



HAL
open science

Didactique de l'informatique et recherche d'information sur le web : quelle (s) perspective (s) ?

Béatrice Drot-Delange

► To cite this version:

Béatrice Drot-Delange. Didactique de l'informatique et recherche d'information sur le web : quelle (s) perspective (s) ?. Sciences et technologies de l'information et de la communication en milieu éducatif : Analyse de pratiques et enjeux didactiques., Oct 2011, Patras, Grèce. pp.51-63. edutice-00676168

HAL Id: edutice-00676168

<https://edutice.hal.science/edutice-00676168v1>

Submitted on 3 Mar 2012

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.

Didactique de l'informatique et recherche d'information sur le web : quelle (s) perspective (s) ?

Béatrice Drot-Delange

beatrice.drot-delange@univ-bpclermont.fr

Université Blaise Pascal, Clermont-Ferrand, France

Résumé. Le champ de la réflexion de la didactique informatique peut s'étendre aux outils de recherche d'information sur le Web. Il semble pourtant que la recherche en didactique de l'informatique n'ait que peu progressé dans ce sens. Notre communication interroge la nature des liens entre informatique et culture informationnelle dans un premier temps, puis nous présenterons les résultats d'une expérimentation auprès d'étudiants de licence information-communication sur leur compréhension du fonctionnement d'un moteur de recherche. Nous avons sollicité le cadre théorique de Norman pour tester dans quelle mesure l'image du moteur de recherche permet ou non à ces utilisateurs d'avoir un modèle mental correct de son fonctionnement. Il s'avère que tant l'interface que l'aide en ligne ne suffisent pas aux utilisateurs pour déterminer les traitements opérés par celui-ci sur leurs requêtes.

Mots-clés : modèle mental, moteur de recherche, culture informationnelle, aide en ligne

Introduction

Le précédent colloque Didapro (Paris) souhaitait étendre le champ de la réflexion de la didactique informatique à des systèmes tels que les « outils de communication et d'accès aux documents et aux savoirs » (Baron & Pochon, 2009, p 14). Ce champ inclut les outils de recherche d'information sur le web. Il semble pourtant que la recherche en didactique de l'informatique n'ait que peu progressée dans ce sens. A lire les recommandations internationales, les textes réglementaires en France, les discours des chercheurs dans le domaine de la formation à la recherche d'information sur le web, on constate qu'informatique et culture informationnelle sont liées. Notre communication interroge la nature de ces liens dans un premier temps, puis nous présenterons les résultats d'une expérimentation auprès d'étudiants de licence information-communication sur leurs compréhensions du fonctionnement d'un moteur de recherche. Nous discuterons enfin la question de savoir si la recherche d'information sur le web peut être un objet pour la didactique de l'informatique.

Informatique et maîtrise de l'information

L'UNESCO dans son programme Information pour tous (PIPT) définit la maîtrise de l'information comme un droit humain de base dans un monde numérique.

« Les personnes qui maîtrisent l'information sont en mesure d'accéder à des données relatives à leur santé, leur environnement, leur éducation et leur travail, ce qui leur permet de prendre des décisions cruciales concernant leur propre existence [...] »

Cette maîtrise de l'information lie les compétences informationnelles et informatiques. Elle « suppose que les intéressés possèdent les compétences requises pour utiliser les TIC et leurs applications afin d'accéder à l'information et de créer de l'information. Ainsi, pour pouvoir naviguer dans le cyberspace et se servir des documents multimédias disponibles par lien hypertexte, il faut posséder les compétences techniques nécessaires pour utiliser l'Internet et celles requises pour interpréter l'information ». Ce texte sépare nettement les compétences manipulatoires de celles liées à l'information. Il s'agit de savoir utiliser l'outil technique et de savoir interpréter l'information. Autrement dit, l'un n'est pas pensé avec l'autre.

Dans l'« Introduction à la maîtrise de l'information », l'UNESCO (2007) définit la famille des « maîtrises pour la survie » au XXI^e siècle comprenant (1) la maîtrise fonctionnelle fondamentale ou de base des pratiques (compétences) que sont la lecture, l'écriture, l'expression orale et le calcul ; (2) la maîtrise de l'informatique ; (3) la maîtrise des médias ; (4) l'enseignement à distance et le cyberapprentissage ; (5) la maîtrise culturelle ; (6) la maîtrise de l'information.

