on-one to narrow down searches and clarify expectations for assignments.*

Prenons par exemple, parmi les enquêtes citées, la plus récente d'entre elles, celle qui porte sur les universités de Bretagne (HENRIET et al., 2008). Si le constat global correspond bien à l'impression générale, certains résultats particuliers viennent cependant nuancer le tableau. S'ils ne contestent pas la pertinence du constat global, ils en limitent la portée au point de mettre en question la valeur explicative du modèle qui le sous-tend.

Nous relèverons quatre points de discussion :

- Si 96% des doctorants bretons utilisent prioritairement pour leurs recherches les moteurs de recherche et si parmi ceux-ci, c'est à 85% de *Google* qu'il s'agit, ils sont cependant 37% à utiliser *Google Scholar*. Or pour porter l'étiquette « Google », *Google Scholar* est un outil très différent du moteur généraliste. Il est spécifiquement destiné au public universitaire et chercheur et offre accès à une documentation généralement pertinente pour la recherche⁶. Mais les 37% d'utilisation d'un outil spécialisé viennent nuancer le constat massif des 96% d'utilisation d'outils généralistes. D'autant plus que *Google Scholar*, qui souffre de la méfiance des médiateurs à l'endroit de la maison *Google* et de la concurrence qu'il vient faire aux outils mis en place par ceux-ci (interfaces « documentation électronique » des SCD en particulier), n'est pas particulièrement promu par ces médiateurs.
- « 94% des doctorants n'utilisent jamais ou rarement les blogs ». C'est-à-dire que la dynamique « web 2.0 » semble ne pas toucher les doctorants, du moins dans le cadre de leurs activités académiques. Or les caractéristiques essentielles de ce qu'on appelle « web 2.0 » (*user generated content*, participation, horizontalité, démediation...) *métonymisent* une part essentielle des nouvelles pratiques prêtées aux *digital natives*⁷.
- L'enquête constate une importante demande de formation à la recherche, l'analyse et l'exploitation de l'information (plus de 60%), constatation qui vient en contradiction avec celles d'enquêtes antérieures qui indiquaient un manque d'intérêt pour la formation documentaire⁸. Ainsi, pour les doctorants bretons au moins, *Google* n'est pas une évidence auto-suffisante.
- Enfin l'enquête reflète une forte demande (plus de 60% encore) de sélection de ressources et d'*intermédiation*.

Ainsi il apparaît que si les déficits de compétences informationnelles des doctorants bretons sont réels, ils semblent conscients de ce manque. La situation n'est donc pas celle d'une confrontation binaire entre deux « littératies », l'ancienne et la nouvelle. Tout en apparaissant en moindre contradiction avec le cadre général que les résultats de l'enquête californiens, l'enquête bretonne impose une conclusion somme toute identique⁹.

Une enquête « flash »

Nous avons réalisé en préparation de cet article une enquête « flash » par questionnaire auprès des chercheurs du laboratoire I3M de l'université de Nice Sophia-Antipolis¹⁰. La faible population sondée et les conditions de l'enquête interdisent toute généralisation. Elle n'en fait pas moins apparaître, chez des chercheurs trop âgés pour être classés comme *digital natives*, une grande diversité de stratégies de recherche où l'on pourra cependant reconnaître les caractéristiques des stratégies attribuées aux étudiants *digital natives* (prépondérance de Google en particulier) mais plus ou moins modulées, parfois par des

6 En particulier *Scholar* va servir de mode d'accès à la documentation électronique acquise par l'université d'appartenance. Le prix fort de ce type d'acquisition pose d'autres problèmes qui sortent du cadre du présent article.

7 Un rapport de PEW Internet (Lenhart, A., Madden, M., Macgill, A. R. & Smith, A. *Teens and social media*. URL http://www.pewinternet.org/pdfs/PIP\Teens_Social_Media_Final.pdf.) montre que dans leur vie sociale, les jeunes générations, américaines au moins, participent largement à cette tendance:

8 La contradiction s'explique sans doute en partie par la différence des publics cibles : on peut penser que les doctorants sont plus sensibles à leurs besoins en compétences informationnelles que des étudiants de licence, pour des raisons de sélection, d'expérience et de nature d'activité académique;

9 A majority of students do not understand what quality research resources are and how to locate them. As a result, students seek a balanced approach to course-related research, leveraging both online and offline resources.

