HAL
open science

Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques

Jean Ravestein, Caroline Ladage, Samuel Johsua

To cite this version:

Jean Ravestein, Caroline Ladage, Samuel Johsua. Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques. Revue française de pédagogie, 2007, pp.71 - 83. 10.4000/rfp.495 . hal-01444796

HAL Id: hal-01444796

https://amu.hal.science/hal-01444796

Submitted on 24 Jan 2017

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.

Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques

Jean Ravestein, Caroline Ladage, Samuel Johsua

Utiliser des moteurs de recherche d'information sur Internet devient aujourd'hui possible et se voit recommandé dans des contextes pédagogiques (B2i, C2i). Toutefois, l'usage des outils de recherche dans les classes ne va pas sans poser des problèmes dont cet article voudrait rendre compte selon deux aspects majeurs : celui du manque d'entendement des élèves et des enseignants concernant le fonctionnement des moteurs de recherche, qui leur reste opaque ; celui de la pertinence et de la nature des résultats fournis par les moteurs en réponse à une requête. Ainsi, notre travail porte sur le fonctionnement même des automates de recherche standards qui fournissent des résultats en situation réelle de consultation dans des classes. Nous montrons d'une part une grande difficulté d'exploitation des ressources de la toile par des enfants de l'école élémentaire avec les moteurs habituellement utilisés, pour des raisons techniques, et, d'autre part, nous mettons à jour un questionnement éthique auquel nous renvoie cette nouvelle pratique pédagogique.

Descripteurs (TEE) : compétence en recherche d'information, école primaire, éthique, Internet.

UTILISER INTERNET À L'ÉCOLE : ENTRE VOULOIR ET POUVOIR FAIRE

Les activités de recherche documentaire informatisée sur Internet se développent aujourd'hui largement dans tout le système éducatif et de formation, vivement encouragées par les institutions jusqu'à faire partie des référentiels de compétence comme le « Brevet informatique et Internet » (B2i), de l'école élémentaire au lycée, et le « Certificat informatique et Internet » en formation continue et à l'université (C2i). Différents travaux ont pointé depuis longtemps les aménagements nécessaires à l'utilisation de machines en éducation et formation (Baron & Bruillard, 1996 ; Baron & Glikman, 1996), et en particulier les singularités des rapports des usagers avec ce nouveau type de média (Linard, 1990, 1995 & 2001 ; Bélisle & Linard, 1996). Aujourd'hui, les hypermédias

éducatifs proposent aux élèves des fonctionnalités d'accès multiples à des informations hétérogènes (Baron & Bruillard, 2002), en particulier via Internet.

Le discours institutionnel à propos de l'usage d'Internet à l'école est sans ambiguïté : « à mesure que l'Internet intègre la vie quotidienne, sa maîtrise s'avère indispensable à l'ensemble des élèves. L'école a un rôle crucial à jouer comme lieu d'apprentissage et de familiarisation à l'Internet. Dans le contexte plus récent de développement d'Internet, élèves et enseignants doivent à la fois s'approprier ces technologies et leurs enjeux et en faire un véritable outil au service de la réussite scolaire. Il est donc nécessaire de favoriser une meilleure connaissance de l'Internet afin de développer l'utilisation de ces technologies par l'ensemble des acteurs scolaires » (Darcos, 2003).

Or, l'usage d'Internet à l'école n'est en rien comparable à celui d'une simple base de données qui n'a pas d'intention à notre égard et dont la constitution reflète une logique qui peut être transparente pour l'utilisateur. En effet, les automates de recherche (moteurs) manifestent d'une certaine manière l'intentionnalité d'entrer en contact avec l'utilisateur à partir d'une interprétation de ses requêtes. Ainsi, renvoient-ils souvent des informations et des contenus non désirés, filtrés et qui semblent parfois erratiques, selon des fonctionnements obscurs pour l'utilisateur. Alors qu'aujourd'hui les moteurs prennent de plus en plus cette place d'intermédiaire entre l'utilisateur et l'information disponible sur Internet, il y a lieu de se poser d'une part la question de la valeur des contenus publiés sur le *Web*, et, d'autre part, de questionner la pertinence des résultats fournis par ces moteurs.

La confiance que l'on accorde à ces résultats soulève des questions concernant leur usage en éducation et l'utilisation d'Internet par les élèves devient un sujet de préoccupation pour parents et enseignants.

On sait que la constitution d'un « *milieu pour l'étude* (1) » (Johsua & Félix, 2002) passe par l'établissement de rapports de l'élève avec des objets de savoir explicités dans l'institution ; l'usage des automates de recherche sur Internet semble, aujourd'hui encore, un objet peu explicité à l'école et donc pose problème si l'on veut s'en servir de constituant du « milieu pour l'étude ». Par ailleurs, l'analyse de l'activité de recherche d'informations avec des machines met en évidence que c'est une activité complexe qui met en jeu des processus cognitifs particuliers (Rouet & Tricot, 1998 ; Rouet & La Passardière, 1999) et l'on connaît les difficultés des élèves lorsqu'ils cherchent des informations dans des bases documentaires de type hypermédia (Beaufils, 1998 & 1999 ; Blondel, 2001 ; Macedo-Rouet *et al.*, 2006). Un certain consensus se dégage de l'ensemble de ces travaux : pour tenter de réduire ces difficultés il faut proposer des solutions méthodologiques ou instrumentales, mettre en place une formation spécifique des enseignants et éventuellement utiliser des outils *ad hoc* avec les élèves comme l'assistant de recherches sur Internet (ARI) de F.-M. Blondel, J.-C. Le Touzé et M. Tarizzo (2002).

À travers le monde, nombre d'expériences sur la fréquentation de l'Internet par des utilisateurs qui en ont au départ une connaissance factuelle et pragmatique ont été conduites et nous commençons à pouvoir regarder le phénomène avec un certain recul.

Ainsi, l'étude récente sur les utilisateurs de moteurs de recherche (*Pew Internet* et *American Life Project*, 2005) indique, d'une part, que les utilisateurs sont très satisfaits de leurs moteurs, et sont, d'autre part, parfaitement ignorants de la manière dont ils fonctionnent et présentent les résultats.

Parmi les résultats les plus marquants on note que 92 % des personnes interrogées (20-30 ans, bon niveau d'éducation) font confiance aux moteurs (52 % leur font « très confiance »). Pour 68 % de la population de l'enquête, les outils de recherche sur Internet sont considérés comme des sources d'information légitimes et non biaisées. D'une part, seulement 38 % des internautes sont conscients de l'affichage de liens sponsorisés et d'informations commerciales dans les résultats fournis et arrivent à les reconnaître. D'autre part, 57 % des personnes interrogées ignorent que les moteurs conservent des données relatives à leur navigation (*tracking*) et que ces données peuvent être réutilisées dans un cadre commercial.

À plus grande échelle, une des conclusions de la mise à disposition des ordinateurs « bruts » de l'expérience « *Hole in the Wall* » (2) menée en Inde est qu'elle ne développe qu'une habileté technique et non une démarche critique susceptible de permettre une évaluation des informations trouvées sur Internet (Warschauer, 2003). Dans son rapport de la « Mission Internet, école et famille » sur « le projet Proxima » (France, 2003b) B. Benhamou parle à ce propos de « forme de stagnation sur des usages techno-ludiques sans réelle portée pédagogique ».

