



REPERE

Ressources Electroniques
Pour les Etudiants,
la Recherche et l'Enseignement

<http://reperre.enssib.fr>

2011

La brochure REPERE 2011 a été révisée, essentiellement dans les liens, à jour au 1^{er} juin 2011.

La brochure REPERE est désormais uniquement disponible sous forme électronique, et peut être librement consultée, téléchargée et diffusée sous sa forme électronique (pdf ou via l'application ISSUU) disponible sur ce site : <http://reper.enssib.fr>

Nous vous souhaitons une bonne lecture.

Sommaire

• Avant-propos	05
----------------------	----

Le contexte

• Besoins – Les besoins d'information	07
• Compétences – Des compétences documentaires à la culture de l'information	11
• Évolutions – Web collaboratif et gestion de l'identité numérique	15

L'offre

• Acteurs – Acteurs, intermédiaires et médiateurs	19
• Historique – La recherche des ressources électroniques	25
• Économique – L'économie des ressources électroniques	29
• Techniques – Accès aux ressources électroniques et à l'Internet	33
• Législation – Création et utilisation de ressources électroniques	39
• Typologie – Typologie des informations scientifiques	45
• Supports – Les supports de l'information	51

La recherche

• Méthode – Savoir rechercher et interroger	55
• Langages – Les langages documentaires	61
• Évaluation – Évaluer l'information sur Internet	65

L'exploitation

• Plagiat – Lutter contre le plagiat	69
• Gestion – Les logiciels de gestion de références bibliographiques	73
• Veille – La veille informationnelle	77
• Références – Rédaction et citation des références bibliographiques	81
• Favoris – Gestion des signets, favoris ou bookmarks	85
• Publier – Publier sur le Web	91

Les outils

• Généralistes – Les outils de recherche d'information sur le Web	95
• Spécialisés – Les outils de recherche sur Internet spécifiques au domaine scientifique	99
• Multimédia – La recherche des documents multimédia sur Internet	105
• Les adresses	111

Avant-propos

La mode est aux réseaux sociaux, au web collaboratif ; nous imaginons que nos étudiants, en véritables « natifs numériques » (*digital natives*) se retrouvent tous sur Facebook ou Myspace. Méfions-nous de ces *a priori* : l'utilisateur moyen de ces outils a 28 ans¹ !

Certes, ces nouveaux outils du web 2.0 modifient la production de l'information, une information qui est devenue plutôt une conversation. Certes, la recherche d'information se modifie, les outils collaboratifs comme les outils de partage de signets permettent une sérendipité accrue. Mais nos étudiants n'ont pas conscience de la richesse d'utilisation possible de ces outils, qu'ils s'approprient à titre personnel, sur lesquels ils exposent leur vie privée.

Les professionnels de l'information suivent ces évolutions, parfois même les devancent. La formation des usagers aborde souvent l'identité numérique, et l'usage de ces nouveaux outils de web 2.0. Certains établissements documentaires, comme le SCD de Lyon¹ ou la médiathèque de l'ESC de Lille, ont mis en place des univers Netvibes thématiques². Mais les étudiants, finalement, continuent de venir chercher leurs livres en bibliothèques et y travailler, et n'imaginent pas toujours que celle-ci puisse aller au-delà dans son offre de service.

Aussi nous faut-il trouver le juste équilibre entre ce qui sera utile réellement à l'étudiant et ce qui est encore en voie de développement. Nous devons veiller sur les évolutions de ces outils, pour imaginer ce que cela nous apportera, en terme de recherche dans les sources électroniques ou d'offre de services.

À l'heure où le réseau social Facebook explose et dépasse même Google en audience, où il est question de droit à l'oubli numérique, nous devons maintenir nos actions de formations. Les Unités Régionales de Formation à l'Information Scientifique et Technique restent, avec les Services Communs de la Documentation, des acteurs incontournables de cette éducation à l'information. Nous espérons que cette version de REPERE, mise à jour essentiellement dans ses ressources, vous sera utile.

Élisabeth NOËL

Conservateur des bibliothèques

École nationale supérieure des sciences de l'information et des bibliothèques

Directrice de publication de REPERE

1. SocioGeek. *Résultats* [en ligne]. < <http://sociogeek.admin-mag.com/resultat/Echantillon.html> >

2. Bibliopedia. *Bibliothèques sur netvibes* [en ligne].

< http://www.bibliopedia.fr/index.php/Bibliothèques_sur_netvibes >

Les besoins d'information

Auteur :

Olivier LE DEUFF

Certifié en documentation – Lycée professionnel De Vinci, Mayenne.

Le guide des égarés : <http://www.guidedesegares.info/>

- 08** Définir le besoin d'information
- 08** De la prise de conscience du besoin d'information à la recherche
- 09** Les nouvelles sources d'information



Le besoin d'information – besoin qui ne s'exprime pas toujours de manière évidente – est, par essence, de nature personnelle et diffère donc selon les individus. Même si son existence peut être initiée par un enseignant auprès d'une centaine d'étudiants, chacun possède une connaissance du sujet fort différente selon ses connaissances, compétences et autres vécus personnels.

