

Les pratiques de recherche d'informations sur Internet

Alain Beaufiles, François-Marie Blondel

TECHNOLOGIES NOUVELLES ET ÉDUCATION
INRP

Quelles difficultés les élèves éprouvent-ils lors de leurs recherches ? Quelles compétences devront-elles être développées ?

La Toile apparaît à de nombreux utilisateurs comme une source d'informations importante en volume, variable en qualité, dont les contenus sont très dispersés, mais qui présente l'immense avantage d'offrir un accès direct à des documents souvent absents des centres de documentation. En ce sens, elle peut être considérée comme un complément majeur de la documentation traditionnelle.

Internet, un fonds documentaire différent

L'évolution des contenus de la Toile est très rapide et encore très dépendante des choix et des orientations prises par ceux qui ont la capacité de mettre en ligne des ressources – à savoir, les institutions, les entreprises et les personnes privées. On peut d'ailleurs se demander si un Internet libre, gratuit et coopératif sur le modèle académique de ses origines aura encore sa place dans les années à venir quand on sait que d'ores et déjà 80 % du contenu de la Toile a une origine commerciale.

Bien qu'il soit encore difficile de se faire une idée précise de ce qui est utile ou intéressant pour des élèves, nombreux sont ceux qui font l'hypothèse que les informations présentes sur la Toile pourront d'une manière ou d'une autre aider les élèves dans leurs activités documentaires. Or, ces activités connaissent un regain d'intérêt avec les nouveaux programmes de lycée et notamment la mise en place des TPE¹. Il devient par conséquent plus urgent de s'interroger sur les conditions d'exercice de ces activités et en particulier sur la manière dont les élèves les pratiquent.

– Comment les élèves vont-ils rechercher dans ce vaste ensemble bien peu structuré ? Y a-t-il

des méthodes de recherche générales applicables à tous les domaines de la connaissance ou faut-il adopter des méthodes plus spécifiques ?

– Comment les élèves mènent-ils leurs recherches ? Quelles sont les difficultés habituellement constatées ? Ces difficultés sont-elles générales à tout type de recherche documentaire ou spécifiques à Internet ?

– Comment aider les élèves dans cette activité ? Les consignes et les outils utilisés pour la recherche dans les bases documentaires fermées peuvent-ils être adaptés à cet ensemble plus vaste et plus complexe que constitue Internet ? Comment faciliter l'accès aux informations ? Quelles sont les compétences sous-jacentes à cette activité ?

Observer les élèves pour les aider à mener une recherche d'information sur la Toile

Dans la lignée de travaux portant sur les usages des hypermédias et la recherche d'information dans les textes², nous nous intéressons depuis plusieurs années aux problèmes de l'activité documentaire dans le secteur scolaire. Nous avons cherché à décrire les difficultés que rencontrent des élèves de différents niveaux lorsqu'ils recherchent des informations dans des bases documentaires complexes, notamment les bases hypermédias. Pour certaines de ces difficultés, nous avons proposé quelques solutions instrumentales ou méthodologiques³. Pour tenter de répondre aux questions que pose la recherche sur Internet, nous avons effectué plusieurs observations détaillées dans des situations de recherche encadrée, au collège et au lycée.

Dans une partie de ces observations, les élèves de lycée ont traité des questions précises, relatives à la *production d'énergie*⁴. Ils ont travaillé pendant deux séances d'une heure et demie chacune en utilisant moteurs et annuaires, après une rapide présentation de la question et des outils de recherche sur la Toile. Ils ont ensuite rédigé chez eux un document de synthèse. Leur activité de recherche a été observée par un logi-

1. Voir à ce propos les nombreuses références sur le site TPE-TICE : www.educnet.education.fr/TPE/

2. ROUET J.-F. et TRICOT A., 1998.

3. BEAUFILS, A., 1998, 1999.

4. BLONDEL François-Marie, KEMPF Olivier, SCHWOB Monique, 2000.

ciel qui mémorise les principaux événements de la navigation, en particulier les appels aux moteurs et les visites de pages.