10 Recherche floue/Enquête flash (I3M)

éléments inattendus (importance du livre acheté en librairie), et donc sans qu'il soit possible d'en rendre compte par l'adéquation à un modèle unique.

Les limites du modèle « fracture générationnelle numérique »

Ces observations nous amènent à remettre en cause la valeur explicative et descriptive du modèle de la « fracture numérique générationnelle ». Une étude réalisée par *University College London* pour la *British Library* et le JISC, parue en janvier 2008 (UCL, 2008), si elle souligne bien les carences en compétences informationnelles des jeunes générations, remet en cause la notion de *digital natives* en montrant que les traits comportementaux censés caractériser ces derniers tendent à se répandre aujourd'hui dans tous les groupes d'âges.

Par ailleurs, si la fracture générationnelle n'est pas absolue, les pratiques ne sont pas homogènes au sein de la *Google generation* des *digital natives*. Un article récent de Danah Boyd (BOYD, 2007) montre ainsi que le « choix » d'un réseau social est sociologiquement déterminé. La synthèse de l'enquête sur la lecture et les loisirs multimédias des adolescents réalisée en 2007 pour le CNL et la direction du livre (ITHAQUE, 2007) montre également une forte détermination sociologique des pratiques de l'internet au sein de la « génération numérique ». Si les adolescents qui lisent le moins de livres (25%) sont aussi ceux qui passent le plus de temps sur internet, leurs parents n'utilisent jamais un ordinateur. A l'inverse les parents des adolescents qui lisent le plus de livres (8%) utilisent souvent un ordinateur et régulent l'usage de l'internet et des jeux vidéos de leurs enfants¹¹. Les caractéristiques sociologiques discriminantes mises en lumière par ces travaux déterminent vraisemblablement des différences analogues dans l'appropriation des outils du web dans le cadre universitaire¹².

Recenser les stratégies

Au-delà d'un modèle, certes en grande partie pertinent mais qui tend à réduire unilatéralement la réalité, notamment en occultant les déterminations socio-culturelles, il est nécessaire de reconnaître, d'ordonner et d'évaluer des stratégies diverses. Ce recensement ne peut se contenter d'être descriptif : il y a en perspective de nos investigations des visées pratiques, en particulier l'adaptation des stratégies de formation à l'IST aux pratiques, attentes et besoins des acteurs universitaires. L'évaluation est donc nécessaire.

La plupart des enquêtes citées ci-dessus confrontent les pratiques observées à ce que les médiateurs (enseignants, formateurs, bibliothécaires) considèrent comme les "bonnes pratiques". En somme elles évaluent en mesurant l'écart à une norme. Comme le public visé est principalement celui des étudiants, cet écart peut facilement être interprété en termes générationnels. C'est supposer que les pratiques informationnelles des chercheurs actifs correspondent bien à cette norme et de façon plus générale à un "état de l'art" relativement stable. Tout nous indique au contraire (voir notre enquête "flash" ou l'étude de UCL (UCL, 2008)) que ces pratiques sont en pleine évolution.

Plutôt que sur une norme implicite, nous proposons de baser l'évaluation sur la finalité des pratiques informationnelles, en d'autres termes d'évaluer les stratégies de recherche d'information en se posant la question de leur efficacité au regard des finalités académiques elles-mêmes. Une telle évaluation suppose l'analyse des résultats en termes de bruit et de silence et ne saurait être réalisée sur des échelles comparables à celles des enquêtes normatives. Elle est cependant possible sur des échantillons représentatifs d'une typologie. Cette approche de l'évaluation permettrait de viser dans la même observation apprenants avancés et chercheurs, d'examiner autrement qu'en termes d'écart les rapports

11 c'est-à-dire qu'ici la corrélation négative entre lecture et pratique de l'internet correspond à une transmission de capital culturel et d'une première expérience du numérique accumulés par la génération des parents.