Définir le besoin d'information

Parfois, la conscience d'un besoin d'information n'existe tout simplement pas, ce qui explique que des personnes puissent se retrouver en manque de connaissances, d'information sur un sujet sans en avoir conscience, soit par manque de connaissances préalables, soit par le fait d'une trop grande certitude. Chez les plus jeunes notamment, cette inconscience du besoin d'information laisse place à un besoin d'affirmation.

Le besoin d'information apparaît alors comme une véritable **compétence documentaire** à acquérir, compétence qui n'est pas nécessairement naturelle,

comme l'indique la définition donnée par l'*American Library Association* : « Être compétent dans l'usage de l'information signifie que l'on sait reconnaître quand émerge un besoin d'information et que l'on est capable de trouver l'information adéquate, ainsi que de l'évaluer et de l'exploiter ».

Cette capacité à prendre conscience de son besoin d'information implique ainsi la volonté de vouloir résoudre une incertitude, une insatisfaction par rapport à sa connaissance d'un sujet, ce que montre parfaitement la définition d'André Tricot : « Avoir besoin d'information implique que l'on ait de l'incertitude et donc des connaissances ».

De la prise de conscience du besoin d'information à la recherche

La suite logique de la prise de conscience du besoin d'information consiste à satisfaire ce besoin par le biais de la

recherche d'information. Cette dernière a considérablement évolué et ne cesse de se transformer, ainsi que les ressources et les



• LE COADIC Yves-François. *Le besoin d'information : formulation, négociation, diagnostic*. Paris : ADBS, 2007 (Sciences et techniques de l'information).

pratiques documentaires, depuis l'apparition des premières bases de données en ligne dans les années 1970. Du fait de la complexité des modes d'interrogation et des coûts engendrés, ces premiers outils étaient avant tout l'apanage des professionnels de l'information, bibliothécaires et documentalistes. Puis le développement des cédéroms a permis une plus large diffusion de contenus multimédias, directement utilisables par les usagers des campus (étudiants, enseignants, chercheurs), même si des formations spécifiques restaient nécessaires pour permettre aux usagers de se servir efficacement de ces outils.

L'émergence d'Internet a bouleversé encore plus fortement les pratiques documentaires, avec la mise à disposition, en ligne, d'un nombre croissant de ressources. Le recours à une médiation par les bibliothécaires et les documentalistes diminuant alors, la nécessité d'une formation à la recherche de l'information est devenue de plus en plus prégnante.

En faisant du lecteur un fournisseur d'information et un auteur potentiel, Internet provoque des changements profonds dans

la recherche et l'exploitation de l'information, ainsi qu'une grande évolution dans les pratiques intellectuelles. La recherche d'information constitue maintenant un apprentissage permanent.

La recherche d'information suppose ainsi l'utilisation de méthodes spécifiques à l'interrogation des bases de données ou des moteurs de recherche, avec l'usage d'opérateurs booléens associés à un choix pertinent de mots-clés. La connaissance de la diversité des outils disponibles et de leur fonctionnement constitue un atout pour une recherche d'information réussie.

Néanmoins, la complexité des outils et de l'Internet contraste avec l'apparente facilité d'usage des moteurs de recherche. Pourtant, ces derniers, dont les possibilités ne cessent de s'accroître au point de devenir de véritables médias à part entière, n'indexent pas la totalité des données présentes sur le Web. Il est nécessaire d'être conscient de l'existence de ce qui est nommé le **Web invisible** – l'ensemble des données non prises en compte par les moteurs de recherche – dès lors que la recherche d'information menée se veut exhaustive.

Les nouvelles sources d'information

Internet facilite grandement la possibilité de production directe d'information. La diversité des ressources disponibles s'en trouve accrue, depuis le simple avis posté sur un forum, à l'article d'un blog consacré à un sujet très précis, en passant par les banques de photos ou d'images. La chaîne éditoriale traditionnelle (sélection avant publication) n'opère plus, même si certaines ressources peuvent être de qualité (**blogs de spécialistes par exemple**). Un travail d'évaluation de l'information incombe dès lors au chercheur d'information.

La publication scientifique évolue également avec le développement de systèmes d'archives ouvertes, permettant aux chercheurs de mettre en accès libre leurs publications sur des serveurs dédiés. De plus en plus d'articles sont ainsi disponibles gratuitement. Toutefois un grand nombre d'articles restent encore diffusés par les éditeurs scientifiques traditionnels, une grande partie d'entre eux pouvant être consultés en ligne grâce aux abonnements souscrits par les bibliothèques universitaires.