La maîtrise de l'informatique y est définie comme « l'aptitude à bien savoir comment utiliser et exploiter les ordinateurs en tant que machines de traitement de l'information. Il s'agit d'une moitié de la maîtrise des TIC, l'autre moitié étant la maîtrise des médias ». Elle inclut la maîtrise du matériel, la maîtrise du logiciel et la maîtrise des applications, c'est-à-dire des logiciels d'application spécifiques. La maîtrise du logiciel « renvoie à l'ensemble « invisible » de procédures et d'instructions d'application générale dont l'ordinateur ou le matériel de télécommunications a besoin pour exécuter convenablement ses fonctions ». Dans ce texte, l'approche de l'informatique reste centrée sur l'utilisation des machines et des logiciels. Mais elle inclut aussi une dimension plus conceptuelle, en se fixant pour objectif des connaissances sur la part « invisible » du fonctionnement des machines.

En France, la « maîtrise des techniques usuelles de l'information et de la communication » est l'un des sept piliers du socle commun de connaissances et de compétences au niveau du collège, défini par le bulletin officiel de l'éducation nationale du 20 juillet 2006. « La culture numérique implique l'usage sûr et critique des techniques de la société de l'information. Il s'agit de l'informatique, du

multimédia et de l'internet, qui désormais irriguent tous les domaines économiques et sociaux ». Elle fait également l'objet d'un domaine du B2i (Brevet informatique et Internet), « s'informer, se documenter », qui vise entre autres capacités, celles de l'utilisation principale ou avancée d'un moteur de recherche sur le web. Enfin, on trouve également une certification au niveau des universités avec les C2i (Certificat Informatique et Internet). On peut noter, à l'instar de Baron (2009), que la centration est faite « sur l'appropriation des compétences en faisant l'impasse sur les savoirs en jeu ». Cette critique des formations focalisées sur les savoir-faire et les procédures est également présente dans (Serres, Le Deuff, 2009). Tout comme elle est soulignée par Simonnot (2009a), « des formations plus systématiques et plus complètes à l'information sont nécessaires à chaque niveau d'étude scolaire et universitaire. Il est évidemment essentiel de les différencier des formations strictement informatiques et manipulatoires qui ont marqué les débuts de l'intégration du numérique dans les programmes de formation pour en faire percevoir le sens et l'utilité » (p. 57).

Les textes internationaux et les textes règlementaires français semblent donc bien entériner l'idée d'un lien fort entre informatique et maîtrise de l'information, même si cette liaison semble par trop réduite à ses dimensions utilitaires plutôt que conceptuelles. Ce lien est mentionné par des chercheurs du champ des sciences de l'information et de la documentation.

Ainsi, Baltz (1998), initiateur de l'expression « culture informationnelle », soulignait le fait que les intellectuels commençant à se préoccuper de la société de l'information avaient par trop tendance à la réduire à ses éléments techniques. La culture informatique est nécessaire, mais ne peut pas être limitée à un « seul contact technologique », qui n'apporte pas de compréhension profonde de cette société de l'information. Selon lui, l'informationnel devait passer avant l'informatique. L'informatique n'étant qu'une réponse à des problèmes informationnels avant tout.

Pour certains, les compétences en informatique sont nécessaires à cette culture informationnelle. Ainsi Juanals (2003) écrit que « les compétences permettant de formuler une recherche d'information dans un dispositif d'hypermédia en réseau nécessitent l'acquisition préalable de notions appartenant à plusieurs champs incluant l'informatique, les environnements hypermédiés et la documentation » (p.177)

Serres et Thiault (2010) insistent sur cette culture technique. « Ainsi, pour prendre le seul domaine de la recherche d'information, nous n'avons cessé, avec d'autres, de plaider pour mettre au premier plan, dans la formation documentaire et informatique des élèves et des étudiants, la compréhension d'internet et des outils de recherche, la « déconstruction » de Google par exemple [Serres, Le Deuff, 2009], la maîtrise des notions de la recherche d'information mais aussi de l'informatique, maîtrise conceptuelle largement absente aujourd'hui chez les

« digital natives » (p. 59). Le rapport 2009 du Groupe de Recherche sur la Culture et la Didactique de l'Information (GRCDI) élabore un curriculum infodocumentaire qui propose entre autre « l'acquisition d'une véritable culture technique, permettant notamment de comprendre le sens et le fonctionnement des objets techniques et informationnels » (p. 87). Une autre proposition passe par la distinction d'objets didactisés empruntés au monde réel des réseaux tels que les moteurs de recherche, etc. La didactisation de ces objets doit permettre « aux élèves de dépasser le stage de l'intuition ou de l'usage irréflecti pour aller vers la capacité à comprendre les principes de fonctionnement des objets techniques et leur logique interne » (p. 92).