Dans d'autres observations effectuées au collège ou au lycée sur de petits groupes constitués de volontaires, les élèves ont traité de questions dont le sujet est plus large, comme *la violence des adultes envers les enfants dans la littérature, la chanson et le cinéma ; le statut et l'image de la femme dans la société d'aujourd'hui* ou *les effets et les dangers du dopage dans le sport*. Pour ce faire, les élèves ont utilisé le métamoteur Copernic, leur travail étant guidé par une méthode comportant des modèles et des consignes. Au terme de quatre à six sessions hebdomadaires d'une heure, ils ont produit une synthèse personnelle de leur travail. À la fin de chaque session, les traces de l'activité de chaque élève étaient recueillies : le résultat des recherches effectuées sur Internet (archives du logiciel Copernic), les références des documents visités (historique de la navigation) et l'état de la production écrite (fichier individuel de prise de notes).

Dans tous les cas, les élèves ont effectué leurs recherches très librement, sans que des sites leur soient suggérés ni des contraintes imposées, à l'exception de la contrainte horaire, largement compatible avec l'objectif de recherche fixé.

Présentés ci-dessous, les premiers constats issus de ces observations sont partiels et probablement discutables ; ils demanderaient à être confirmés par des études plus précises et plus approfondies sur les conditions qui affectent l'utilisation des outils, sur l'effet d'une formation ou d'un apprentissage initial, sur le niveau de connaissances dans le domaine exploré.

La capacité des élèves à trouver est indéniable

Les élèves trouvent des documents, et, dans ces documents, ce qu'ils estiment être des réponses à la question qui leur a été posée.

Ainsi, au cours de la recherche sur les filières de production électrique, les lycéens de classe de Première ont dans leur grande majorité trouvé des informations sur le principe du fonctionnement de ces filières, sur les schémas associés, et sur les puissances et les rendements correspondants. Cependant, en examinant ce qu'ils ont trouvé, c'est-à-dire ce qu'ils ont utilisé dans le document final rendu à leur enseignant, on se rend compte que *l'efficacité* de leur recherche est assez faible. En effet, le nombre de documents utilisés par rapport au nombre de sites consultés est en moyenne dans un rapport de un à cinq. Autrement dit, sur l'ensemble des pages (une

trentaine) appartenant à la douzaine des sites qu'ils ont consultés, en moyenne, ces lycéens n'ont retenu que deux documents.

Sur une question plus générale comme *la violence des adultes envers les enfants* posée à des élèves de Quatrième, la recherche a été plus incertaine. Chaque élève a formulé en moyenne douze requêtes pour lesquelles le métamoteur a renvoyé plus de deux cents documents. Une trentaine d'entre eux ont été sélectionnés puis explorés, et sur cette trentaine, cinq documents seulement ont été exploités.

Cette déperdition suggère plusieurs explications dont la plus évidente est probablement un manque de satisfaction des utilisateurs, opinion souvent exprimée lors des entretiens ou des observations directes.

Mais savent-ils chercher ?

En effet, il apparaît que les résultats obtenus à la suite d'une requête adressée à un moteur ou un métamoteur ne correspondent pas toujours aux attentes des élèves. Ce qui dénote une maîtrise pour le moins imparfaite des techniques de recherche sur la Toile.

Quelques indices recueillis dans les traces de navigation contribuent à formuler ce constat.

Tout d'abord, une analyse du temps consacré à chaque phase de la recherche pour la classe de Première travaillant sur l'énergie, permet de constater que si les lycéens consacrent plus de la moitié de leur temps de recherche à la consultation des sites, la formulation des requêtes et la consultation des résultats rendus par les moteurs ou les annuaires occupent un peu plus du tiers de leur temps. Le reste correspond à des temps morts intégrant les délais de réception des documents, délais peu importants au demeurant (voir Figure 1).

En second lieu, une analyse des requêtes formulées permet de constater que le choix des termes et leur combinaison avec des opérateurs pour formuler une requête efficace, et par conséquent susceptible de produire des résultats correspondants à leurs attentes, restent encore très approximatifs et laborieux. Sur une cinquantaine de requêtes, 80 % comprennent moins de trois termes (voir Figure 2). Quant au lexique utilisé, il se limite souvent aux mots ou aux termes de la question posée.