Enfin, l'utilisation de fils RSS permet de mener une recherche d'information liée

à un besoin constant et ainsi d'effectuer une veille informationnelle. Ce travail peut aussi être facilité par les sites de partage de signets (favoris) en ligne, où il est possible d'effectuer la veille de manière collaborative. La simplification des outils de publication (plate-forme de blogs, réseaux sociaux, microblogging du style twitter), ainsi que les possibilités de tra-

vail collaboratif, avec les wikis permettant de travailler à plusieurs sur un même document, entraînent un passage de plus en plus fréquent de la recherche d'information à la production d'information. Au final, l'identification des besoins de communication devient ainsi aussi importante.



- AMERICAN LIBRARY ASSOCIATION
Presidential Committee on Information Literacy : Final Report, 1989 [en ligne].
<http://www.ala.org/ala/mgrps/divs/acrl/publications/whitepapers/presidential.cfm>
- MANUE. *Bibliothécaire et blogeuse...* Biblioacid, octobre 2005 [en ligne].
http://www.nicolasmorin.com/BiblioAcid_revue/BAv2n3.pdf
- TRICOT André. *La prise de conscience du besoin d'information : une compétence documentaire ?* Les interviews d'E-Doc, 2004 [en ligne].
<http://docsdocs.free.fr/spip.php?article70>

Des compétences documentaires à la culture de l'information

Auteur :

Claire DENECKER

Conservateur des bibliothèques – URFIST de Lyon

- 12** Les compétences documentaires
- 13** Les apprentissages
- 14** Information, culture, savoir



Trouver des documents pour compléter un cours, exploiter les ressources disponibles sur les Espaces Numériques de Travail (ENT), rechercher et évaluer l'information, élaborer une bibliographie pertinente, citer correctement les textes consultés... autant d'usages documentaires attendus à l'Université.

Bien qu'elles soient rarement explicitées, ces compétences sont précieuses et constituent, selon certains travaux, l'un des facteurs clé de la réussite du parcours universitaire.

Les compétences documentaires

Dans les années 1990, les compétences documentaires étaient définies comme la capacité à mobiliser des connaissances et des savoir-faire pour traiter et exploiter la documentation. Cela englobe, bien sûr, l'activité de recherche documentaire, mais aussi les aptitudes à déchiffrer, comprendre, assimiler et restituer les contenus. La frontière entre documentation et information devenant plus perméable, ces compétences sont désormais élargies à la notion d'information.

L'information literacy

L'anglais dispose du concept d'**Information literacy** pour désigner

cette faculté de comprendre et d'utiliser l'information. La notion a germé dans les bibliothèques académiques américaines qui proposèrent à leurs lecteurs divers services d'accompagnement pour se repérer dans les collections (*Library orientation*) ou utiliser les instruments bibliographiques (*Bibliographic instruction*). Dans les années 1980, l'essor des technologies documentaires a fait évoluer l'entreprise de *Library instruction* vers un projet plus large d'éducation à l'information. De nos jours, l'*Information literacy* est considérée comme un enjeu de société en Amérique du Nord, en Scandinavie ou en Australie.

- CHAPRON Françoise, DELAMOTTE Éric. *L'Éducation à la culture informationnelle*. Villeurbanne : Presses de l'enssib, 2010.
- COULON Alain. *Le métier d'étudiant : l'entrée dans la vie universitaire*. 2^e éd. Paris : Anthropos, 2005.
- *Culture de l'information : des pratiques... aux savoirs*. 8^e congrès de la FADBEN. Paris : Nathan, 2009.
- DENECKER Claire. *Les compétences documentaires : des processus mentaux à l'utilisation de l'information*. Villeurbanne : Presses de l'enssib, 2002.
- *Éduquer à/par l'information*. Esquisse, janvier 2007, n° 50-51.



Éducation à l'information

En France, l'expression anglaise d'*Information literacy* peine à trouver une traduction explicite, bien que le terme de « culture de l'information » commence à se répandre. C'est probablement l'une des raisons pour laquelle la notion de maîtrise de l'information est longtemps restée cantonnée à l'univers documentaire.

Depuis les années 1990, documentalistes et bibliothécaires se mobilisent pour faire reconnaître ces notions dans les établissements d'enseignement. Tandis que le concept de compétences était introduit à l'école, les professionnels de la documentation ont élaboré des outils pour

développer les compétences documentaires de leurs étudiants, dans l'espoir de pouvoir élaborer, un jour, un projet de *curriculum* qui s'étendrait « de la maternelle à l'université ».

Récemment, plusieurs travaux en didactique de la documentation ont ouvert des perspectives interdisciplinaires plus larges. L'explosion des technologies numériques a précipité les choses. La « maîtrise des techniques de l'information » est désormais inscrite au socle commun des compétences de base à l'école. Mais il reste à clarifier les confusions entre informatique et information et à travailler les rapports entre culture numérique et culture de l'information.

Les apprentissages

Pour exploiter l'information à bon escient, il faut tout à la fois apprendre à manipuler des outils (ordinateur, interface, dictionnaire par ex.), comprendre des notions (indexation, source, etc.) et piloter des activités mentales (lire, raisonner, décider, etc.).