La question de la qualité des contenus d'Internet est intimement liée aux modes d'utilisation des outils de recherche. Il est très difficile de déterminer la qualité d'une page *Web* à laquelle l'internaute accède en cliquant *via* un lien de la liste de résultats fournis par un moteur de recherche. Les critères du positionnement

ordonné des résultats sont opaques pour l'usager. Pourtant, l'arrivée en tête de liste leur confère parfois abusivement un statut de qualité. Seuls ceux qui connaissent le fonctionnement de l'application d'algorithmes de classement des informations indexées par les moteurs de recherche sont conscients des effets biaisés. Si cette connaissance des techniques d'indexation de l'information ne garantit pas une recherche d'information efficace, elle permet à l'usager de mieux cibler et de mieux formuler ses requêtes et, partant, d'améliorer d'une part sa perception du bruit, page proposée n'ayant aucun rapport avec le sujet recherché, et, d'autre part, du silence, nombre de références utiles mais absentes (Ravestein & Ladage, 2006). Cette connaissance permet à toute personne productrice d'informations sur le *Web* de mieux comprendre la manière dont ses pages seront traitées, ou évitées, par les robots des moteurs de recherche (Ladage & Ravestein, 2005).

Concernant la perception des contenus sur Internet, une étude du Centre de liaison de l'enseignement et des moyens d'information (Clemi, 2003) montre que les jeunes internautes sont généralement confiants sur la qualité des contenus que l'on trouve sur le *Web* et que c'est généralement grâce aux informations données par les parents ou l'école qu'ils ont appris à visiter les sites avec discernement et à mieux déterminer la qualité des contenus.

Il est important de souligner qu'il se dégage de l'ensemble des enquêtes nationales et internationales (3) que la question de la validité des informations n'est pas un aspect sur lequel les internautes s'interrogent eux-mêmes spontanément.

D. Boullier et F. Ghitalla (2004) posent clairement cette problématique générale : « le *Web* est – par nature et fondation historique – sans référent central, sans source ni origine, et, partant, sans processus de validation des pages, des sites, laissant place à la multiplication *ad hoc* des formats sociotechniques de document. C'est cette absence d'autorité qui laisse la place à la domination d'un système technique opaque, aux réactions intempestives et aléatoires. C'est alors le système général en tant que tel qui doit d'abord être maîtrisé ».

D'autres chercheurs ont également bien pointé ces problèmes et ont produit des recommandations.

Ainsi, P. Brandt-Pomares (2003), à la lumière de différentes études, relève que la tâche de recherche sur Internet n'est pas d'un abord si aisé par des élèves et qu'il ne suffit pas de disposer de l'équipement matériel afin d'effectuer des recherches efficacement.

S. Proulx (2001) et B. Lelong (2002) ont montré qu'une démarche critique est nécessaire et passe par une sensibilisation des jeunes usagers aux techniques qui sous-tendent le fonctionnement d'Internet pour permettre une forme d'objectivation à travers l'analyse des informations recueillies.

L. Le Roux (2004) s'interroge à partir d'études menées auprès de documentalistes sur l'appréciation délicate des documents trouvés sur le *Web* : « L'équation traditionnelle : document = support + information devient, avec les technologies en réseau, caduque et ne permet pas encore une définition précise et consensuelle des aspects polymorphes du document numérique. Comment pourrions-nous le définir : sur sa structure (*forme*), son sens (*signe*) ou sa légitimation (*médium*) ? »

L.-J. Gurak (2001) a dégagé une démarche analytique d'évaluation des informations recueillies sur Internet qu'elle lie à celle qui prévaut lors de la mise en place des informations sur un site *Web*. Elle propose une liste d'éléments à prendre en considération, en particulier :

- l'auteur ou la source de l'information : le fait que cette mention soit absente ou qu'elle soit particulièrement difficile d'accès interroge la validité de la source consultée ;
- la fréquence et la nature des mises à jour : l'information sur ces points permet d'apprécier la pérennité de la source, critère d'évaluation important sur le *Web* ;
- clarté de l'objet du site ainsi que les cibles visées : elles permettent d'évaluer le niveau d'information que l'on peut attendre *a priori* (grand public, public professionnel, etc.) ;
- l'adresse d'un site, son URL (pour *Uniform* ou *Universal Resource Locator*) : en plus de la localisation géographique (« .fr », « .de », « .jp », etc.) l'adresse d'un site permet de déterminer s'il est lié à une entreprise commerciale, une association, une institution (« .com », « .org », « .biz », « .info », « .museum », « .aero », etc.) ou s'il émane d'un particulier ;
- le type des liens qui émanent d'un site : il permet d'évaluer le degré de précision des informations car ces liens donnent de nombreuses informations sur les sites considérés comme des références pour les auteurs. Cette information désigne les contours de la « sphère » à laquelle appartient le site considéré ;
- l'observation des liens qui renvoient vers un site donné : elle permet d'évaluer la « place » qu'occupe

ce site sur le *Web*. La notoriété d'un site se traduit presque toujours par un surcroît de liens qui renvoient vers ses pages ;

- la conception éditoriale, technologique et graphique : elle donne de nombreux éléments d'appréciation sur le « professionnalisme » de la démarche des éditeurs d'un site.

Si rechercher *via* les moteurs est une pratique répandue chez les jeunes internautes, l'appréciation de ce qu'ils trouvent et donc l'utilisation de ces informations dans un contexte pédagogique reste problématique.

En effet, tant que des moyens intellectuels d'analyse, fondés sur des méthodologies du type celle de L.-J. Gurack, ne seront pas promu(e)s à l'école, l'utilisation légitime des informations fera défaut et l'usage d'Internet par les élèves restera suspect aux yeux des enseignants.

Après que la question d'une utilisation d'Internet sans danger pour les enfants ait déjà été largement débattue depuis plusieurs années et ait fait l'objet de diverses études nationales et internationales, les initiatives qui ont été mises en œuvre jusqu'à ce jour pour améliorer son utilisation ne semblent pas encore avoir répondu aux problématiques sans cesse changeantes qu'engendrent les différents usages de la toile.

Le plus souvent, c'est seulement l'accès des enfants à certains sites de nature particulière qui inquiètent les éducateurs (Soëtard, 2000). Les initiatives nationales, avec le soutien de la Commission européenne, se sont multipliées sous la forme d'actions de protection des enfants utilisateurs d'Internet et de sensibilisation des différents groupes d'utilisateurs, parents, enseignants et étudiants, mais aussi fournisseurs de contenus et de services.

Ces initiatives ne concernent que les problèmes d'accès au contenu et n'évoquent que marginalement la protection des mineurs à l'égard de leurs données personnelles comme le constate en 2001 un rapport de la CNIL (France, 2001).

Dans un souci d'assurer que la navigation sur Internet s'effectue avec une plus grande sécurité, différentes solutions techniques ont été développées. Des systèmes de filtrage (4) sont le moyen le plus répandu mais ne sont pas simples à mettre en place pour les parents et les éducateurs et ne présentent pas une fiabilité absolue. Leur utilisation n'est donc pas suffisante et ne pourra aucunement remplacer le rôle des éducateurs ou des parents, ce qui repose la question d'un apprentissage à l'utilisation d'Internet.

Citons la réponse institutionnelle du programme européen « *Safer Internet* » qui a été récemment mis en place dans le cadre du programme d'éducation aux médias « Educanet » particulièrement centré sur les risques liés à l'usage d'Internet. Son objectif est d'apprendre aux jeunes à naviguer en toute responsabilité sur les réseaux. Il s'adresse aux enseignants, parents et éducateurs et propose une démarche et une série d'activités éducatives. Chaque enseignant peut ainsi y trouver des ressources.

Citons enfin les chercheurs qui, avec une approche différente du programme précédemment cité, mettent l'accent sur la nécessité d'une normalisation de l'information mise en ligne et de son indexation. Ainsi pour M. Grandbastien (2002), dans le sillon des défenseurs du *Web* sémantique, se mettre d'accord au niveau mondial sur des formes de description de ressources est un enjeu important.

On le voit, il y a loin entre vouloir faire et pouvoir faire en matière d'utilisation d'Internet à l'école.