Dénonçant, dans la démarche de recherche d'information, l'accent parfois mis exclusivement sur la rapidité et l'obtention d'un résultat, Simonnot (2009b) pense que « [la culture informationnelle] doit amener à questionner la manière dont sont obtenus ces résultats et les procédés que cela induit, par exemple, quand, au nom de la sécurité, on garde des traces des activités informationnelles des individus » (p. 35). Elle conclue en insistant sur la nécessaire « prise de conscience quant aux dispositifs complexes qui sont mis en place pour collecter, traiter et diffuser l'information à travers les réseaux » (p. 36).

Il ressort de ces analyses que les chercheurs invoquent la nécessité d'avoir une maîtrise conceptuelle de l'informatique, de comprendre le sens et le fonctionnement des objets informatiques. Il ne s'agit pas en effet de mieux faire mais de mieux comprendre pour garder le contrôle sur des artefacts cognitifs utilisés quasi-quotidiennement.

Une expérimentation concernant la compréhension du fonctionnement du moteur de recherche

Une didactique des moteurs de recherche passe par la distinction de deux types d'apprentissage celui de la manipulation de surface et celui d'une compréhension profonde du système. Selon Norman (1987), les utilisateurs, en interagissant avec un système cible (t), développent un modèle mental de ce système (M(t)). Il leur permet, même s'il n'est pas complet, d'anticiper les résultats d'une action. Les concepteurs, mais aussi les enseignants, développent un modèle conceptuel du système cible (C(t)). Idéalement, les modèles mental et conceptuel coïncident pour une utilisation optimale. La seule manière qu'ont les concepteurs du système de communiquer avec les utilisateurs est le système lui-même. Norman (1986) définit la notion d'image du système. Le système nous donne ou devrait nous donner à voir toutes les informations nécessaires à son utilisation. L'image du système comprend l'artefact lui-même mais également les systèmes d'aides, les manuels, les messages d'erreur, etc.

Après avoir rappelé la méthodologie et les principaux résultats obtenus lors de notre expérimentation (Drot-Delange, 2011), nous analyserons sur quels éléments

s'appuient les utilisateurs réguliers d'un moteur de recherche pour se prononcer sur certaines de ses fonctionnalités. Il s'agit donc de tester dans quelle mesure l'image du moteur de recherche permet à ces utilisateurs d'avoir un modèle mental correct de son fonctionnement. Hendry et Efthimiadis (2008) ont déjà montré que même des étudiants avancés en science de l'information utilisent peu de concepts pour décrire le fonctionnement d'un moteur de recherche. De plus, ces concepts sont souvent les plus évidents : requête et résultats.

Les fonctionnalités retenues sont les suivantes : l'ordre des mots et la troncature. Ces fonctionnalités ont été sélectionnées parce qu'elles n'appellent pas de réponse évidente ni immédiate et qu'elles impliquent d'avoir une représentation des fonctions essentielles d'un système de recherche d'information. Un tel système se compose de deux fonctions principales (Ihadjadene & Fondin, 2004) :

- l'indexation qui vise à fournir une représentation du contenu des documents, c'est-à-dire l'extraction et le stockage de la sémantique du contenu des documents. Le résultat est une liste de termes significatifs, généralement pondérés, pour mesurer au mieux leur représentativité.
- l'appariement document-requête, c'est-à-dire la comparaison entre requête et document qui évalue par un score la pertinence du document par rapport à la requête. Cette fonction d'appariement est liée à la manière dont les documents et les requêtes sont indexés et à la pondération des termes.

Ainsi le critère « ordre des mots » fait-il intervenir les représentations en matière d'indexation et d'appariement. Comment le moteur indexe-t-il les pages ? Quels sont les éléments de la page qui sont concernés par cette indexation (titre, description, lignes de textes, etc.) ? Comment l'algorithme détermine-t-il qu'un terme est significatif ? Comment le moteur procède-t-il pour déterminer que tel document est « pertinent » par rapport à telle requête ?

La troncature intervient dans la manière dont le moteur indexe documents et requêtes et éventuellement dans l'interface de recherche. La troncature ou *stemming* est la méthode qui consiste à supprimer des suites de lettres à la fin ou au début des mots avant de les indexer ou de les rechercher. Par exemple, les formes *déménageurs*, *déménagement*, *déménagera*, *déménager* seront ramenés à la pseudo-racine *déménag* (De Loupy & Creastan, 2004). Il n'y a généralement pas de contrôle du lien entre les termes considérés. La troncature peut donc intervenir tant dans l'utilisation via l'interface que dans le fonctionnement du moteur. Quelle est la visibilité pour l'utilisateur de ce traitement effectué par le moteur sur sa requête ?