Les savoir-faire

Les savoir-faire servent à mettre en œuvre des gestes ou des actions comme manier un outil. Ils sollicitent des connaissances spécifiques, conservées en mémoire dite procédurale. Les savoir-faire s'acquièrent par une pratique régulière et répétitive. Une fois assimilés, ils ressemblent parfois à une routine, c'est pourquoi ils restent souvent indissociables de la conduite et difficiles à exprimer verbalement.

Les connaissances déclaratives ou sémantiques

Les connaissances déclaratives mobilisent, par contre, une mémoire consciente et explicite, qui stocke les connaissances abstraites sur le monde et les objets. Les connaissances

sémantiques sont organisées par les concepts et leurs propriétés. Cela signifie qu'une nouvelle notion ne s'apprend pas *ex nihilo*, mais s'agrège au réseau de concepts déjà existant. Ici, la qualité de l'apprentissage dépendra de la capacité à enrichir sa base de connaissances et à tisser des associations riches et fondées.

Les processus cognitifs

Toute utilisation de l'information suppose, par exemple, de savoir lire, trier ou prendre des décisions. Oublier de considérer ce registre reviendrait à séparer l'information du savoir, autrement dit les données de leur traitement intellectuel.

Les processus cognitifs évoluent avec les stades de développement. Selon certaines études, le maniement de documents complexes ne pourrait s'acquérir avant l'âge adulte. Certains mécanismes cognitifs peuvent être améliorés par l'auto-régulation. Contrôler, planifier ou évaluer ses propres activités mentales est aussi un excellent moyen pour faire progresser ses compétences documentaires.

Information, culture, savoir

Les technologies de l'information ont largement contribué à démocratiser des activités détenues auparavant par les seuls professionnels de l'information ; le développement du Web 2.0 accélère encore ce phénomène.

Aussi les enjeux d'une culture de l'information apparaissent maintenant de manière cruciale. Les savoirs à mobiliser sont nombreux, ils relèvent de plusieurs domaines et concernent diverses « littéracies » : l'*information literacy*, mais aussi la culture des médias, les techniques de communication, la culture numérique et informatique.

Savoir utiliser l'information demande donc des compétences variées, ...et du temps. Il faut composer avec la complexité des notions, l'évolutivité des documents et des outils, ainsi qu'avec les rythmes nécessaires à la structuration des connaissances. Il faut, enfin, pratiquer pour comprendre, car c'est dans l'exercice de ce « savoir d'action » qu'une culture de l'information peut se construire.

Or les techniques et les outils n'instruisent pas sur le savoir lui-même. Les processus d'assimilation et de construction des connaissances procèdent d'un registre cognitif spécifique ; ils demandent une patiente maturation, qui semble parfois fort éloignée des mouvements et des flux d'information.



- **L'Éducation à la culture informationnelle : communications du colloque**, Université Lille 3, 16-18 octobre 2008 [en ligne].

<http://archivesic.ccsd.cnrs.fr/ERTE/fr/>

- ENDRIZZI Laure. *L'éducation à l'information*. Lettre de la VST, avril 2006, n° 17 [en ligne].

<http://www.inrp.fr/vst/LettreVST/avril2006.htm>

- ROUET Jean-François. *Les activités documentaires complexes : aspects cognitifs et développementaux*. HDR, 2001 [en ligne].

http://www.mshs.univ-poitiers.fr/laco/Pages_perso/Rouet/Textes.htm

- SERRES Alexandre. *Maîtrise de l'information : la question didactique*. Les dossiers de l'ingénierie éducative, avril 2007, n° 57 [en ligne].

<http://www.cndp.fr/archivage/valid/89418/89418-14447-18257.pdf>

- SICD2, Université Grenoble 2 et 3. *EruDist : votre référentiel de compétences documentaires* [en ligne].

<http://domus.grenet.fr/erudist/index.php>

Les évolutions de l'information :

Web collaboratif et gestion
de l'identité numérique

Auteur :

Olivier ERTZSCHEID

Maître de conférences – Université de Nantes

Affordance : <http://affordance.typepad.com/>

- 16** Le Web 2.0 : partage et collaboration
- 16** La gestion de l'identité numérique
- 17** Qu'est-ce que l'identité numérique ? Une collection de traces



Les outils variés du Web 2.0 permettent à tous de partager et de produire de l'information. Leur utilisation doit cependant s'accompagner d'une prise de conscience des traces laissées et de l'apprentissage de la gestion de son identité numérique.

Le Web 2.0 : partage et collaboration

Les blogs (plateformes éditoriales « clé en main » comme Wordpress), les wikis (sites collaboratifs dont chacun peut modifier le contenu, comme Wikipedia), les fils RSS (permettant de suivre les dernières nouveautés publiées sur un site), mais également ces gigantesques réservoirs de contenus que sont Flickr (pour les photos), DailyMotion ou YouTube (pour les vidéos) sont autant de sites et de technologies emblématiques de ce Web 2.0, expression qui désigne, pour les outils d'accès, de production et de recherche d'information, la part de plus en plus grande qui est laissée à l'utilisateur.