Ainsi, dans le but d'apporter une contribution à la description de l'état des pratiques concrètes de l'usage de l'Internet par des élèves, notre travail porte sur le fonctionnement même des automates de recherche qui rendent des résultats de requêtes en situation réelle de consultation par de jeunes élèves.

QUELLES REQUÊTES, QUELS MOTEURS ?

En contexte pédagogique, les requêtes concernent des objets de savoir qui sont vifs dans la classe. On admet en effet qu'un certain contrôle de la part de l'enseignant s'exerce sur cette pratique, ce qui lui donne une certaine légitimité dans et hors de la classe.

Nous avons donc enquêté dans des classes d'école élémentaires équipées d'ordinateurs connectés et dans lesquelles la pratique d'interrogation d'Internet fait partie des activités usuelles. On pourrait dire que l'usage de la requête est dans ces classes assez « naturel », mais pas « sauvage », c'est-à-dire qu'il s'effectue selon des besoins et le plus souvent se trouve inclus dans une situation didactique. Dans ces classes, conformément aux recommandations du B2i (5), à l'occasion d'enquêtes, de travaux pratiques encadrés, les élèves recherchent des informations sur Internet, individuellement ou par petits groupes. Ces informations sont relatives à des sujets variés et les requêtes sont formulées sous formes d'expressions tapées dans les fenêtres des moteurs de recherche. Ces expressions comprennent des mots-clés concer-

nant le domaine de l'enquête entreprise (Histoire, Géographie, Sciences de la vie et de la terre, etc.). Les informations recueillies sont ensuite exploitées pour des exposés, la réalisation de posters, le plus souvent avec l'aide de l'enseignant.

Nous avons cherché à lister une dizaine de requêtes par expression-clé qui sont « courantes », voire fréquentes à ce niveau de l'école élémentaire. Notre intention, en établissant cette liste, est de sélectionner des mots-clés qui constituent un échantillon assez représentatif d'un fonctionnement ordinaire de l'utilisation des moteurs dans une classe.

Nous avons procédé à une sélection à partir d'une liste plus vaste (59 mots-clés) recueillie auprès de différentes sources (élèves, enseignants, intervenants), à différents moments (milieu et fin d'année scolaire) et lieux (six classes de cours moyen d'écoles ordinaires : trois urbaines, trois rurales, hors ZEP).

Notre échantillonnage ne couvre certainement pas l'ensemble des requêtes vraisemblablement possibles à l'école (en premier cycle, par exemple) et ignore la variable de sectorisation (ZEP, REP, etc.). Toutefois, notre recherche ne portant pas directement sur l'influence de ces variables sur l'expression des requêtes, nous considérons que nos choix sont légitimes pour notre problématique et ne limiteront *a priori* que dans une faible mesure la portée de nos résultats.

Dans un premier temps nous avons éliminé les requêtes à champ très vaste comme « les soldats », ou trop restreint, comme « montagne Sainte Victoire », ou bien encore celles dont la polysémie peut laisser prévoir des résultats particuliers, comme « les poilus », en histoire.

Dans un deuxième temps nous avons criblé la liste restante à partir des critères suivants :

- les requêtes doivent couvrir un large champ des programmes de l'école élémentaire ;
- elles doivent être assez courtes pour correspondre à la manière spontanée d'interroger les moteurs observée chez les jeunes élèves de notre enquête (mots plutôt que phrases) ;
- elles doivent être peu exotiques, c'est-à-dire interroger des champs du savoir bien présents dans le milieu de la classe.

Ainsi, nous avons retenu les neuf expressions de requêtes suivantes :

Lombric – Charlemagne – Cycle de l'eau – Démocratie – Jacques Prévert – Homme préhistorique – Polygones – Polyèdres – Fables de La Fontaine.

Certes, à ce stade de construction de l'échantillon, on aurait pu être tenté d'intégrer dans la liste des mots-clés plausibles en classe et choisis judicieusement par le chercheur pour les probables résultats spectaculaires qu'ils auraient produits (« femme préhistorique », par exemple), ou bien pour équilibrer les disciplines (en rajoutant « Mozart », par exemple). Nous avons choisi de partir uniquement de la liste de base, dressée sur le terrain, afin de conserver à cette recherche un caractère exploratoire.

Ces neuf requêtes ont été proposées aux trois moteurs les plus utilisés dans la réalité (6) : Google, Yahoo et MSN. Il faut préciser qu'il ne s'agit en aucun cas de vouloir par cette recherche établir un palmarès des moteurs choisis pour l'expérimentation.

L'objectif méthodologique était de recueillir les informations de la première page de résultats données par ces trois outils de recherche à chacune des requêtes retenues pour notre expérimentation. Quotidiennement pendant trois mois nous avons enregistré ces pages de résultats sous format HTML afin de pouvoir les récupérer dans un tableur et effectuer différents niveaux d'analyse. Celle-ci ne comportera pas une description des ressources trouvées, qui reverrait à une méthode plus qualitative devant exploiter une masse extrêmement importante d'informations, ce qui constitue une autre piste très intéressante de recherche.

Nous avons pu identifier différents éléments importants à vérifier :

- la qualité « technique » des résultats selon les deux critères suivants :
  - *bruit* : adresses ne correspondant pas du tout à la requête (7) ;
  - *code d'erreur et URL inactif* : toute adresse inactive lors de l'interrogation et tout code ne permettant pas l'affichage, (page introuvable, etc.).

Trois autres critères pourraient êtres retenus, mais nous les avons négligés pour les raisons suivantes :

- *silence* : absence d'une adresse pertinente (8) (ou qui se retrouve au-delà de la première page de résultats) : ceci supposerait de pouvoir analyser la pertinence de milliers de sites, ce qui est hors de portée de moyens logistiques ordinaires ;
- *doublon* : liens différents pointant vers des pages ayant le même contenu (site-miroir ou migration d'un site) : il semble que ce phénomène devient de plus en plus rare et peut être négligé ;
- *dysfonction* : d'un champ, d'une fonction ou de tout autre élément du langage de recherche (9) : pour les mêmes raisons que pour « doublon ».

- la stabilité des résultats dans le temps pour chacun des moteurs : mesure de l'apparition et de la disparition des liens, en particulier des liens pertinents ;
- la « qualité » du site en première position sur la page de résultats correspondant véritablement à la question exprimée (ou sous-entendue) par les mots de la requête (10), selon trois modalités : site institutionnel contrôlé, site « privé », bruit (URL inactive ou site hors sujet).

Notre analyse prendra en compte également la présence de liens pointant des pages au contenu douteux vis-à-vis de notre population : sites à caractère pornographique, politiquement tendancieux, etc.

PERFORMANCES DES MOTEURS

Dans un premier temps, le dépouillement des données journalières nous indique que l'intérêt d'en faire un traitement statistique complet était réduit vis-à-vis de nos questions. En effet, la stabilité des résultats dans le temps pour chacun des moteurs montrait une variabilité évidente seulement sur un délai d'un mois. Aussi avons-nous choisi de traiter nos questions (qualité technique, stabilité, qualité du champ sémantique) à partir d'une base de 810 lignes correspondant aux dix premiers résultats des trois moteurs sur neuf requêtes distinctes à trois dates : 5 juillet 2005, 5 août 2005, 5 septembre 2005.

J. Véronis (2005) a attiré notre attention sur le fait qu'à des dates précises et sans préavis, les moteurs rectifient leurs index et que donc nos captures pouvaient « tomber » dans une époque de remaniement et ainsi biaiser nos résultats. Nous nous sommes donc assurés avec l'outil de veille approprié (*TrendMapper*) (11) que les dates que nous avions choisies pour « arrêter » nos observations ne correspondaient pas à une modification drastique des index, mais plutôt à une période de fonctionnement ordinaire.