12 voir aussi : Granjon, F. & Lelong, B. Capital social, stratifications et technologies de l'information et de la communication: Une revue des travaux français et anglo-saxons. Réseaux 139, 147-181 (2006). URL http://www.cairn.info/resume.php?ID_ARTICLE=RES_139_0147.

entre les compétences académiques et les pratiques extra-universitaires et peut-être ainsi de proposer des pistes pour les enseignants, pour les formateurs à l'IST et pour les bibliothécaires.

Le critère d'efficacité cependant ne peut suffire à articuler une typologie. Comme nous venons de le noter, une typologie, même imparfaite, est plutôt une condition pratique d'une telle évaluation. Pour ordonner les observations sur une typologie, il est nécessaire de proposer des pistes théoriques pour la compréhension des mécanismes cognitifs mis en oeuvre dans les pratiques informationnelles.

II - Quels ancrages cognitifs ?

Flou et flouté

Qu'il s'agisse d'une adaptation de la technique aux usages en vigueur ou d'un apprentissage aux méthodologies documentaires, l'universitaire se trouve dans un espace certainement flou mais qui peut de surcroît être flouté par la méconnaissance des outils, des techniques et des méthodes, méconnaissance qui provoque une troncature ou une surcharge informationnelle. Si « flou » peut renvoyer à l'imprécision des connaissances d'un individu (acquises ou données) qui permettent d'ajuster son regard, il peut également conduire à une aperception de la recherche comme bricolage selon l'expression de Claude Lévi-Strauss. Par contre « flouté » ne peut renvoyer qu'à l'ensemble des indices sémiotiques de l'interface qui viennent, avec la complicité plus ou moins active de l'internaute, brouiller l'intelligibilité de la source ou la mise en application des connaissances. Ainsi en est-il de l'énonciation d'un moteur de recherche comme Google qui vient cacher la complexité des opérations du *Page Ranking* et simultanément le caractère nécessairement arbitraire des résultats fournis. Dans les deux cas c'est l'ancrage cognitif qui est questionné.

L'ancrage cognitif

« des documentalistes à la fois praticiens et formateurs font des choses qu'ils n'enseignent pas et enseignent des choses qu'ils ne font pas »
« Le travail d'élaboration d'un référentiel de compétences documentaires dit ce qu'il faudrait faire plutôt que ce qui se fait. »
« Ils (enseignants en documentation) ne considèrent ni les connaissances acquises dans le cadre de l'enseignement de la documentation en lycée et en collège, ni les compétences documentaires acquises lors de l'utilisation de moteurs comme Google. Pour eux, utiliser une base de données [en droit] relèverait d'une compétence spécifique.
« Ils ne savent pas comment fonctionne cette base, inconnue d'eux. Mais ils savent l'utiliser. Ces mêmes étudiants n'ont pas non plus conscience qu'il existe des annuaires, des portails et des moteurs de recherche, même s'ils savent utiliser ces trois outils... sans véritablement les différencier. » (TRICOT, 2007)

Ce constat plutôt sévère pour les formateurs pose frontalement la question des ancrages cognitifs nécessaires aux individus plutôt que celle d'un référentiel stabilisé qui ne prendra jamais suffisamment en compte le tropisme des utilisateurs d'outils documentaires.

En terme de recherche d'information, on peut distinguer avec (TRICOT, 2006) quatre registres de compétences documentaires:

- les compétences en lecture ;
- les compétences en compréhension ;
- les compétences instrumentales ;
- les compétences informationnelles .

Derrière chacun de ces registres un ensemble de concepts, notions, et pratiques doit être articulé dans un curriculum à la « maîtrise de l'information » pour que chaque étudiant/stagiaire dispose des compétences instrumentales et informationnelles nécessaires à la recherche d'information. Sans proposer un curriculum, d'ailleurs en phase de recherche/conception (cf. travaux Erté¹³), nous souhaitons pointer une notion : la *relevance*.