Auparavant, les contenus Web provenaient de communautés relativement réduites (enseignants, universitaires, entreprises, institutions...), et n'offraient pas beaucoup de possibilités d'interaction et de partage.

Le point commun des sites et des technologies « 2.0 » est de permettre aux internautes de partager, de participer, de « remixer », de noter, de diffuser et d'accéder à des contenus qu'ils produisent majoritairement.

Le passage au Web 2.0 se caractérise donc par une frontière de plus en plus floue et perméable entre les producteurs et les utilisateurs de contenus. Ce dernier point est aussi le plus controversé : l'abolition des filtres éditoriaux habituels, la prédominance de contenus « amateurs » font du monde numérique un terrain d'observation privilégié de la démocratie directe dans ce qu'elle a de meilleur (partage, réciprocité, interaction, collaboration) mais également de pire (dictature de la masse, notoriété l'emportant sur l'expertise, logiques quantitatives plutôt que qualitatives).

La gestion de l'identité numérique

En lien avec l'avènement du Web 2.0, l'utilisateur est aujourd'hui au cœur, non seulement des contenus circulant sur le

Net, mais également des préoccupations des grands acteurs propriétaires de ces sites, qui, pour rentabiliser leur modèle

- FAYON David. *Web 2.0 et au-delà. Nouveaux internautes : du surfeur à l'acteur*. Paris : Economica, 2008.
- ITEANU Olivier. *L'identité numérique en question*. Paris : Eyrolles, 2008.
- PISANI Francis, PIOTET Dominique. *Comment le web change le monde : l'alchimie des multitudes*. Paris : Village mondial, 2008.

économique et continuer d'offrir un accès souvent gratuit à ces services, ont besoin de connaître le plus finement possible leurs utilisateurs, pour proposer des publicités contextuelles parfaitement ciblées. L'une des grandes nouveautés du Web actuel est qu'un très petit nombre de sociétés collectent (via les

services qu'elles offrent) et peuvent recouper un très grand nombre d'informations sur chacun d'entre nous, aussi bien sur la sphère publique (le métier par exemple), la sphère privée (les relations, les amis), que sur la sphère intime (préférences politiques, sexuelles ou religieuses).

Qu'est-ce que l'identité numérique ?

Une collection de traces

La question de l'identité numérique doit être distinguée de celle des identifiants utilisés pour accéder à différents services. Sur ce dernier point, différents projets (comme *OpenID*) existent pour proposer un standard d'identification unique qui, via un seul *login* et mot de passe, permettrait de s'enregistrer sur toute une gamme d'applications et de pouvoir les utiliser.

L'identité numérique peut être définie comme la collection des traces (écrits,

contenus audio ou vidéo, messages sur des forums, identifiants de connexion...) laissées, consciemment ou inconsciemment, au fil des navigations sur le réseau. Cet ensemble de traces, une fois « remixé » par les moteurs de recherche, peut parfois produire des résultats gênants (et ce malgré le recours à des pseudonymes d'identification). Ainsi, la pratique consistant à taper le nom de quelqu'un dans un moteur de recherche pour s'informer sur lui est devenue de plus en plus cou-



- **blogs.regionsjob.com.** *L'identité numérique en question*, 2009 [en ligne].

<http://www.scribd.com/doc/14983641/Lidentite-numerique-en-question>

- **Bulletin CLIC.** *Identité numérique : être ou ne pas être sur le Web ?* Numéro 69, janvier 2009 [en ligne].

<http://clic.ntic.org/cgi-bin/aff.pl?page=article&id=2110>

- **CHAIMBAULT Thomas.** *Web 2.0 : l'avenir du web ?* Dossiers documentaires de l'enssib, septembre 2007 [en ligne].

<http://enssibweb.enssib.fr/bibliotheque-numerique/document-2>

- **GOOGLE.** *Supprimer du contenu de l'index Google* [en ligne].

<http://www.google.com/intl/fr/remove.html>

- **GUILLAUD Hubert.** *Le Web à la puissance 2 : le Web 2.0 cinq ans plus tard* [en ligne].

<http://www.internetactu.net/2009/09/01/le-web-a-la-puissance-2-le-web-20-cinq-ans-plus-tard/>

rante, notamment dans le cadre d'un recrutement. La question est d'autant plus importante qu'il est question aujourd'hui de « natifs numériques » (*digital natives*), une génération qui utilise différents services Web dès sa prime enfance.

Que faire après coup ?

En attendant des avancées sur la question du droit à l'oubli numérique, pour faire disparaître certains contenus regrettables, la personne concernée peut s'adresser d'abord aux gestionnaires des sites qui disposent desdits contenus. Cela peut concerner la suppression d'une vidéo ou de photos, mais également des messages sur différents forums de discussion. La deuxième étape est de s'adresser directement aux moteurs de recherche : ceux-ci conservent en effet souvent dans leur « cache » une version des contenus précédemment supprimés. Il faut alors leur demander la suppression de ces contenus.

Comment rester maître de son identité numérique ?