Le protocole

L'expérimentation a été menée auprès d'étudiants de première année de licence en information et communication dans le cours intitulé « Initiation à Internet ». Nous avons demandé aux étudiants, seuls ou en binôme, de se prononcer sur le

fonctionnement du moteur de recherche Google concernant les deux fonctionnalités présentées plus haut. Il s'agissait pour eux de concevoir les investigations qu'ils jugeaient nécessaires pour être à même de statuer. Les consignes données à l'oral et par écrit incitaient les étudiants à chercher soit dans la documentation de l'outil soit dans une autre documentation et à valider ou non le jugement lié à l'expérimentation. L'image du système, permettant d'élaborer un modèle mental du système, est en effet constituée des informations présentes à l'écran lors des interactions de l'utilisateur avec le moteur mais aussi de la documentation en ligne de l'outil. Les étudiants devaient produire un document écrit réalisé lors des séances, véritable journal de bord, de leurs expériences: investigations menées, documentations consultées le cas échéant, réponse apportée, commentaires (difficultés rencontrées par exemple). Ces réponses ont ensuite fait l'objet d'une analyse de contenu.

Résultats

Nous rappelons succinctement les résultats obtenus lors de cette expérimentation pour les deux fonctionnalités.

- Ordre des mots

Globalement, les étudiants affirment majoritairement (environ les deux tiers) que l'ordre des mots a une importance. Cependant l'analyse des explications fournies montre un certain embarras pour trancher. Ainsi, certains groupes évoquent « *le léger changement dans l'ordre des résultats* », le « *léger chamboulement dans l'ordre des 10 premiers résultats* », « *les résultats sont sensiblement similaires* », « *le moteur ne tient pas tellement compte de l'ordre des mots* ». On constate donc une relative incertitude que la plupart des groupes n'explique pas.

Les méthodes d'investigation des étudiants relèvent de deux approches. La première consiste à mélanger des mots sans se soucier de savoir si cela a un impact sur le sens de la requête : *communication média consommation, beurre noisette salé, climatique réchauffement, etc.* La seconde consiste à trouver des exemples *a priori* pour lequel le changement d'ordre modifie le sens de la requête : *Moyen Age/âge moyen, Haïti séisme/séisme Haïti, etc.* Pour quelques étudiants, Google interprète leur requête et agit en conséquence (souligné par nous) :

« *Google fait la différence : « fleur rouge jardin » et « jardin fleur rouge », le résultat est différent* ».

« *En tapant un nom propre désignant une région, malgré l'inversement des termes, Google les remet dans l'ordre* ».

La documentation du moteur précise, à la date de l'expérimentation, que « *Google accorde (...) une grande importance à l'ordre dans lequel vous saisissez les termes et au fait qu'ils soient regroupés. Nous ne nous transgressons jamais cette règle sans raison valable* ». Les étudiants se sont également appuyés sur la page qui

précise que « *Google signale uniquement les pages qui comportent TOUS ces termes, mais pas nécessairement à la suite des uns des autres* ».

- La troncature

Concernant la troncature, environ les deux tiers des groupes considèrent que Google l'autorise, un tiers qu'il ne l'autorise pas. Aucun n'a consulté (ou trouvé l'information) dans la documentation en ligne de Google. Celle-ci explique la troncature (<http://www.google.com/intl/fr/help/basics.html>) dans le paragraphe sur la lemmatisation. « *Exemple : Si vous entrez le terme « cheval » ou « cheva* », Google ne fait pas porter la recherche sur « chevaux », « chevaline » ou « chevalerie », mais uniquement sur le terme « cheval » ou « cheva* » (soit la chaîne de caractères « cheva » suivie d'un astérisque)* ». En fait, le caractère * n'est pas un opérateur de troncature sur Google mais un caractère joker qui remplace un ou plusieurs mots.

Peu d'étudiants se sont documentés pour connaître la disponibilité ou non de la troncature sur Google et l'opérateur lui correspondant. Trois quarts des groupes ont testé la troncature sans employer d'opérateur. Ils ont donc en fait testé les « suggestions » de Google qui apparaissent dans la barre de recherche au fur et à mesure de la saisie des mots clés. Les explications données concernent la capacité d'anticipation de Google et son intelligence (souligné par nous :

« *Il prend la racine « biblio » pour la transformer en bibliothèque, moteur de recherche **intelligent qui trouve** la suite de la racine* ».

« *Google **comprend** immédiatement les abréviations. Google est **efficace** dans le sens où il **interprète** notre recherche grâce à un répertoire d'abréviations. Son interprétation est donc assez ouverte* ».

« *On tape l'abréviation com et Google a **directement compris** le sens de notre recherche et nous dirige sur des sites de communication* ».