Cette notion permet d'appréhender de manière complexe et non triviale le lien entre le comportement de l'utilisateur, la mesure des systèmes d'information et la terminologie d'usage associée à la recherche d'information. Elle couvre en même temps la performance des systèmes et la pertinence intrinsèque du document pour le lecteur. Elle renvoie aux notions d'indexation, de taxonomie, d'ontologie, de pertinence, de *serendipity*, de *ranking*, de bruit, de silence, de corpus... autant de notions clés qui permettent non seulement de comprendre les outils documentaires et du web mais surtout de posséder des repères cognitifs capables d'engendrer une appropriation instrumentale et informationnelle.

Le schéma ci-dessous établi par S. Mizzaro et plus encore l'ensemble de son article « *Relevance: The whole history* » (MIZZARO, 1997) propose un état des lieux très précis de cette notion qui continue d'être un sujet important en science de l'information, travaillé par d'autres disciplines comme les mathématiques, la logique, l'informatique, la psychologie, les sciences cognitives...

« [...] each relevance can be seen as a point in a four-dimensional space, the values of each of the four dimensions being: (i) *Surrogate, document, information*; (ii) *query, request, information need, problem*; (iii) *topic, context, and each combination of them*; and (iv) *the various time instants from the arising of problem until its.* » (MIZZARO, 1997)

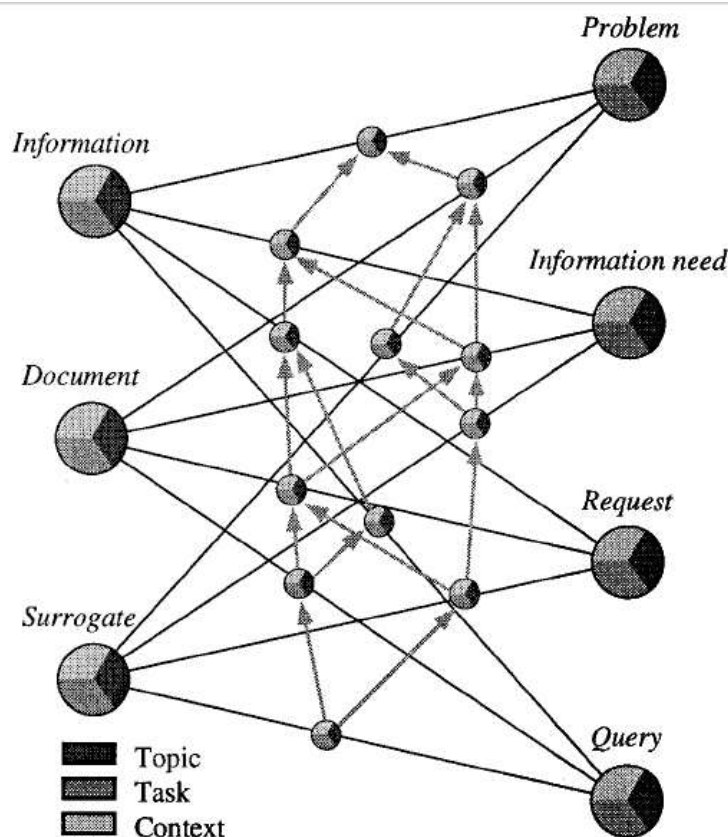


FIG. 1. The partial order of relevances.

Embrasser le concept de *relevance*, c'est certainement posséder les ancrages cognitifs pour appréhender un univers documentaire en perpétuel mouvement (et donc flou). Il s'agit de prendre appui sur ceux-ci pour faire évoluer de manière individuelle les pratiques documentaires dans le temps et pour provoquer des intégrations conceptuelles (*blending*) (TURNER, 2000) et ainsi construire les schèmes cognitifs

engagés dans la recherche d'information. C'est également, dans le même mouvement, y déterminer la place d'une subjectivité.