L'insatisfaction des élèves quant aux résultats fournis par les moteurs est assez fréquente. Comme le montre la Figure 3, la moitié des requêtes sont infructueuses, en ce sens qu'elles ne sont pas suivies de la consultation d'au moins un site appartenant à la liste des résultats affichés.

5. DINET, J., ROUET, J.-F. & PASSERAULT, J.-M., 1999.
6. BOURDONCLE F., 1999.

Comment identifier les références pertinentes ?

Il faut reconnaître que la simple consultation des listes de résultats fournies par les moteurs ne donne qu'une idée approximative du contenu réel des documents référencés. On y observe couramment de nombreuses erreurs ou confusions dans l'appréciation de la pertinence des références trouvées⁵, ce qui peut s'expliquer par l'opacité des modèles de classement adoptés par les moteurs⁶. Sans assistance ou sans dispositif adéquat, les élèves ne peuvent utiliser de manière rationnelle les listes de résultats que leur proposent les moteurs. Ainsi, nos observations des emplois de Copernic montrent qu'un tiers des documents sélectionnés, en moyenne, ne sont pas consultés.

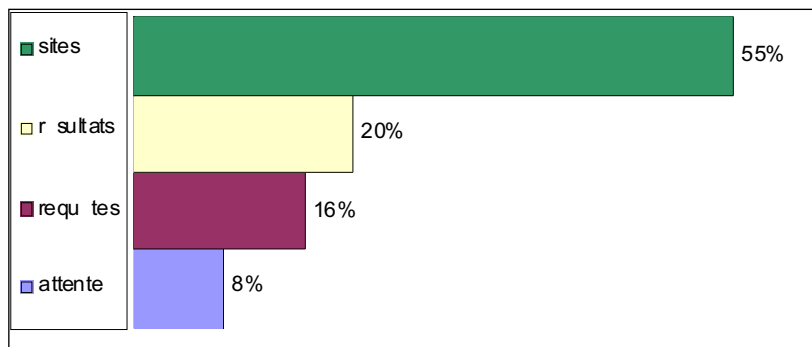
La question se pose aussi de savoir si ce qu'ils consultent est toujours pertinent. La réponse semble dépendre du sujet et du domaine dans lequel s'opère la recherche. En reprenant l'exemple de la recherche sur les filières énergétiques, nous avons observé que les élèves ont visité essentiellement des sites appartenant soit à des institutions comme le ministère de l'Environnement du Canada, soit à des établissements de formation comme des universités, soit à des entreprises du secteur. Sur un sujet précis comme celui-ci, la pertinence des sites visités peut être considérée comme très bonne. Sur la question de la violence, en revanche, le taux de documents pertinents est d'environ 30 % pour ceux sélectionnés sur références et explorés et de 80 % pour ceux exploités dans la synthèse finale.

Choisir et maîtriser des stratégies de recherche

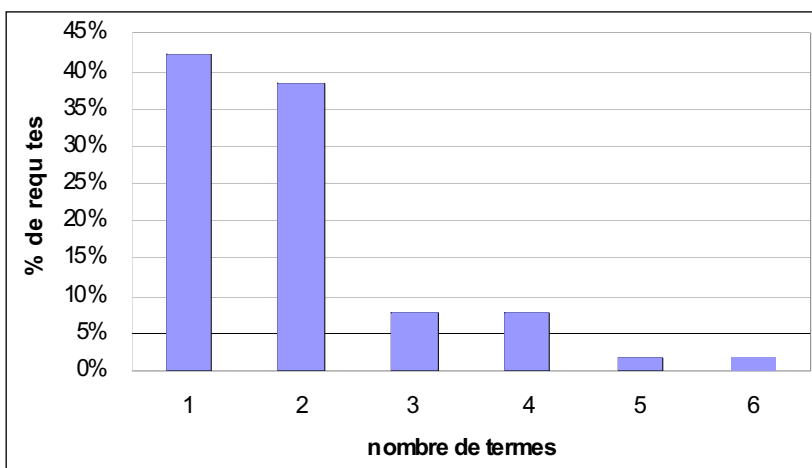
La plupart des élèves rencontrent à peu près les mêmes difficultés au cours des étapes de leur travail. Certaines sont liées à la recherche d'informations en général, quel que soit l'outil utilisé, d'autres sont propres à la manière de conduire cette activité sur Internet. La mise en œuvre de stratégies de recherche adaptées et efficaces demande que certaines conditions soient remplies, ce qui est loin d'être toujours le cas.