Discussion

L'aide en ligne, partie intégrante de l'image du système qu'est un moteur de recherche, n'a été que peu consultée par les étudiants. Ce n'est en effet pas un comportement naturel pour des utilisateurs réguliers et qui estiment que ce moteur est convivial, c'est-à-dire facile et plaisant à utiliser comme l'a montré l'enquête menée par Simonnot (2009a). Malgré cela, ceux qui ont utilisé l'aide de l'outil nous ont fait part de la difficulté à naviguer et à trouver l'information recherchée. Ils ont observé des comportements estimés « aléatoires », des liens d'une page pouvant aboutir à une documentation en anglais alors que dans un autre contexte le même intitulé de lien avait abouti à une page en français. Chang et al. (2008) ont passé en revue les aides en ligne d'une vingtaine de moteur de recherche et concluent à leur insuffisance et notamment leur incapacité à s'adapter à l'utilisateur. La documentation est surtout construite sur des instructions du type

comment faire. Elle ne fournit pas de justifications à ces règles. Ce faisant, elle ne permet pas à l'utilisateur de comprendre davantage le fonctionnement du moteur.

L'expérimentation sur l'ordre des mots menée par les étudiants ne leur permet pas de conclure de manière définitive. Ils formulent des requêtes et se prononcent en fonction du feed-back du moteur qui se traduit par l'affichage d'une page de résultats. Ainsi, les critères retenus par les étudiants sont que, suivant l'ordre des mots, les réponses obtenues sont les mêmes ou non, l'ordre des résultats est identique ou non, ou bien encore le nombre de résultats affiché varie ou non. Notons que ce contrôle visuel est sommaire car ne portant que sur la première page de résultats. Les opérations effectuées par le moteur ne sont pas explicitées et les utilisateurs n'en ont pas connaissance.

L'incertitude des étudiants est encore plus grande avec la troncature. Là encore, ils se fondent sur la lecture de la page de résultats, prenant pour preuve du traitement de la troncature le fait que des mots entiers se substituent à leur requête. Pourtant, cette fonctionnalité n'est pas disponible pour l'utilisateur comme l'indique l'aide en ligne de google.fr déjà citée. Mais le moteur de recherche, lui, utilise la troncature de manière automatique (voir par exemple les billets du blog officiel de Google, « *Our international approach to search* » du 21 novembre 2008 ou « *helping computers understand language* » du 19 janvier 2010 sur <http://googleblog.blogspot.com/>). La mise en gras des mots qui ne sont pas forcément ceux de la requête est expliquée par Google comme le produit de cette troncature ou de la lemmatisation.

Reconnaissons une limite sérieuse pour une didactique de l'informatique dans le cas des moteurs de recherche : les traitements opérés par les moteurs sont cachés et l'image donnée par ce système ne permet pas d'en suivre les opérations.

Jeanneret (2000) soulignait déjà que « *l'utilisateur ne voit plus une partie des opérations dont procède ce qui lui est rendu visible, qu'il abandonne à la machine, et à ses concepteurs, une partie des activités qu'il pratiquait lui-même dans la recherche documentaire. (...) C'est particulièrement vrai dans le cas des moteurs de recherche, puisque les principes techniques sur lesquels repose leur fonctionnement sont des secrets de fabrication, auxquels même les chercheurs n'ont pas accès* » (p. 49). On peut s'en réjouir dans le sens où cela rendrait plus démocratique l'accès à ce type d'outil. Simonnot (2008) écrit ainsi que « *la prise en charge par les algorithmes de l'interprétation des requêtes et de leur transformation en équations de recherche a élargi l'usage des moteurs à tous ceux qui ne connaissaient pas les langages documentaires et leurs opérateurs* ». Mais on peut aussi considérer que cela nécessite justement un apprentissage pour permettre l'autonomie des utilisateurs : « *la recherche d'information avec un moteur peut être perçue par l'utilisateur comme automatisée alors qu'il s'agit d'un filtrage de l'information très sophistiquée. Ce filtrage, souterrainement imposé, sans cesse changeant, soulève la question de l'autonomie de l'utilisateur* »

(Ravenstein & Ladage, 2006, p. 372). C'est pourquoi ces auteurs concluent à la nécessité de faire du référencement un objet d'enseignement, pour ne pas laisser les seuls experts de la sphère marchande, détenteurs de ces savoirs.

On constate que les traitements effectués par le moteur ne sont guère rendus visibles via l'interface de l'outil. En cela on peut dire que l'image renvoyée à l'utilisateur par le système, tant par son interface que par son aide en ligne, n'est pas suffisamment explicite pour que celui-ci puisse élaborer un modèle mental qui coïncide avec le modèle conceptuel.