III - Redocumentarisation et introduction d'une subjectivité

La subjectivité introduit un flou au sens où chaque individu (se) projette son expérience. Même en situation de maîtrise instrumentale et informationnelle, il convient d'accepter cette part de flou qu'il convient sinon d'apprivoiser, du moins d'intégrer, voire de provoquer. Il s'agit ici, loin de l'effet de *flouté* produit par l'homogénéité des résultats d'une recherche d'information tronquée, de distinguer des éléments capables de faire émerger des structures mentales. Concernant ce flou contrôlé il faut alors à considérer deux axes :

- celui des pratiques informationnelles avancées (permettent-elles une véritable intelligibilité des résultats de la recherche d'information?)
- celui des capacités représentationnelles

Redocumentarisation

« Redocumentariser c'est documentariser à nouveau un document ou une collection en permettant à un bénéficiaire de réarticuler les contenus sémiotiques selon son interprétation et ses usages à la fois selon la dimension interne (extraction de morceaux musicaux pour les ré-agencer avec d'autres, ou annotations en marge d'un livre suggérant des parcours de lecture différents...) ou externe (organisation d'une collection, d'une archive, d'un catalogue privé croisant les ressources de différents éditeurs selon une nouvelle logique d'association) » (PEDAUQUE, 2006), (ZACKLAD, 2007)

L'utilisation avancée des ressources et des outils qui permettent de redocumentariser (en particulier ceux qui autorisent l'annotation, le *mashup*, la tertiarisation, la percolation, le tri subjectif...) génère une appropriation documentaire étendue. Du sens émerge alors à partir de liaisons inattendues. Ils permettent une représentation personnalisée de corpus là où les outils traditionnels de recherche d'information uniformisent.

Ces composés sont érigés, *via* des artefacts, par et pour un individu dans un « espace privé ». En passant dans l'espace public, ces unités documentaires vont plus ou moins se fixer dans le temps par « hasard-sélection » selon un processus de « darwinisme documentaire » (PEDAUQUE, 2006). Le document perçu *a posteriori* ne sera intelligible que par les traces et le cadre de référence des éléments qui le composent. Loin d'objectiver n'importe quel ensemble de résultats d'outils de recherche, il s'agit de « jouer » avec les possibles.

Schèmes cognitifs et sérendipité

Ce qui est à l'oeuvre ici, ce sont les représentations mentales des individus dont l'analyse mettrait en perspective les variances des schèmes cognitifs et ainsi peut-être éclairerait les éléments générateurs des pratiques informationnelles.

Les situations de *flouté* peuvent être corrigées (mais non résolues) par la réintroduction d'un questionnement sur les résultats obtenus. Il s'agit de permettre au sujet, à l'utilisateur, de « reprendre la main » :

- par compréhension du principe de fonctionnement des outils, les usagers n'accordent plus aveuglément le même crédit aux résultats;
- par des artefacts techniques que l'utilisateur, les usagers appréhendent alors autrement les résultats.

Les schèmes cognitifs changent, on ne « joue » plus de la même manière avec les outils de recherche. Là

où l'usager fonctionnait sur le mode de la « consommation » en tenant pour équivalentes les notions de relevance, de pertinence et de sérendipité ($R = P = S$), il peut fonctionner sur le mode de la « création » où la sérendipité est l'écart entre relevance et pertinence ($S = P - R$, d'où $P = R + S$).

La sérendipité est définie ainsi : « *découverte par chance ou sagacité de résultats que l'on ne cherchait pas* ». Au plus proche du conte original et du terme « sagacité », il convient de relever que cela ne peut arriver qu'aux esprits bien préparés. De même, avec Sylvie Catellin (CATELLIN, 2001), nous pouvons souligner une analogie avec l'abduction où une activité sémio-cognitive est liée à des dispositifs indiciaires. Il s'agit de donner un sens aux faits observés. La notion de sérendipité, au-delà de son acception couramment retenue (« trouver par hasard »), peut être comprise comme un ensemble de capacités à détecter des indices et à générer des heuristiques pertinents à la recherche en cours (FOSTER & Ford 2003).