1. Identifier la nature des informations à trouver

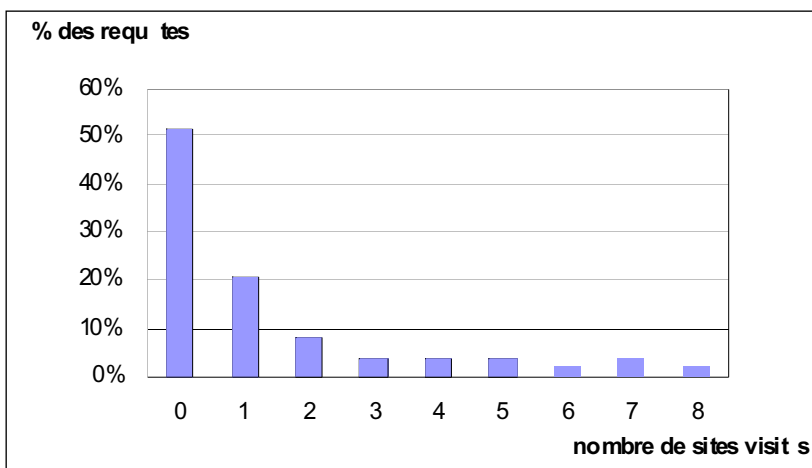
Cela implique d'une part de bien comprendre la question posée, d'autre part de disposer de quelques connaissances initiales sur le domaine concerné. Généralement, le temps consacré par les élèves à la préparation d'une recherche d'informations est plutôt bref, voire inexistant, si on ne leur impose pas de consignes précises qui les



1. Répartition du temps consacré à la consultation de sites, à la lecture de résultats de moteurs, à la formulation de requêtes et à l'attente de la réponse des serveurs, temps exprimé en % de la durée totale de la navigation (classe de première S, recherche sur l'énergie, avril 1999)



2. Répartition du nombre de termes par requête, effectif 52 requêtes (classe de première S, recherche sur l'énergie, avril 1999)



3. Répartition du nombre de sites visités par requête, effectif 52 requêtes (classe de première S, recherche sur l'énergie, avril 1999)

incitent à effectuer quelques travaux préliminaires : noter les indices contenus dans la question, élargir le champ sémantique au-delà des mots de l'énoncé, éventuellement construire un plan provisoire de la réponse. Toutefois, même s'ils y consacrent un certain temps, les élèves ne peuvent pas évaluer correctement le besoin d'informations s'ils n'ont aucune connaissance préalable du sujet qu'ils doivent traiter.

2. Se faire une représentation précise du fonds documentaire à explorer

Ce n'est pas possible avec Internet dont aucune description générale de contenu, telle que table des matières ou index, n'est directement consultable. Les annuaires qui offrent un classement hiérarchique d'une partie du contenu de la Toile ne semblent pas compris comme tels par les élèves. En règle générale, il apparaît que les élèves n'ont qu'une connaissance très approximative du contenu de la Toile et, par conséquent, font preuve d'une assez grande naïveté sur ce qu'on peut y rechercher. Ils se comportent souvent comme si la réponse à leur question *devait* se trouver dans un document ou un autre. D'où une obstination quelquefois excessive à rechercher en variant les mots de la question, en variant

« Les élèves se comportent souvent comme si la réponse à leur question devait se trouver dans un document ou un autre.. »

les moteurs appelés. Mais il faut noter aussi que, dans le cadre d'une activité imposée, ils sont en droit de faire l'hypothèse – celle du contrat didactique habituel – que la réponse à la question posée doit pouvoir être trouvée avec les méthodes proposées par l'enseignant.