Koenemann et Belkin (1996) proposent une typologie des niveaux d'interactivité :

- opaque (boîte noire) : le système ne fournit aucune information sur le traitement opéré sur la requête de l'utilisateur ;
- transparent (boîte blanche) : le système informe l'utilisateur sur les transformations opérées ;
- pénétrable : en plus de la transparence ci-dessus, le système fournit les moyens de contrôler ou d'ajuster les transformations.

Même si l'interface de Google évolue très rapidement et régulièrement, il ne permet pas de saisir toutes les modifications opérées sur la requête de l'utilisateur par le moteur lui-même de manière automatique (Muramatsu & Pratt, 2001). Ces modifications ont pourtant une influence sur la sélection des pages, la présentation des résultats et l'évaluation par les utilisateurs de leur pertinence.

L'interface actuelle de Google, en mars 2011 Google Jazz, en donnant à voir explicitement quelques-unes des options de filtrage des résultats pourrait s'inscrire dans cette volonté de rendre le système « pénétrable ». Ainsi en est-il de la géolocalisation des résultats : elle apparaît clairement dans la colonne de gauche sur google.fr. On peut si on le souhaite modifier la localisation de la requête, mais on ne peut pas l'empêcher, contrairement à ce que permet l'interface de google.com. Les étudiants interrogés sur ce point n'en perçoivent pas la portée sur l'obtention des résultats. En somme, il ne suffit pas de rendre accessible certains paramètres pour que ceux-ci soient compris et utilisés par l'utilisateur. De Souza (2005) montre que l'interface d'un système renseigne l'utilisateur sur les fonctionnalités ou les possibilités de celui-ci à un moment donné. Mais, par exemple, l'interface ne renseigne pas les utilisateurs sur le raisonnement qui a conduit à rendre inopérant tel menu dans telles circonstances. Le fait que certains soient en grisé indique que, dans le contexte de la tâche de l'utilisateur, en fonction des variables d'état du logiciel, ils ne sont pas applicables. Mais cela ne dit rien sur le pourquoi et encore moins quoi faire pour qu'ils le soient.

Les perspectives ouvertes par la recherche d'information en contexte et l'adaptation à l'utilisateur (Chiaromella & Mulhem, 2007) laissent entrevoir de plus en plus d'automatisme dans le traitement et surtout dans la modification de la

requête. Ainsi Marchionini (2008) indique que les « interactions homme-information », qu'elle propose en lieu et à place d'interactions homme-machine, sont déjà intégrées au fonctionnement du web. Les classements des résultats des moteurs de recherche, les recommandations sur les sites de e-commerce sont variables en fonction de la situation et du moment de la consultation ou de la recherche de l'information. Ces modifications peuvent être instantanées et basées sur des conditions externes éventuellement historisées. Ainsi les filtres de Google Jazz sont proposés de manière contextuelle : si Google détecte dans la requête de l'internaute un lien avec un évènement, il mettra en avant les filtres Actualités, Blogs ou Vidéos par exemple. Mais la navigation des autres influe également sur les informations présentées par le système. L'UNESCO (2007) envisageait même dans le futur « *des systèmes capables d'établir automatiquement de nouveaux liens entre pages de la Toile, ou d'en supprimer sur la base du comportement des usagers : l'Internet ainsi conçu serait capable de se mettre à jour automatiquement en intégrant les choix des acteurs* » (p 61).

Marchionini (2008) explique que les comportements conditionnels des objets web permettent aux designers d'impliquer les utilisateurs dans l'annotation, l'édition, le classement et l'affichage personnalisé. Il nous semble qu'en matière de didactique de l'informatique, une question intéressante serait celle de la gestion personnelle de l'information ainsi recueillie, annotée, etc. Par exemple, le constat a été fait en France d'une très faible utilisation de logiciels permettant par exemple la gestion bibliographique par les doctorants (Albert, 2009 ; Henriët, Malingre & Serres, 2008). Quelques travaux se sont intéressés par exemple à l'usage du copier/coller comme moyen de conserver, de structurer cette information, mais cette phase du processus de la recherche d'information reste très peu étudiée (Boubée & Tricot, 2010, p. 66). Le constat de Duchâteau en 2002 semble en effet encore d'actualité : « *on pouvait espérer que des usages plus soutenus d'Internet allaient réveiller l'intérêt pour le reste des instruments logiciels, ceux qui permettent le traitement des informations recueillies ou mises en ligne* ».

Conclusions

Les moteurs sont un moyen quasi-obligé d'accès à l'information sur le web. Connaître leur logique de fonctionnement devient alors fondamental tant elle touche des domaines divers de la vie quotidienne des utilisateurs soit comme producteurs d'informations sur le web, incluant les traces laissées par toute navigation, posant avec insistance la question de l'identité numérique, soit comme chercheurs d'informations. Si la recherche d'information sur le web est une activité solitaire, elle prend néanmoins place dans un collectif dont le comportement devient une variable de l'interaction « individuelle ».