Pour évaluer la qualité technique des moteurs, nous avons tout d'abord procédé à un « nettoyage » et à un codage des URL présents sur chaque première page de résultats pour chaque moteur et chaque requête. En effet, certains liens renvoient sur les mêmes pages avec une URL légèrement différente. La base de données que nous allons traiter comprend ainsi, après ce travail, des URL qui renvoient sur des pages distinctes.

Ce que nous appelons « bruit » dans le tableau récapitulatif ci-après correspond au nombre moyen de liens inopérants techniquement et/ou totalement hors sujet qui apparaissent sur la première page de résultats de chacun des trois moteurs pour chacune des neuf requêtes sur les trois captures à un mois d'intervalle.

On remarque une évidente et parfois importante variabilité entre chaque moteur et entre chaque requête. Google semble en moyenne le plus fiable. Certaines requêtes engendrent une telle proportion de bruit (Charlemagne pour MSN, 96,7% ; par exemple) qu'on peut se demander si sur certains moteurs des algorithmes d'analyse du langage sont réellement mis en place.

Pour un élève en situation ordinaire de consultation du *Web*, on peut donc dire que sa recherche va être considérablement ralentie, voire générer des égarements ou découragements à cause de la quantité de bruit qui reste importante et surtout, pour lui, aléatoire (12).

Tableau I. – **Pourcentages de bruit par requête et par moteur**

Requêtes	Google	MSN	Yahoo	Total bruit moyen
Cycle de l'eau	0,0 %	0,0 %	3,3 %	1,1 %
Fables de La Fontaine	0,0 %	3,6 %	0,0 %	1,2 %
Démocratie	3,6 %	3,3 %	0,0 %	2,3 %
Polygones	8,3 %	3,3 %	10,0 %	7,1 %
Jacques Prévert	3,3 %	33,3 %	10,0 %	15,6 %
Homme préhistorique	12,5 %	33,3 %	30,0 %	26,2 %
Polyèdres	25,9 %	24,1 %	36,7 %	29,1 %
Charlemagne	10,0 %	96,7 %	33,3 %	46,7 %
Lombric	33,3 %	66,7 %	60,0 %	53,3 %
Total bruit moyen	10,9 %	29,6 %	20,4 %	20,5 %

Tableau II. – **Mouvance moyenne des moteurs par requête**

Requêtes	Google	MSN	Yahoo	Total moyen
Cycle de l'eau	0,80	2,24	1,15	1,40
Polyèdres	1,17	2,43	1,73	1,78
Lombric	0,92	1,62	2,93	1,82
Homme préhistorique	1,60	2,75	1,88	2,08
Fables de La Fontaine	2,07	2,43	2,00	2,17
Charlemagne	1,77	2,56	2,37	2,23
Polygones	1,20	2,81	2,72	2,24
Jacques Prévert	1,42	2,84	2,53	2,26
Démocratie	0,88	3,35	2,78	2,34
Total Moyen	1,31	2,56	2,23	2,04

On pourrait suggérer ici l'utilisation d'un type de moteur fonctionnant sur des index particuliers comme le projet Google « Scholar », adapté au domaine de recherche de l'internaute, fut-il un élève de l'école élémentaire. Un moteur qui s'adapterait au « niveau » de l'internaute est à l'ordre du jour comme en témoignent les progrès récents de certains moteurs comme *Exalead* (13).

Pour mesurer la stabilité des résultats dans le temps, nous avons calculé un indice moyen en attribuant à chaque déplacement d'une URL un score (par exemple, si URL1 pour « lombric » passe de 1er rang au 6e rang entre le 5 juillet et le 5 août, elle a un score de 5, etc.). Nous avons développé un outil d'analyse *ad hoc* (programme en PERL) qui a permis de traiter notre base de données importante sans erreur de calcul.

Ce que nous cherchons à mettre en évidence c'est un aperçu du taux de fluctuation des URL sur les pages de résultats qui se présentent aux élèves suite à une même requête faite à plusieurs moments, ce qu'on pourrait nommer « mouvance » des moteurs.

Nos calculs nous donnent le tableau récapitulatif suivant (où « 1 » correspond à la variation moyenne d'un rang) :

D'une manière générale, on peut conclure à une mouvance acceptable, avec toujours un petit avantage pour Google. On voit par contre qu'il y a une variabilité inter-requêtes, des mots-clés semblent plus stables que d'autres et ce phénomène est variable selon les moteurs. Par exemple, si « démocratie » est celui qui bouge le plus en moyenne, il est très stable sur Google (moins d'un rang), le terme est-il « sensible » chez Microsoft (MSN, plus de deux rangs) ? Ainsi, si le calcul moyen renvoie l'image d'une certaine robustesse temporelle globale du fonctionnement des algorithmes des moteurs qui, au fond, assure leur succès légitime, on peut néanmoins remarquer qu'à l'échelle d'un utilisateur ordinaire, le fait que tel lien classé en n° 1 le 5 juin 2005 qui passe n° 2 le mois suivant n'est pas sans importance.

En effet, on sait que les internautes commencent massivement par tester le premier lien avant de retourner à la page de résultats s'ils ne sont pas satisfaits, pour ensuite tester le second, etc. (s'ils ne se sont pas déjà égarés à partir des liens proposés sur la page correspondant au premier lien).

Par exemple, le 5 juin 2005, la requête « lombric » sur Google renvoie en n° 1 sur un site qui traite effectivement de l'animal, alors qu'un mois plus tard il renvoie, toujours en n° 1, sur le site d'une association suisse qui n'a rien à voir avec le ver de terre, le site du mois de juin étant passé en n° 2.

Une analyse plus précise sur les liens qui disparaissent de la première page de résultats et migrent dans le silence du *Web* d'une part et ceux qui « montent » en première page d'autre part, nous montre que les techniques de référencement peuvent agir et « forcer » les algorithmes des automates de recherche (14).

Si on rapporte ces résultats aux précédents, on peut dire que les moteurs sont relativement stables mais de manière hétérogène.

Pour affiner notre description nous avons calculé les mouvements moyens du lien n° 1 des premières pages de résultats pour les deux périodes de notre étude, pour chaque requête et pour les trois moteurs :

Tableau III. – **Taux moyens de fluctuation par moteur**

Moteur	Période 1	Période 2	Moyenne
Google	0,1	2,0	1,1
MSN	3,0	3,9	3,4
Yahoo	1,8	0,0	0,9
Total	1,6	2,0	1,8

Là encore, on remarque des fluctuations, opaques pour l'internaute. On note également dans le détail la disparition totale de certains liens de notre échantillon au cours de la période, invisible dans ce calcul de moyenne.

Comment expliquer que Yahoo passe d'une instabilité de 1,8 en première période à une stabilité totale en deuxième, alors même que son index a plutôt plus varié en taille dans la deuxième période (source *TrendMapper*) ?

En tout état de cause, la logique interne des classements des moteurs semble aujourd'hui parfaitement inaccessible à l'utilisateur. Alors, comment bien se servir d'un outil dont on ignore le mode opératoire ?

Pour aller plus avant dans ce questionnement, nous avons ensuite tenté d'évaluer la qualité des pages où nous envoie le premier lien de chaque page de résultats pour nos trois « captures » (vingt-sept sites évalués pour chaque moteur).

Nous avons distribué l'échantillon selon trois critères :

- site institutionnel contrôlé qui traite de la requête : émanation d'une organisation scientifique ou « légale » (CNRS, université, gouvernement, etc.) ;
- site privé : émanation d'une initiative personnelle ou collective qui traite de la requête (pages personnelles, associations, fondations d'entreprises, etc.) ;
- bruit : site qui ne traite pas du tout de la requête ou qui ne fonctionne pas.