Le fait pour l'usager d'utiliser des mots clés approximatifs, de ne pas connaître les techniques documentaires et de mal connaître l'étendue du champ qu'il interroge, lui donne d'autant plus l'impression d'obtenir des réponses pertinentes ; la sérendipité est donc l'équivalent pour lui de la magie du moteur de recherches. C'est en ce sens, ou plutôt avec ce détournement de sens, que Google l'utilise (bouton « j'ai de la chance »). Des modèles de représentations montrent que la « correspondance » simple n'est jamais en phase avec le déroulement d'une recherche d'information, ainsi le modèle documentaire de Rouet (ROUET, 2004) met en avant les représentations des sources et la variation des contenus, ou encore le modèle d'intégration conceptuelle (ou *blending*) de Turner et Fauconnier (TURNER, 2000) – décrit *une opération cognitive non séquentielle* qui permet de générer une structure mentale nouvelle.

Des bifurcations cognitives aidées ou non d'artefacts techniques sont toujours nécessaires, le principe de la sérendipité, c'est d'être préparé à ces sauts mentaux.

Conclusion

Observation participante, préenquête, parcours de la littérature et examen du contexte conceptuel nous amènent à poser en amont de notre enquête l'hypothèse que les stratégies de recherche les plus efficaces, c'est-à-dire celles dont les résultats sont non seulement les plus pertinents mais les plus féconds, sont celles qui dépassent l'opposition des deux fausses évidences, celle d'une objectivation des résultats des interfaces faciles et puissantes de type « Google-Wikipedia » comme celle de la stabilité des « bonnes pratiques », retrouvant ainsi des processus (subjectivité, sérendipité) oblitérés par ces évidences.

Ce dépassement suppose de nouveaux ancrages cognitifs et une subjectivation assumée de ses pratiques informationnelles. Ces nouveaux ancrages cognitifs ne se résument pas à l'apprentissage de procédures et à la connaissance des ressources mais supposent une compréhension de l'écologie informationnelle dans son évolution et on peut s'attendre à les retrouver dans des configurations multiples, reflétant des stratégies et des priorités spécifiques.

Ce cadre d'hypothèses sera mis à l'épreuve, précisé et affiné par l'enquête que nous mettons en œuvre, au moment de l'élaboration du protocole d'observation puis lors de l'analyse des résultats.

Références

- (ARASZKIEWIEZ, à paraître) cf. Jacques Araszkieviev, Le grand avaleur, Google : une propédeutique à l'usage des dispositifs, A paraître.
- (BATTELLE, 2006) BATTELLE, J., -La révolution Google, Eyrolles, 2006, pp. 133-163
- (BOYD, 2007) Boyd, D. Viewing american class divisions through Facebook and Myspace. Apophenia Blog Essay (2007). URL <http://www.danah.org/papers/essays/ClassDivisions.html>.