Il s'agit bien ici d'intégrer la notion de sens dans le comportement des machines qui prennent de plus en plus en compte les actions voire les intentions des

utilisateurs, les caractéristiques des situations, le moment de l'interaction pour générer des résultats à « la volée ».

Se pose donc avec acuité la question de l'enseignement et de la didactique des connaissances informatiques nécessaires à donner aux usagers pour satisfaire leurs besoins d'utilisateurs. Les notions d'algorithme, d'indexation en sont quelques exemples, mais elles ne semblent pas suffisantes. De même que se pose aussi la question des références possibles de ces savoirs et donc des modèles informatiques à considérer : ceux des concepteurs ne sont pas ceux des scientifiques, ceux des référenceurs ne sont pas ceux des documentalistes.

Des recherches complémentaires seront menées sur les représentations des usagers concernant les moteurs de recherche, notamment sur le référencement, et sur la perception et la compréhension des innovations proposées par les interfaces. De même, des recherches sur la gestion de l'information recueillie, collectée, annotée, etc. pourrait constituer un autre pan de la recherche en didactique de l'informatique.

Plus généralement, la question qui se pose est bien celle de la compréhension de la manière dont nous intégrons ces dispositifs dans nos modes de pensée.

Bibliographie

- Albert, F. (2009). *BCIU Clermont-Ferrand - Enquête sur les besoins des doctorants clermontois en formation à la recherche documentaire*. Document consulté le 21 juillet 2010 de <http://www.enssib.fr/bibliotheque-numerique/notice-40779>
- Baltz, C. (1998). « Une culture pour la société de l'information ? Position théorique, définition, enjeux ». *Documentalistes – Sciences de l'information*, 35(2),75-82.
- Baron, G.-L. (2009). « Introduction. Pratiques, compétences et savoirs dans l'enseignement supérieur ». In G.-L. Baron, E. Bruillard & L.-O. Pochon (dir.), *Informatique et progiciels en éducation et en formation* (pp. 9-17). Lyon: INRP.
- Baron, G.-L., & Pochon, L.-O. (2009). « Enjeux didactiques de l'informatique et de ses outils. Vingt ans après : regards sur un cheminement ». In G.-L. Baron, E. Bruillard & L.-O. Pochon (Ed.), *Informatique et progiciels en éducation et en formation* (p. 9-17). Lyon: INRP.
- Boubée, N., & Tricot, A. (2010). *Qu'est-ce que rechercher de l'information ?* Villeurbanne :Presses de l'enssib.
- Chiarabella, Y., & Mulhem, P. (2007). La recherche d'information - De la documentation automatique à la recherche d'information en contexte. *Document numérique*, 10(1), 11-38.
- Chang, Y., Arroyo, M. A., Spink, A., Aung, M. T., Thu, K., Lwin, Y., & Htike, Z. W. (2008). Search Engines' Help Systems. In *Proceedings of the Fifth international Conference on information Technology: New Generations* (April 07 - 09, 2008) (p. 309-314). ITNG. IEEE Computer Society. Washington, DC.
- De Souza, C. S. (2005). *The Semiotic Engineering of Human-Computer Interaction*. Cambridge, Mass : USA.
- De Loupy, C., & Crestan, E. (2004). « SRI et traitement du langage naturel ». In Ihadjadene M. (ed.), *Les systèmes de recherche d'informations: modèles conceptuels*. (p. 139-161). Paris : Hermès Lavoisier.
- Drot-Delange, B. (2011). « Comprendre l'interrogation des moteurs de recherche : une expérimentation auprès d'étudiants en information et communication. ». In H2PTM : Hypertextes, Hypermedia. Produit, Outils et Méthodes. Metz (France),12-14 Octobre 2011.