L'idéal pour maîtriser son image numérique consiste à en devenir l'acteur principal. Ainsi, le fait de tenir un blog

sous son nom, d'acheter « son » nom de domaine pour son site Web, d'y déposer son CV sont des éléments d'information qui apparaîtront en premier dans les moteurs de recherche. Mais il est difficile de demander à des adolescents alimentant leur *skyblog* ou leur *MySpace* aujourd'hui de penser à l'image que cela renverra d'eux dans 10 ou 15 ans.

Les règles du bon sens demandent donc de conserver les identifiants utilisés lors de la création d'un contenu ou d'un service, pour pouvoir, le moment venu, supprimer ce contenu ou ce service ; éviter également de déposer des messages ou des contenus qui pourraient s'avérer nuisibles à son image, quels que soient le service utilisé et son niveau de « confidentialité » affiché. Enfin, avoir présent à l'esprit cette idée qu'une réputation numérique se construit de la même manière qu'une réputation « non-numérique » – c'est-à-dire patiemment et sur la durée – représente encore le meilleur moyen d'agir en conséquence.



- O'REILLY Tim. *What Is Web 2.0*, 30 septembre 2005 [en ligne].

<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

Traduction en français :

<http://www.eutech-ssii.com/ressources/1>

- Wikipédia [en ligne].

<http://www.wikipedia.org>

- Dailymotion [en ligne].

<http://www.dailymotion.com>

- Flickr [en ligne].

<http://www.flickr.com>

- OpenID [en ligne].

<http://openid.net/>

- Youtube [en ligne].

<http://www.youtube.com>

- Wordpress [en ligne].

<http://www.wordpress.com>

Acteurs, intermédiaires et médiateurs de l'information scientifique numérique

Auteur :

Ghislaine CHARTRON
Professeur – CNAM/INTD

- 20** Les services
- 21** Quelques acteurs importants
- 23** Réorganisation de la chaîne de la documentation scientifique électronique



Un aperçu des différents services et acteurs qui interviennent dans l'information spécialisée permet de mieux comprendre et appréhender l'évolution liée à l'apparition du Web et ses répercussions sur la recherche d'information dans les ressources électroniques.

Les services

Au cours des trente dernières années, chaque support dominant à une période donnée (banques de données ASCII, cédéroms, Internet) fut associé à une organisation particulière de la chaîne de production-diffusion de la documentation scientifique, et parallèlement, à l'émergence de nouveaux acteurs. La chaîne de la documentation scientifique

implique en effet des acteurs variés, qui proposent, à différents niveaux, des services à valeur ajoutée. Le tableau ci-dessous résume les principaux services et propose quelques exemples d'acteurs en correspondance. Plusieurs niveaux sont distingués : contenu, dispositif technique, distribution-commercialisation, usager.

	Types de services	Acteurs
Contenu	Élaboration du contenu scientifique	Chercheurs, laboratoires, sociétés savantes, éditeurs
	Mise en support et diffusion du contenu	Éditeurs
	Production d'information secondaire : banques de données bibliographiques, de sommaires	Exemples : INIST, NIH Éditeurs : IEE pour Inspec. Universités, chercheurs
Dispositif technique	Accès aux infrastructures de réseau Universités	Fournisseurs d'accès Internet, Universités
	Production électronique, système d'accès	Serveurs (Dialog, Questel...), Éditeur, Éditeur électronique (HighWire Press...), Universités, Laboratoires
	Interface fédératrice pour l'utilisateur	Bibliothèque ou moteurs spécialisés
	Agrégation de différents contenus primaires ou secondaires	Exemples : Proquest, Erudit, Ovid
	Archivage numérique	Exemples : OCLC, JSTOR, éditeurs
	Archives ouvertes	Services universitaires (HAL)

- AUBRY Christine, JANIK Joanna. *Les archives ouvertes : enjeux et pratiques. Guide à l'usage des professionnels de l'information*. Paris : Édition ADBS, 2005 (Sciences et techniques de l'information).

	Types de services	Acteurs
Distribution, commercialisation	Distributeur	Exemples : OCD, Bibliopolis, agences d'abonnements
	Fournisseur de documents à l'unité	Exemples : Ingenta, British Library, INIST, PEB. Éditeurs
Usager	Gestion des ressources en ligne	Bibliothèques, universités
	Formation présentielle ou à distance	Bibliothèques ou enseignants, producteurs, éditeurs

Toutefois, les acteurs impliqués opèrent à des niveaux divers, car il existe une tendance forte à l'intégration des services, accélérée avec le développement des technologies Internet. Ainsi, de nombreux éditeurs de revues scientifiques ont développé des services d'accès aux archives électroniques de leurs revues, ainsi que

des dispositifs d'accès par Internet (plateforme **ScienceDirect** de l'éditeur **Elsevier**) et assurent de plus en plus une diffusion directe de leurs produits aux bibliothèques, sans intermédiaire commercial. De surcroît, des guides d'auto-formation aux services sont également disponibles sur ces plate-formes.