On pourrait penser proposer d'autres critères, dérivés d'une analyse plus poussée des pages trouvées, mais une telle démarche conduirait à une distribution moins nette des sites dans les catégories, certaines pages pouvant se distribuer dans plusieurs catégories. Les trois critères que nous avons choisis se justifient en particulier par le fait que le rafraîchissement des informations des sites privés est bien plus faible que dans les sites administrés d'une part, et, d'autre part, les sites institutionnels, mieux référencés, sont sous les feux de la critique et des outils collaboratifs y sont souvent présents pour permettre de prendre en compte des remarques, critiques, ou même les amendements des internautes. Ainsi, la qualité du contrôle de la valeur des contenus est très probablement meilleure dans les sites administrés, plus dynamiques, plus visibles sur la toile.

Nous obtenons les résultats suivants, toutes requêtes confondues :

Tableau IV. – **Répartition des premiers liens par genres**

Moteurs	Site institutionnel contrôlé	Site privé	Bruit
Google	7	18	2
Yahoo	3	20	4
MSN	3	11	13
TOTAL	23	78	34

On observe premièrement une grande quantité de bruit sur les liens en première position, ce qui est relativement gênant pour un jeune internaute ; deuxièmement une fiabilité des ressources au moins discutable, puisque le lien n° 1 renvoie trois fois plus souvent sur des sites privés.

Si, dans ces sites privés, l'information est souvent en rapport direct avec la requête, on peut citer l'exemple pour le lien « démocratie » où six liens sur neuf renvoient à des sites de partis politiques.

On peut remarquer également la présence de plus en plus grande du site *Wikipedia* (15), qui est l'exemple même de source instable avec tous les atours de la source sûre (16).

Par ailleurs, la requête « Jacques Prévert », renvoie huit fois sur neuf sur un site « personnel », à l'ergonomie défaillante, difficilement exploitable par un élève d'école élémentaire.

Par contre, ce même élève n'aura pas de peine à utiliser le site de la catégorie « bruit » qui renvoie sur un jeu en ligne (« jurassiclife.com ») à partir de la requête « homme préhistorique » sur MSN !

À partir de ces premières mises à plat de résultats, on peut donc faire l'hypothèse d'une grande difficulté d'exploitation des ressources du *Web* par des enfants de l'école élémentaire sur des requêtes ordinaires avec les automates habituellement utilisés.

On peut ici relever une incidence éducative de notre travail : il met en évidence qu'un rapport de confiance « naïf » des enseignants et des élèves avec les outils de recherche sur Internet doit être sérieusement questionné car, faute d'alternative, ils risquent bien de faire de nécessité vertu et de continuer d'utiliser les principaux moteurs sans recul critique.

Nous avançons que dans l'avenir, l'examen de cette problématique passera par poser l'hypothèse de la nécessité d'une *catachrèse outil-instrument* dans le processus d'appropriation des TIC par les enseignants. L'étymologie grecque indique que la catachrèse consiste, en quelque sorte, à « tirer la couverture à soi ».

Pour Y. Clot (2003), l'attribution de fonctions nouvelles aux outils ou leur usage détourné et inventif ne s'effectue pas en intériorisant les fonctionnements extérieurs ou les significations existantes mais en les recréant pour leur donner une « deuxième vie » au service de l'activité présente. L'école russe de psychologie historico-culturelle a proposé le concept d'*appropriation* pour désigner cette « double vie » des significations (Leontiev, 1984). Une telle approche transforme l'activité apparemment passive et soumise en activité inventive et créative.

Il y aurait ainsi dans ce cadre théorique des sortes de rhétoriques de l'action de l'enseignant au prise avec les TICE : elles détourneraient les machines de leur fonction officielle ou plutôt développeraient des fonctions pour réaliser, malgré tout, avec elles, une activité locale.

Mais l'observation plus attentive de nos « captures » de pages de résultats ainsi qu'une pratique de la veille technologique ont attiré notre attention sur des phénomènes nouveaux, peut-être importants et redevables de recherches futures. Nous en donnons un bref aperçu ci-après.

INTERNET DANS LA CLASSE : DANGER ?

Dire que l'utilisation d'Internet sans contrôle par de jeunes enfants entraîne des risques importants en terme de « mauvaises rencontres (17) » est une trivialité, pour preuve les démarches commerciales qui assaillent les parents attentifs sous la forme de logiciels de « contrôle parental d'accès à Internet ».

Toutefois, on est en droit d'attendre de nos institutions éducatives qu'elles utilisent des outils pédagogiques qui ne comportent pas ce type de risque, voire même qu'elles garantissent une certaine neutralité laïque des informations disponibles dans la classe, quel que soit le média.

Or, nous nous sommes aperçus qu'une simple approche du problème par une censure en amont, au niveau des requêtes, n'est pas suffisante.

En effet, nous avons relevé des exemples de premières pages de résultats à nos neuf requêtes, légitimes dans le champ d'une classe, dans lesquelles des liens, de surcroît mis en évidence, renvoient vers des sites dont le contenu n'est pas *a priori* « recommandable » pour de jeunes esprits.

On imagine la gêne probable d'un professeur d'école qu'un élève interpellerait en classe alors qu'il vient d'ouvrir le site du premier lien « sponsorisé » (« rencontrez des hommes ») présent sur la page de résultat de MSN suite à la requête « homme préhistorique ».

Que répondre à cet autre élève qui va cliquer sur le lien du site d'un parti politique bien connu alors qu'il a tapé le mot clé « démocratie », etc. (18)

Cette étude nous renvoie à des questionnements plus précis : quel degré tolérable de perméabilité des institutions didactiques aux savoirs qui ne sont pas issus d'une transposition « classique » (au sens de Chevallard, 1991) qui garantit une correspondance avec un savoir légitime et robuste dans le champ de référence scientifique ?

Les liens n° 1 qui pointent vers des sites comme *Wikipédia* ou des sites de particuliers experts et passionnés posent la question concrètement. Le tollé des conservateurs de la Bibliothèque Nationale de France contre l'entreprise « Google Print », ou des « intellectuels » contre *Wikipédia* montrent la vivacité du problème dans la noosphère. Pour le coup, certains des fameux « savoirs sociaux de référence » (Martinand, 1983) ou « experts » (Johsua, 1998) qui entrent d'habitude avec difficulté par la grande porte dans la classe, entrent par la fenêtre via le *Web* sous la forme d'espèces d'*Almanach Vermot* « en ligne ».

Ce phénomène de perte de contrôle est à notre sens un problème fondamental à moyen terme car les savoirs académiques sont traditionnellement organisés selon une logique de catégorisation et de hiérarchisation qui est plutôt celle qui est présente à l'école. C'est une logique d'arborescence descendante, témoin de l'exercice de la pensée de ceux qui savent vers ceux qui ne savent pas et qui doivent l'acquérir

pour accéder à la connaissance. Dictionnaires, index, thésaurus, fonctionnent ainsi et il y a souvent une activité cognitive surajoutée à la requête (maîtriser l'ordre alphabétique parfaitement, par exemple). L'usage de plus en plus répandu d'outils informatisés de recherche d'information avec la possibilité donnée à l'utilisateur d'associer une « étiquette » à tel ou tel mot, tel ou tel texte, telle ou telle image rencontrée sur le *Web* (*tagging*) ou mise en ligne par lui-même, change l'organisation des connaissances et modifie son accès. C'est le phénomène nouveau du « *tag* » (19) qui semble prendre une ampleur que personne n'avait supposée. En effet, plus personne ne contrôle l'appartenance des objets du monde à une catégorie, plus personne ne décide en amont des liens logiques entre ces mêmes objets. On dirait qu'on assiste à une sorte de processus darwinien en direct accéléré en ce qui concerne les outils d'intervention sur le *Web*. Les utilisateurs essayent, testent, adoptent, rejettent avec un pragmatisme surprenant les outils qu'on leur propose : ne survivent que ceux « qui marchent » (facilement), d'où l'importance croissante des préoccupations ergonomiques des concepteurs de contenus en ligne.