- Duchâteau, C. (2002). Mais qu'est la didactique de l'informatique devenue ? Présenté à *Technologies informatiques en éducation : Perspectives de recherches, problématiques et questions vives*. Paris.
- Hendry, D. G., & Efthimiadis, E. N. (2008). « Conceptual Models for Search Engines ». In A. Spink, M. Zimmer (ed.) *Web Search, Information Science and Knowledge Management*, (14), 277-307. Springer Berlin Heidelberg.
- Henriet, O., Malingre, M., & Serres, A. (2008). *Enquête sur les besoins de formation des doctorants à la maîtrise de l'information scientifique dans les Ecoles doctorales de Bretagne*. Document consulté le 21 juillet 2010 de <http://www.enssib.fr/bibliotheque-numerique/notice-1793>
- Ihadjadene, M., & Fondin, H. (2004). « Le modèle booléen ». In Ihadjadene M. (ed.), *Les systèmes de recherche d'informations: modèles conceptuels*. (p. 19-34). Paris : Hermès-Lavoisier.
- Jeanneret, Y. (2000). *Y a-t-il (vraiment) des technologies de l'information ?* Presses Universitaires Du Septentrion.
- Juanals, B. (2003). *Culture de l'information. Du livre au numérique*. Paris : Hermès.
- Koenemann, J., & Belkin, N. J. (1996). "A case for interaction : a study of interactive information retrieval behavior and effectiveness". In *Proceedings of the SIGCHI conference on Human factors in computing systems common ground - CHI '96* (p. 205-212). Vancouver, British Columbia, Canada. doi:[10.1145/238386.238487](https://doi.org/10.1145/238386.238487)
- Marchionini, G. (2008). "Human-information interaction research and development". *Library & Information Science Research*, 30(3), 165-174.
- Muramatsu, J., & Pratt, W. (2001). « Transparent Queries ». In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '01* (p. 217-224). New Orleans, Louisiana, United States.
- Norman, D. A. (1987). "Some observations on mental models". In R. M. Baecker (Ed.), *Human-computer interaction* (p. 241-244). San Francisco, CA, USA : Morgan Kaufmann Publishers Inc.
- Norma, D.A. (1986). "Cognitive Engineering" . In Norman, D. A., & Draper, S. W. (Ed.). *User Centered System Design; New Perspectives on Human-Computer Interaction* (p. 31-65). Hillsdale, NJ, USA: L. Erlbaum Associates Inc.
- Nicolle, A. (2001). « La question du symbolique en informatique ». *Colloque interdisciplinaire en Sciences Cognitives, ARCo'2001*, 12-14 décembre 2001, Institut des Sciences cognitives, Lyon.
- Nicolle, A. (2002). « L'informatique en éducation entre science et technique ». In Georges-Louis Baron, Éric Bruillard (éds.), *Les technologies en éducation. Perspectives de recherche et questions vives*. Actes du symposium international francophone, Paris, 31 janvier et 1^{er} février 2002, INRP, MSH et IUFM de Basse-Normandie, p. 177-190.
- Nicolle, A. (2003). « Etude préliminaire à une théorie des processus interactifs infinis ». *Cahiers du Greyc | UMR6072 - GREYC - Groupe de REcherche en Informatique, Image, Automatique et Instrumentation de Caen*.
- Pochon, L.-O. et Bruillard, E. (2006). « Présentation ». In L.-O. Pochon, E. Bruillard & A. Maréchal (ed.), *Apprendre(avec) les progiciels. Entre apprentissages scolaires et pratiques professionnelles*. (p. 15-24). Neuchâtel. Lyon: IRDP-INRP.
- Ravenstein, J., & Ladage, C. (2006). « Pour une didactique de l'usage d'Internet ». In A. Piolat (ed.), *Lire, Ecrire, Communiquer et Apprendre avec Internet* (pp. 489-507). Marseille: Solal.
- Serres, A., & Le Deuff, O. (2009). « Outils de recherche : la question de la formation ». In B. Simonnot & G. Gallezot (ed.), *L'entonnoir. Google sous la loupe des sciences de l'information et de la communication*. (pp. 93-111). Caen: C&F Editions.
- Serres, A., & Thiault, F. (2010). « Six questions vives de la culture informationnelle ». In Serres, A., Duplessis, P., Le Deuff, O., Ballarini-Santonocito, I., Kerneis, J., Thiault, F., & Montaigne, A. (2010). *Culture informationnelle et didactique de l'information. Synthèse des travaux du GRCDI, 2007-2010*.
- Simonnot, B. (2008). « Être usager de l'information en ligne nécessite-t-il de nouvelles compétences? ». In J. Dinet (Ed.), *Usages, usagers et compétences informationnelles au 21e siècle* (p. 21-39). Paris: Hermès Lavoisier.

- Simonnot, B. (2009a). « De l'usage des moteurs de recherche par les étudiants ». In B. Simonnot et G. Gallezot (Ed.), *L'entonnoir. Google sous la loupe des sciences de l'information et de la communication* (p. 31-57). Caen : C&F Editions.
- Simonnot, B. (2009b). « Culture informationnelle, culture numérique : au-delà de l'utilitaire ». *Les cahiers du numérique*, 5(3), 25-37.
- UNESCO (2005). Rapport mondial de l'UNESCO. Vers les sociétés du savoir, Paris, Unesco.
- UNESCO (2007). Introduction à la maîtrise de l'information. Paris : UNESCO.