Quelques acteurs importants

Au niveau du contenu

Les **auteurs** (chercheurs, étudiants, organismes divers) produisent l'information primaire. Ils travaillent en collaboration avec des éditeurs qui organisent la validation scientifique, assurent la mise en support et la diffusion des contenus. Le développement d'Internet n'a généralement pas supprimé cette fonction intermédiaire, mais a élargi le potentiel d'acteurs concernés : les institutions publient directement différents documents sur leur site Web par exemple, les bibliothèques peuvent devenir éditrices.

Éditeurs scientifiques

Ils sont très nombreux et d'origines diverses, partagés entre l'édition de revues et d'ouvrages. On distingue les grands groupes internationaux (comme **Reed-Elsevier**, **Wolters Kluwer**) particulièrement présents en médecine et en droit, les grandes sociétés savantes américaines et européennes présentes dans les sciences exactes (**American Chemical Society**,

European Physical Society), les sociétés savantes nationales plus nombreuses en sciences humaines et sociales, ainsi que les institutions, les associations diverses ayant investi des activités d'édition.

De nombreux éditeurs ont développé ces dernières années une version en ligne de leurs revues et de certains de leurs ouvrages. Des revues exclusivement numériques ont également été développées. Le **rôle d'éditeur organisant la sélection et l'évaluation scientifique** reste essentiel dans la communication scientifique, même si des formes de circulation directe des articles se développent aussi de façon parallèle (voir les archives ouvertes).

Producteurs d'informations secondaires

Ce sont des organismes qui, par un **travail systématique de collecte et de sélection** des informations, alimentent le contenu des banques de données.

Avec les technologies Web, entreprises, associations professionnelles, organismes publics ont aussi initié divers services d'informations secondaires, diffusés souvent gratuitement.

Au niveau de la production électronique et du dispositif d'accès

Acteurs informatiques

Nombreux sont les grandes sociétés et start-up spécialisées qui proposent des services de conception et de réalisation multimédia pour le support cédérom ou le Web. À noter en particulier l'émergence d'acteurs technologiques spécialisés pour la mise en ligne des revues scientifiques dans le bassin anglophone (**HighWire Press, Ingenta**).

Diffuseurs de banques de données

Les diffuseurs sont chargés, par les producteurs de données, de la **structuration des banques et du développement des outils d'interrogation**.

Agrégateurs de contenu

Ces acteurs se sont positionnés sur la fonction d'agrégation de contenus, ils offrent des services qui permettent notamment d'**interroger plusieurs banques de données**, de consulter des collections de revues de différents domaines.

Prestataires de services pour les archives ouvertes

Les archives ouvertes sont des **réservoirs d'articles en accès libre**, sans barrière économique ni juridique, organisés de façon parallèle aux autres supports de communication scientifique (revues, ouvrages...). Le mouvement des archives ouvertes est proche du mouvement du logiciel libre, il intègre notamment des protocoles techniques d'interopérabilité

devant permettre à un usager d'interroger de façon transparente plusieurs archives ouvertes. En France, le Centre pour la Communication Scientifique Directe (CCSD) du CNRS propose des archives ouvertes par type de documents : HAL, Hyperarticle En Ligne, mais aussi TEL – Thèses En Ligne ou CEL – Cours En Ligne, par institutions ou par thématiques. Le portail Archives-ouvertes.fr a pour ambition de proposer une archive ouverte pour l'ensemble des communautés scientifiques françaises.

Au niveau de la distribution commerciale

Distributeurs commerciaux

La distribution commerciale des produits et des services documentaires peut être assurée soit par le producteur directement (en extension avec Internet), soit par un même intermédiaire ayant développé des interfaces d'accès techniques, soit par un autre intermédiaire impliqué dans des réseaux de distribution élargis. Dans le cas des banques de données, les serveurs cités précédemment assurent ce rôle de distributeur. De la même façon, les intermédiaires techniques spécialisés dans la production de cédéroms en assurent aussi la commercialisation, tout en travaillant avec d'autres réseaux de distribution.

Fournisseurs de documents à l'unité

La fourniture d'articles à l'unité s'effectue encore majoritairement sous forme papier.

De nombreux acteurs proposent aujourd'hui des services de fourniture de documents à l'unité :

- les producteurs de banques de données (**INIST**) ;

- les bibliothèques dans le cadre du prêt entre bibliothèques ;
- les éditeurs, directement à partir de leur site ;
- de nouveaux entrants, tels que Ingenta au Royaume-Uni.

Au niveau des service aux usagers

Un ensemble de tâches sont liées à la mise à disposition des ressources électroniques pour l'utilisateur et à son accompagnement pour le développement de bonnes pratiques de recherche d'information.