L'accroissement des *tags* sur Internet, bien évidemment répercuté sur les résultats fournis par les moteurs de recherche qui ne peuvent « comprendre » ce qu'il faut ou non indexer, fait émerger, par effet de nombre, une véritable *folksonomie* (20) planétaire. Ainsi, on assiste à une sorte de renversement de l'arbre de la connaissance : il pousse par ses branches. Si on fait la requête « chien », on va pouvoir accéder à toutes les photos de chiens que les internautes ont mis en ligne avec cette étiquette, ce qui donne une étonnante diversité de chiens, de situations concrètes « vivantes » (parfois des vidéos) où on voit des chiens « normaux » (le plus souvent « bâtards ») au sens de la statistique : les plus fréquemment observés. Ce ne sont plus uniquement des chiens de race, davantage présents dans les encyclopédies, présentés dans des poses convenues, finalement les plus rares dans la réalité. Tout cela nous conduit à poser deux questions :

- quelle alternative aux moteurs commerciaux pour les enseignants (21) ? Peut-on penser à des moteurs institutionnels, qui n'auront vraisemblablement pas les crédits suffisants pour rivaliser avec des géants comme Google ? ;
- une entreprise d'envergure de type *Wikipédia* « semi-ouverte », qui serait pilotée par les seuls enseignants, implantée sur un serveur INRP, et référencée sur la toile par d'excellents professionnels, pourrait-elle combler le vide laissé par les encyclopédies traditionnelles difficilement exploitables pédagogiquement, en prolongement des initiatives institutionnelles comme *Educasource* (22) ?

Sur un autre plan plus général, comment ne pas voir que, sous l'apparent consensus concernant une désirable « *société des savoirs* », les enseignants, les formateurs, les responsables pédagogiques et les élèves eux-mêmes sont progressivement contraints de s'engager dans des formes d'enseignement dérivées de l'autoformation et mettant l'apprenant (de force) au centre du dispositif ? En d'autres termes, trop souvent, loin d'être un choix décidé suite à une concertation, des dispositifs coûteux d'usage des ordinateurs et de l'Internet à l'école obéissent à un faisceau « *d'ardentes obligations* » pour « *réduire la fracture numérique* » (Compaine, 2001 ; Castells, 2002), qu'il importe d'identifier et de comprendre.

Ainsi, dans ce contexte où des décisions politiques comme, par exemple en France, l'opération « Ordina 13 » (23), obligent d'une certaine manière les institutions éducatives à s'adapter à l'intrusion des TIC, de quel poids pèseront demain les professionnels de l'enseignement ou les institutions de recherche en pédagogie pourtant détenteurs d'expérience et de savoir-faire en la matière ?

Trouver et utiliser des informations sur Internet à l'école renvoie bel et bien à des problèmes techniques pour lesquels on pourrait proposer des solutions, mais pose également de nouvelles questions éthiques auxquelles la communauté éducative devra s'efforcer de répondre, peut-être en amont des grands choix d'équipements et des stratégies de leur déploiement.

Jean Ravestein
jean.ravestein@up.univ-mrs.fr
Université de Provence – Aix-Marseille 1
UMR « Apprentissage, didactique, évaluation, formation » (ADEF)

Caroline Ladage
ladage@analyze.fr
Université de Provence – Aix-Marseille 1
UMR « Apprentissage, didactique, évaluation, formation » (ADEF)

Samuel Johsua
johsua@univ-aix.fr
Université de Provence – Aix-Marseille 1
UMR « Apprentissage, didactique, évaluation, formation » (ADEF)

NOTES

(1) Si la responsabilité finale de l'étude appartient à l'élève, il n'en demeure pas moins que son activité d'apprentissage est fortement déterminée par son insertion dans un espace organisateur de l'étude, où sont désignés, dans une collectivité, des objets de savoirs proprement dits ainsi que des objets pertinents pour l'étude de ces savoirs (le cercle et le compas, par exemple). Cet espace, cet environnement du système didactique, où le triptyque professeur/élèves/savoirs déploie ses effets, nous le désignons sous le vocable de « milieu pour l'étude ».

(2) Des ordinateurs avaient été placés directement dans les murs de certaines villes indiennes (à la manière des distributeurs de billets), après les premiers temps d'exploration, les enfants n'ont pas acquis une maîtrise de l'information mais seulement une dextérité technique.

(3) Une étude internationale menée en 2000-2001 conjointement dans sept pays (Belgique, Espagne, France, Italie, Portugal, le Québec et la Suisse), sur les usages d'Internet par les adolescents, permet de conforter ces observations. (Clemi, 2003).

(4) On annonce pour 2006 en France, selon des technologies développées en Chine où la censure d'Internet est très performante, un filtrage proposé obligatoirement (par la loi) par chaque fournisseur d'accès selon trois niveaux : enfant, adolescent, adultes.

(5) Brevet informatique et Internet (B2i) pour l'école élémentaire dont on peut résumer les objectifs relatifs à l'usage d'Internet : « Trouver sur la toile des informations scientifiques, historiques, géographiques, artistiques et culturelles, les activités sportives, les apprécier de manière simple et critique et les comprendre ».

(6) Nous aurions pu proposer ces requêtes à des moteurs dédiés « éducation » comme *Spinoo*, mais nous avons constaté qu'ils sont utilisés de manière confidentielle sur le terrain.

(7) Ce phénomène est le résultat de la mauvaise prise en compte de l'homographie ou bien au dysfonctionnement de l'opérateur de proximité. Il est dû parfois à la technique de référencement consistant à inclure une liste de mots cachés ne correspondant pas au sujet de la page ou encore des mots au pluriel alors que le sujet est singulier, des faux synonymes, des paronymes ou des homonymes.

(8) Une réponse est dite pertinente lorsque la page pointée contient le sujet principal de la requête de manière évidente (mots titres qui reprennent la requête, illustration qui s'y rapporte, etc.), sans clic supplémentaire.

(9) Par exemple avoir une page en anglais alors que la langue paramétrée est le français.

(10) Une nouvelle étude réalisée par un groupe de chercheurs de la Cornell University tente de comprendre le comportement des internautes lorsqu'ils utilisent un moteur de recherche. Selon ces travaux, plus de 40 % des internautes cliqueraient en premier sur le premier lien proposé, 16 % sur le deuxième, 10 % sur le troisième et 5 à 6 % sur les liens situés de la quatrième à la sixième place. Cf. Joachims *et al.*, 2005.

(11) http://www.trendmapper.com

(12) Nous pourrions développer les raisons qui expliquent ces variations du bruit, mais il faudrait alourdir considérablement notre travail par des considérations techniques qui nous semblent ici inopportunes.

(13) Ce moteur innovant fonde son succès sur une réponse positive aux questions suivantes : (Traduction des auteurs) : Devez-vous accéder à l'information appropriée rapidement ? Avez-vous votre propre manière de classer des choses ? Avez-vous des difficultés pour vous rappeler l'orthographe des noms de personnes ? Ne souhaiteriez-vous pas faire des recherches à la manière dont fonctionne votre cerveau, par sérendipidité ? Détestez-vous révéler vos données personnelles à un « Big Brother » chaque fois que vous recherchez une information ? Et que diriez-vous d'obtenir l'information qui est vraiment pertinente pour vous, et pas celle que les moteurs de recherche principaux pensent est pertinente pour vous ? Information disponible sur http://www.exalead.com (consulté le 9 juin 2006).

(14) Pour une information scientifique quotidienne sur ces sujets : http ://aixtal.blogspot.com.

(15) Ce site semble « remonter » dans les pages de résultats sur l'ensemble de la période observée, toutes requêtes confondues, en particulier avec Yahoo.

(16) Cf. « Wikipedia, une encyclopédie libertaire sur le Net ». *Le Monde*, 2 septembre 2005. Wikipédia est une encyclopédie dont les rubriques sont rédigées et modifiées par les internautes, en ligne, de manière « libre », sans aucun contrôle institutionnel.