3. Être capable d'utiliser des outils de recherche adaptés

Sur Internet, l'utilisation raisonnée de moteurs de recherche demande souvent des compétences d'experts que n'ont pas les élèves. Ceux-ci ont du mal à comprendre comment fonctionnent ces outils et ne parviennent pas à articuler correctement les termes d'une requête pour qu'elle soit efficace (choix et nombre de termes, mots généraux ou précis, utilisation d'opérateurs comme *et, ou, sauf*, etc.). Cet exercice est d'autant plus difficile que les outils de recherche qui prétendent s'adresser à un large public mettent en œuvre des techniques complexes que seuls des experts connaissent.

Des guides et des aides à inventer

La complexité de la tâche de recherche d'information incite à proposer des méthodes et des guides de travail qui, en réduisant les aléas, aideront l'élève à centrer son activité sur des aspects essentiels. Quelques guides de recherche proposés aux élèves ont été publiés à ce jour⁷. Il est possible que ces guides soient trop généraux pour être appliqués tels quels et que leur mise en pratique impose des conditions qui ne sont pas toujours respectées, notamment en matière de compétences d'élèves.

Des guides ou des aides plus spécifiques du domaine de recherche ou plus orientés vers des outils particuliers, comme ceux que nous avons commencé à produire pour une observation détaillée⁸, peuvent présenter un intérêt pour développer les modes d'usage des outils. Bien que les aides de ce type proposées aux élèves soient souvent peu utilisées, un document qui serve à la fois de guide et d'exemple de recherches peut faciliter la tâche de l'élève tout en l'encourageant à certaines pratiques jugées plus efficaces. Et, dans les situations de travail moins encadrées, un guide de référence peut servir à l'utilisateur pour relancer son travail.

Dans le domaine des outils logiciels, les assistants prennent en charge une partie de la tâche ou, du moins, en simplifient l'exécution. Certains métamoteurs comme Copernic présentent l'avantage de conserver non seulement les résultats de la recherche, mais aussi le contenu des pages dont l'utilisateur aura demandé le téléchargement. Cette fonctionnalité permet ensuite de « raffiner » la recherche sur un sous-ensemble des documents recueillis, avec des outils de recherche en texte intégral. La technique est alors proche de ce que permettent certains CD-Rom ou logiciels de recherche documentaire.

Nous travaillons actuellement à un assistant de recherche d'information construit autour d'un historique structuré et d'un système de prise de notes qui ne masque pas les étapes de la recherche mais, au contraire, aide à les organiser.

Des compétences spécifiques?

Si l'on estime indispensables les activités de recherche documentaire sur la Toile, et ce point mérite une discussion sérieuse, alors un apprentissage est certainement nécessaire. À l'évidence, les connaissances et les savoir-faire acquis en matière de documentation au cours de la scolarité antérieure ne semblent pas suffisants pour permettre un usage raisonné et efficace des outils pour une recherche libre sur Internet. Ces connaissances et savoir-faire adaptés à la recherche sur les supports traditionnels sont-ils d'ailleurs transférables à la recherche sur Internet?

Sans revenir sur les compétences générales en matière de documentation, on peut esquisser une première liste de compétences induites par ces activités de recherche d'information sur la Toile. Comme nous avons pu en juger d'après les pratiques d'élèves et en analysant ce qui est publié sur la recherche d'information pratiquée par des experts et par des novices, il apparaît que cette activité met en jeu des savoir-faire particuliers et des connaissances spécifiques (voir encadré).

7. Voir *Formanet guide général* « Les étapes pour chercher et exploiter l'information » www.ac-poitiers.fr/pedago/coll_lyc/formanet/guides/AIDE_WEB/index.htm.

Ou un guide plus spécialisé comme « Recherche d'information et communication sur Internet » rédigé par Hélène Guertin, www.fas.umontreal.ca/ebsi/jetrouve/internet/internet.htm.