- (CATELLIN, 2001) - CATELLIN, S. « Sérendipité, abduction et recherche sur Internet. » In Emergences et continuité dans les recherches en information et communication, Actes du XIIe Congrès national des SIC, UNESCO (Paris), 10-13 janvier 2001. Paris : SFSIC, 2001
- (DESPRES-LONNET & COURTECUISSÉ, 2006) DESPRES-LONNET, M.; COURTECUISSÉ, J.-F., « Les étudiants et la documentation électronique », BBF, 2006, n° 2, p. 33-41, [en ligne] <<http://bbf.enssib.fr>> Consulté le 9 juillet 2008
- (FOSTER & Ford 2003) - FOSTER A. E. et FORD N. – “Serendipity and information seeking: An empirical study”. Journal of Documentation, 2003; 59: 321-340
- (GUESS, 2008) – GUESS A.- « Research Methods ‘Beyond Google’ ». Inside Higher ED, june 17, 2008- URL: <http://www.insidehighered.com/news/2008/06/17/institute>, dernier accès septembre 2008
- (HEAD, 2007) – HEAD, Alison J., “Beyond Google : How do students conduct academic research?” in First Monday, volume 12, number 8 (August 2007), URL: http://firstmonday.org/issues/issue12_8/head/index.html, dernier accès juillet 2008
- (HENRIET et al., 2008) - HENRIET, O MALINGRE, M.-L. SERRES, A. Enquête sur les besoins de formation des doctorants à la maîtrise de l’information scientifique dans les écoles doctorales de Bretagne (2008).URL: http://www.uhb.fr/urfist/files/Synthese_Enquete_SCD-URFIST.pdf
- (ITHAQUE, 2007) - Synthèse de l’enquête sur la lecture et les loisirs multimédia des collégien(ne)s et lycéen(ne)s enquête « centre national du livre / direction du livre et de la lecture » réalisée par ithaque. (2007). URL <http://www.centrenationaldulivre.fr/?Synthese-de-l-enquete-sur-la>
- (MARCUM & GEORGE, 2003) – MARCUM D. B., GEORGE, G. - D-Lib Magazine ISSN 1082-9873 9 (2003). URL: <http://www.dlib.org/dlib/october03/george/10george.html>
- (MARCUS, 2007) - MARCUS, C., et al. , Understanding research behaviors, information resources, and service needs of scientists and graduate students: A study by the university of minnesota libraries (2007) URL: <http://lib.umn.edu/about/scieval/>
- (MARESCA et al. 2005) - MARESCA, B., DUPUY, C., CAZENAVE, A. Enquête sur les pratiques documentaires des étudiants, chercheurs et enseignants-chercheurs de l’université pierre et marie curie (paris 6) et de l’université Denis Diderot (Paris 7), R238 (2005). URL: <http://www.credoc.fr/publications/abstract.php?ref=R238>
- (MITTERMEYER & QUIRION, 2003) D MITTERMEYER D., QUIRION, D. Étude sur les connaissances en recherche documentaire des étudiants entrant au 1er cycle dans les universités québécoises, CREPUQ (2003). URL: <http://crepuq.qc.ca/documents/bibl/formation/etude.pdf>
- (MIZZARO, 1997) MIZZARO, S. « Relevance: The whole history ». Journal of the American Society for Information Science 48, 810-832(1997)
- (PEDAUQUE, 2006) - Pédaque, R., T. Le document à la lumière du numérique, C&F Editions: 2006
- (ROUET, 2004) - ROUET, J.-F. – « Traitement et représentation mentale des documents complexes, bilan et projet de recherche », Oct 2004, [En ligne] http://www.mshs.univ-poitiers.fr/laco/Pages_perso/Rouet/Textes/rapport04-JF-Rouet.pdf
- (THIRION, 2008) - THIRION, P. et al. Enquête sur les compétences documentaires et informationnelles des étudiants qui accèdent à l’enseignement supérieur en communauté française de belgique: Rapport de synthèse. Tech. Rep., Conseil Interuniversitaire de la Communauté française de Belgique (2008). URL: <http://www.edudoc.be/synthese.pdf>
- (TRICOT, 2006) - TRICOT A. L’utilisation d’Internet permet-elle de nouveaux apprentissages documentaires ?, <http://savoirscdi.cndp.fr/pedago/Reflexion/tricot/tricot.htm>
- (TRICOT, 2007) - TRICOT A., « Des pratiques informationnelles aux savoirs documentaires chez les élèves du secondaire : l’exemple de la recherche d’information » http://fadben.free.fr/synth%E8se_politiquedocumentaire_2007.doc
- (TURNER, 2000) - Mark TURNER, « blending », 4 interventions au Collège de France, Juin 2000 <http://markturner.org/cdf.html>
- (UCL, 2008) - UCL Information behaviour of the researcher of the future. British Library, JISC (2008). URL

<http://www.bl.uk/news/pdf/googlegen.pdf>

- (ZACKLAD, 2007) - ZACKLAD, M. – « Réseaux et communautés d'imaginaire documédiatisées », in Skare, R., Lund, W. L., Varheim, A., A Document (Re)turn, Peter Lang, Frankfurt am Main, 2007, pp. 279-297