(17) Sites à caractère pornographique, violents, incitant à la haine raciale, etc.

(18) Le 11 novembre 2005 date de la commémoration de l'armistice qui marqua la fin de la grande guerre (1914-1918), à la requête « les poilus », Google fournit, en première page de résultats, six liens vers des sites traitant le sujet, deux liens vers des sites sans rapport et huit liens commerciaux mis en évidence vers des sites de pornographie.

(19) Ainsi, la possibilité technique très directe (usage simple) donnée à l'utilisateur de laisser des traces (opinion, mots-clés) associées à tel ou tel contenu rencontre un très vif succès sur Internet, à l'instar des « Blogs ».

(20) E. Broudoux et ses collaborateurs (2005) précisent le concept : en utilisant le balisage (*tagging*) l'auteur (ou le lecteur) affecte des mots-clés à un contenu caractérisant de son point de vue l'information présente sur le *Web* (image, texte, lien). On obtient ainsi une classification « par centre d'intérêt » qui peut être destinée à rester dans la sphère privée ou à être ouverte au public. La *folksonomie* engendre donc une sorte de classement populaire (*folk*) de l'information, qui, à cause du fonctionnement interne des moteurs (indexation des mots-clés), finit par s'imposer par effet de nombre.

(21) Beaucoup d'entre eux disent utiliser des méta-moteurs comme « Copernic », mais quiconque sait comment ces automates fonctionnent peut montrer qu'ils génèrent encore plus de bruit, malgré une interface qui paraît plus lisse, ou « propre ».

(22) Cf. le site http://www.educasource.education.fr (consulté le 9 juin 2006).

(23) L'opération, décidée par les instances politiques locales sans concertation avec le terrain, consiste dans le prêt d'un ordinateur portable à tous les collégiens de 4e et de 3e des collèges publics et privés sous contrat du département des Bouches-du-Rhône. Cette machine est équipée d'un lecteur de CD-Rom, d'un modem et une connexion Internet haut-débit Wi-Fi est assurée dans toutes les classes des établissements. Dix heures par mois de connexion gratuite au Réseau téléphonique commuté (RCT), c'est-à-dire « fixe » au domicile sont fournies. Le parc d'ordinateurs portables prêtés par le Conseil général des Bouches-du-Rhône est de 60 000 unités, et, en 2006, trois autres Conseils généraux tentent de s'inspirer de ce dispositif, mais dans une plus faible mesure.

RÉFÉRENCES BIBLIOGRAPHIQUES ET SITOGRAPHIQUES

BARON G.-L. & BRUILLARD É. (1996). *L'informatique et ses usagers dans l'éducation*. Paris : PUF.

BARON G.-L. & BRUILLARD É. [dir.] (2002). *Les technologies en éducation : perspectives de recherche et questions vives : actes du symposium international francophone [tenu à] Paris, Maison des sciences de l'homme, les 31 janvier et 1er février 2002 / [organisé par l'] INRP, [l'] IUFM Basse-Normandie [et la] Fondation Maison des sciences de l'homme, Programme de numérisation pour l'enseignement et la recherche.* Paris : INRP.

BARON G.-L. & GLIKMAN Viviane [éd.] (1996). *Technologies nouvelles et éducation : quatre années de résultats à l'INRP : 1991-1994.* Paris : INRP.

BEAUFILS A. (1998). « Aide à l'exploitation de bases hypermédias ». In A. Tricot & J.-F. Rouet (éd.), *Les hypermédias, approches cognitives et ergonomiques*. Paris : Hermès, p. 191-209.

BEAUFILS A. (1999). « Aide à la recherche d'informations dans les environnements hypermédias ». In J.-F. Rouet & B. de La Passardière (éd.), *actes du quatrième Colloque Hypermédias et apprentissages, Poitiers, 15, 16 & 17 octobre 1998, université de Poitiers, Maison des sciences de l'homme et de la société.* Paris : INRP ; EPI, p. 73-86.

BÉLISLE C. & LINARD M. (1996). « Quelles nouvelles compétences des acteurs de la formation dans le contexte des TIC ? », *Éducation permanente*, n° 127, p. 19-47.

BEVORT E. & BRÉDA I. (2001). *Les jeunes et Internet : représentations, usage et appropriations*. Paris : Centre de liaison de l'enseignement et des moyens d'information [CLEMI], 2001. Disponible sur Internet : <http://www.clemi.org/jeunes_internet.html> (consulté le 9 avril 2007).

BLONDEL F.-M. (2001a). « Pratiques documentaires en sciences avec Internet, outils de recherche et compétences des élèves ». In W. Frindte, T. Köhler, P. Marquet & E. Nissen (éd.), *Internet-Based Teaching and Learning (IN-TELE) 99*. Frankfurt-am-Main : P. Lang, p. 73-78.

BLONDEL F.-M. (2001b). « La recherche d'informations sur Internet par des lycéens : analyse et assistance à l'apprentissage ». In E. de Vries, J.-P. Pernin & J.-P. Peyrin (éd.), *Hypermédias et Apprentissages : actes du cinquième Colloque Hypermédias et apprentissage*. Paris : INRP ; EPI, p. 119-133.

BLONDEL F.-M. ; LE TOUZE J.-C. & TARIZZO M. (2002). « ARI : un assistant logiciel pour accompagner la formation à la recherche ». In C. Frasson & J.-P. Pecuchet (éd.), *Actes du Colloque TICE2002.* Lyon : INSA 2002, p. 167-174. Document disponible au format PDF sur Internet à l'adresse : <http://www.inrp.fr/Tecne/Savoirplus/Rech40123/acdoc.htm> (consulté le 9 avril 2007)

BOULLIER D. & GHITALLA F. (2004). « Le Web ou l'utopie d'un espace documentaire ». *Information, interaction, intelligence : une revue en sciences du traitement de l'information*, vol. 4, n° 1, p. 173-189.

BRANDT-POMARES P. (2003). *Les nouvelles technologies de l'information et de la communication dans les enseignements technologiques : de l'organisation des savoirs aux conditions d'étude : didactique de la consultation d'information.* Thèse de doctorat : psychologie et sciences de l'éducation, université de Provence [Aix-Marseille 1].

BROUDOUX E. ; GRÉSILLAUD S. ; LE CROSNIER H. & LUX-POGODALLA V. (2005). « Construction de l'auteur autour de ses modes d'écriture et de publication ». *Actes de la conférence Hypermédias Hypertexts Products, Tools, Methods.* Document disponible au format PDF sur Internet à l'adresse : <http://archivesic.ccsd.cnrs.fr/docs/00/06/26/61/PDF/sic_00001552.pdf> (consulté le 9 avril 2007).

CASTELLS M. (2002). *La galaxie Internet*. Paris : Fayard.

CHEVALLARD Y. (1991). *La transposition didactique*. Grenoble : Éd. La pensée sauvage.

CENTRE DE LIAISON DE L'ENSEIGNEMENT ET DE L'INFORMATION (2003). *Les jeunes et Internet : synthèse internationale de la recherche*. Paris : CLEMI. Document auparavant disponible sur Internet : <http://www.clemi.org/recherche/jeunes_Internet/ji_international.rtf> (consulté le 14 mai 2005).

CLOT Y. (2003). « La catachrèse entre réel et réalisé : contribution d'un psychologue du travail ». In Y. Clot, & R. Gori (éd.), *Catachrèse : éloge du détournement*. Nancy : Presses universitaires de Nancy, p. 11-27.

COMPAINE B. (2001). *The Digital Divide : Facing a Crisis or Creating a Myth ?* Cambridge [Mass.] : MIT Press.