Ou encore la méthode proposée par Maryvonne Courte-cuisse dans le *Dossier de l'ingénierie éducative* n° 29.

8. Voir l'exemple de guide sur le site ADIPC www.inrp.fr/Tecne/adipc/aides/aide-eleve/aide-eleve.htm.

9. Voir des référentiels de compétences rédigés par des professionnels de la documentation comme ceux qui sont présentés sur le site de Formanet, www.fas.umontreal.ca/ebsi/formanet/maitrise.html#Modeles, ou encore le *Dossier de l'ingénierie éducative* n° 29.

Compétences des élèves

Savoir-faire

- Naviguer : utiliser plusieurs fenêtres, gérer ses signets, utiliser l'historique, copier un document
- Rechercher dans un catalogue ou une base de données bibliographique ou factuelle
- Rechercher dans un annuaire, par la table de classement ou par mots-clés
- Rechercher avec un moteur en employant les options de recherche
- Exploiter et structurer une liste de références
- Analyser un document : structure, origine, ancienneté...

Savoirs

- Le vocabulaire de base : serveur, site, page, document, répertoire, annuaire, moteur...
- La construction d'une adresse
- La structure d'un document
- Les caractéristiques principales d'un annuaire, d'un moteur et d'une base de données
- Les possibilités d'interrogation
- Les types de documents présents sur la Toile : ouvrages, articles de revues scientifiques, articles de vulgarisation, rapports, thèses, cours, documents personnels, documents éducatifs, documents commerciaux...

Références

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

BEAUFILS, A., « Aide à la recherche d'informations dans les environnements hypermédias » in J.-F. Rouet & B. de La Passardière (Eds), Hypermédias et apprentissages 4, INRP, EPI, 1999, p. 73-86.

BLONDEL François-Marie, KEMPF Olivier, SCHWOB Monique (2000).- Recherche d'informations sur la toile : Pratiques d'élèves, Bulletin de l'Union des physiciens, n° 828, novembre 2000, pp. 1819-1846.

BOURDONCLE F., Panorama et perspectives des outils de recherche d'Information textuelle sur Internet, Actes du colloque IDT'99 (10 juin 1999). www.cma.ensmp.fr/Francois.Bourdoncle/idt99.html

DINET, J., ROUET, J.-F. & PASSERAULT, J.-M., « Les nouveaux outils de recherche documentaires sont-ils compatibles avec les stratégies cognitives des élèves? » in J.-F. Rouet & B. de La Passardière (Eds), Hypermédias et apprentissages 4, INRP, EPI, 1999, p. 149-161.

GOFFARD Monique, Les activités de documentation en physique et chimie, Armand Colin, 1998, 119 p.

ROUET J.-F. et TRICOT A., Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs, in Les hypermédias, approches cognitives et ergonomiques, Tricot André & Rouet Jean-François (eds), Hermès, Paris 1998, pp. 57-74.

Encadrer des élèves pour qu'ils recherchent efficacement sur la Toile suppose aussi des compétences propres de la part des enseignants. Risquons-nous ici à en mentionner quelques-unes.

Au-delà des connaissances indispensables pour guider les élèves en matière de méthodes et techniques documentaires, encadrer des élèves suppose une connaissance des sources et des documents disponibles sur la Toile, une connaissance des thèmes, des sujets et des mots-clés efficaces (quels sont les bons et les mauvais sujets de recherche dans un domaine donné?) et des connaissances sur les techniques à privilégier, en fonction du degré de maîtrise des outils et du degré de connaissance des élèves.

Loin de tout maximalisme, nous pensons qu'il ne faut pas sous-estimer le poids des connaissances spécifiques que ces activités impliquent, et que seule une formation *ad hoc* est en mesure de fournir.

Comprendre ce que l'on fait, travailler avec rigueur, accroître son efficacité n'empêchent pas de se laisser surprendre par des documents inédits ou des informations inattendues. Rechercher sur la Toile reste une activité passionnante et motivante. ●