DARCOS X. (2003). « Lettre de mission pour le projet Proxima ». In France : ministère délégué à l'enseignement scolaire ; ministère délégué à la famille. *Mission Internet et Famille : le projet PROXIMA : pour une appropriation de l'Internet à l'école et dans les familles. Rapport remis de B. Benhamou à X. Darcos, ministre délégué à l'enseignement scolaire et à C. Jacob, ministre délégué à la famille.* Paris, 2003. Document disponible au format PDF sur Internet à l'adresse : <http://www.educnet.education.fr/plan/proxima.htm> (consulté le 9 avril 2007).

FRANCE : COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS (2001) : *Internet et la collecte de données auprès des mineurs. Rapport de la CNIL adopté le 12 juin 2001, présentée par C. Alvergnat.* Document disponible au format PDF sur Internet à l'adresse : <http://www.cnil.fr/fileadmin/documents/junior/droits/mineurs.pdf> (consulté le 9 avril 2007).

FRANCE : MINISTÈRE DÉLÉGUÉ À L'ENSEIGNEMENT SCOLAIRE ; MINISTÈRE DÉLÉGUÉ À LA FAMILLE (2003b). *Mission Internet et Famille : le projet PROXIMA : pour une appropriation de l'Internet à l'école et dans les familles. Rapport de B. Benhamou remis à X. Darcos, ministre délégué à l'enseignement scolaire et à C. Jacob, ministre délégué à la famille.* Paris, 2003. Document disponible au format PDF sur Internet à l'adresse : <http://www.educnet.education.fr/plan/proxima.htm> (consulté le 9 avril 2007).

GRANDBASTIEN M. (2002). « Quelques questions à propos de l'indexation et de la recherche de ressources pédagogiques sur le Web ». In G.-L. Baron & É. Bruillard (éd.), *Les technologies en éducation : perspectives de recherche et questions vives : actes du symposium international francophone [tenu à] Paris, Maison des sciences de l'homme, les 31 janvier et 1er février 2002 / [organisé par l'] INRP, [l'] IUFM Basse-Normandie [et la] Fondation Maison des sciences de l'homme, Programme de numérisation pour l'enseignement et la recherche.* Paris : INRP, p. 211-220.

GURAK L.-J. (2001). *Cyberliteracy : Navigating the Internet with Awareness*. New Haven : Yale University Press.

JOACHIMS T. ; GRANKA L. ; PAN B. ; HEMBROOKE H. & GAY G. « Accurately Interpreting Clickthrough Data as Implicit Feedback ». *SIGIR'05*, 15-19 août. Document disponible au format PDF sur Internet à l'adresse : <http://www.cs.cornell.edu/People/tj/publications/joachims_etal_05a.pdf> (consulté le 9 avril 2007).

JOHSUA S. (1998). « Des savoirs et de leur étude : vers un cadre de réflexion pour l'approche didactique ». *L'année de la recherche en education*, p. 79-97.

JOHSUA S. & FÉLIX C. (2002). « Le travail à la maison, un milieu pour l'étude ». *Revue française de pédagogie*, n° 141, p. 89-97.

LADAGE C. & RAVESTEIN J. (2005). « Apprentissage et enseignement sur des espaces de communication informels sur Internet : l'exemple du "référencement" ». In P. Tchounikine, M. Joab & L. Trouche (éd.), *Environnements informatiques pour l'apprentissage humain [EIAH] : actes de la conférence EIAH 2005, Montpellier : 25, 26 et 27 mai 2005.* Lyon : INRP ; Montpellier : université Montpellier 2, p. 371-376. Document disponible au format PDF sur Internet à l'adresse : <http://hal.archives-ouvertes.fr/docs/00/03/19/20/PDF/ac7.pdf > (consulté le 9 avril 2007).

LAVAL C. (2003). *L'école n'est pas une entreprise : le néo-libéralisme à l'assaut de l'enseignement public.* Paris : La Découverte.

LE ROUX L. (2004). *Des "repères" et des "liens" : l'enseignant-documentaliste, la recherche d'information sur Internet ; répercutions sur la formation des élèves*. Document au format PDF disponible sur Internet à l'adresse : <http://archivesic.ccsd.cnrs.fr/docs/00/06/25/44/PDF/sic_00001277.pdf> (consulté le 9 avril 2007).

LELONG B. (2002). « Savoir-faire technique et lien social : l'apprentissage d'Internet comme incorporation et autonomisation ». In C. Chauviré & A. Ogien (éd.), *La régularité : habitude, disposition et savoir-faire dans l'explication de l'action*. Paris : Éd. de l'EHESS, p. 267-292.

LEONTIEV A.-N. (1984). *Activité, conscience, personnalité*. Moscou : Éd. du Progrès.

LINARD M. (1990). *Des machines et des hommes : apprendre avec les nouvelles technologies.* Paris : Éd. Universitaires.

LINARD M. (1995). « La distance en formation : une occasion de repenser l'acte d'apprendre ». In G. Davies & D. Tinsley (éd.), *Open and Distance Learning : Critical Success Factors* [Accès à la formation à distance : clés pour un développement durable]. *Proceedings. International Conférence, Geneva 10-12 October.* Erlangen : FIM, p. 46-55.

LINARD M. (2001). « Concevoir des environnements pour apprendre : l'interaction humaine, cadre de référence ». *Sciences et techniques éducatives*, vol. 8, n° 3-4, p. 22-31.

MACEDO-ROUET M. ; ROUET J.-F. ; ZAMPA V. & BOUIN É. (2006). « Jugement de la qualité de documents Web : rôle de l'expertise et de la tâche ». In A. Piolat (éd.), *Lire, écrire, communiquer et Apprendre avec Internet*. Marseille : Solal, p. 463-488.

MARTINAND J.-L. (1983). « Questions pour la recherche : la référence et le possible dans les activités scientifiques scolaires ». *Actes de l'atelier international de recherche en didactique de la physique*. Paris : Publications du CNRS, p. 227-250.

MERCIER, A. ; SCHUBAUER-LÉONI M.-L. & SENSEVY G. [coord.] (2002), *Revue française de pédagogie*, n° 141 : « Vers une didactique comparée ».

PROULX S. (2001) « Usages de l'Internet : la pensée-réseaux et l'appropriation d'une culture numérique ». In E. Guichard (dir.), *Comprendre les usages de l'Internet*. Cachan : Éd. de l'ENS, p. 171-180.

RAVESTEIN J. & LADAGE C. (2006). « Pour une didactique de l'usage d'Internet ». In A. Piolat (éd.), *Lire, écrire, communiquer et apprendre avec Internet*. Marseille : Solal, p. 489-508.

ROUET J.-F. & LA PASSARDIERE Brigitte de [éd.] (1999). *Hypermédias et apprentissage : actes du quatrième Colloque Hypermédias et apprentissages, Poitiers, 15, 16 & 17 octobre 1998, université de Poitiers, Maison des sciences de l'homme et de la société.* Paris : EPI ; INRP.

ROUET J.-F. & TRICOT A. (1998). « Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs ». In A. Tricot & J.-F. Rouet (éd.), *Les hypermédias, approches cognitives et ergonomiques*. Paris : Hermès, p. 57-74.

SOËTARD J. (2000). *Internet, dangereux pour les enfants ? Étude Ipsos-Libération-Powow.net, en partenariat avec La Cinquième*. Disponible sur Internet : <http://www.ipsos.fr/CanalIpsos/articles/98.asp> (consulté le 9 avril 2007).

VERONIS J. (2005). « Web : Le mystère des pages manquantes de Google résolu ». *Technologies du langage : actualités, commentaires, réflexions*. Document disponible sur Internet : <http://aixtal.blogspot.com/2005/02/web-le-mystre-des-pages-manquantes-de.html> (consulté le 9 avril 2007).

WARSCHAUER Mark (2003). *Technology and Social Inclusion : Rethinking the Digital Divide.* Cambridge [Mass.] : MIT Press.