Bibliothèques, centres de documentation

Les licences d'accès, les relations avec les éditeurs et autres fournisseurs de contenus sont gérées par les professionnels des bibliothèques et des centres de documentation. La mise en place d'interface-

usager fédératrice est très souvent assurée par ces professionnels attentifs aux demandes et aux besoins des usagers. Par ailleurs, ces services accessibles directement sur le poste-usager nécessitent d'être accompagnés de formations méthodologiques permettant une maîtrise efficace de l'ensemble des ressources offertes, formations souvent proposées par les services documentaires.

Auto-formation

Les services d'auto-formation peuvent aussi permettre à l'utilisateur de s'approprier certains services et de parfaire ses stratégies de recherche. Les professionnels de l'information ont développé différents guides avec cet objectif : à titre d'exemple CERISE, JURISGUIDE, BIOGUIDE (accessibles sur URFIST de Paris <http://urfist.enc.sorbonne.fr/>) et certaines ressources signalées dans le site FORMIST <http://www.enssib.fr/formist/ressources>).

Réorganisation de la chaîne de la documentation scientifique électronique

Dans la chaîne de la documentation scientifique électronique, l'évitement d'intermédiaires qui opéraient jusqu'à présent entre l'auteur et le lecteur se vérifie désormais à différents niveaux. Certaines tendances de « désintermédiation » peuvent être repérées, et le lecteur final est au cœur de ces mouvements :

- La **circulation directe des articles de l'auteur au lecteur** qui, même si elle existait déjà pour le support papier dans certaines communautés, se trouve amplifiée par le réseau, voire étendue à d'autres champs disciplinaires. Les structures organisationnelles liées à cette communication directe sont plus ou moins coordonnées et investies par un financement public (soutien à la construction de « nouvelles infrastructures »).

- Les transactions directes entre sociétés savantes et chercheurs se concrétisent par la **proposition d'accès aux revues en ligne** à des tarifs très bas pour les membres de la société.
- Les transactions directes des éditeurs avec les bibliothèques, par le développement des **consortiums** d'achats.
- Les transactions directes de **vente à l'unité** entre les producteurs de banques bibliographiques (qui sont aussi fournisseurs de documents primaires très souvent) et le lecteur.

Enfin, une **intégration verticale** est importante à souligner :

- pour la fonction de diffusion et de mise à disposition qui, auparavant, justifiait le

rôle de plusieurs intermédiaires (diffuseurs, agences d'abonnements, bibliothèques),

- pour les contenus : bases de données bibliographiques, articles, archives, services de veille tendent maintenant à être proposés de façon intégrée à l'utilisateur.

Mais ces évolutions sont variables selon les communautés, les contextes institutionnels, le statut des lecteurs... et même

dans les domaines particulièrement avancés pour la mise en place de dispositifs de communication directe entre chercheurs, il existe des phénomènes d'accumulation de services plus que de substitution. Ainsi, les revues papier continuent à être largement défendues pour deux raisons majeures : le confort du support de lecture et la garantie d'un archivage indépendant.



- **Archives-ouvertes.fr [en ligne].**
<http://hal.archives-ouvertes.fr/>
- **FORMIST - Réseau francophone pour la formation à l'usage de l'information dans l'enseignement supérieur [en ligne].**
<http://www.enssib.fr/formist/>
- **HAL - Hyperarticle en ligne [en ligne].**
<http://hal.archives-ouvertes.fr/>
- **INIST – Institut de l'Information Scientifique et Technique [en ligne].**
<http://www.inist.fr/>
- **Scirus - for scientific information [en ligne].**
<http://www.scirus.com>
- **URFIST de Paris [en ligne].**
<http://urfist.enc.sorbonne.fr/>

La recherche des ressources électroniques : un aperçu historique

Auteur :

Catriona BERGIA
Ingénieur d'étude - ensib

- 26** Les années cinquante
- 27** Les années soixante
- 27** Les années soixante-dix
- 27** Les années quatre-vingt
- 28** Les années quatre-vingt dix
- 28** Les années deux mille



Les débuts puis le développement de la recherche documentaire informatisée ont été intimement liés à l'apparition et au développement de l'informatique et de la télécommunication.

Les années cinquante : premières recherches, premiers chercheurs "experts"

L'apparition des premiers ordinateurs, dans les **années cinquante**, a permis leur utilisation pour assister la recherche documentaire. Initialement, ces machines ne servaient qu'à lire des cartes trouées qui portaient des informations codées, mais une évolution rapide leur a donné la possibilité de consulter, par voie d'impression et de manière extrêmement rudimentaire, des données stockées sur bandes magnétiques.

Pendant cette première période, les recherches nécessitent la présence de l'interrogateur sur place, à côté de la machine, car la seule manière de lire les informations obtenues est la lecture des « sorties » papier (l'écran n'existe pas

encore). Les données contenues dans les premières bases sont initialement fournies sous forme de listings, sans aucun tri, filtre, ou sélection possible. Puis, avec la création d'index, il devient possible de lancer les premières requêtes, parfois groupées afin d'optimiser l'utilisation des machines, car le temps de recherche peut se compter en heures ou même en jours.

La manipulation des machines de cette époque s'effectue par des experts, car elle nécessite de connaître précisément les machines, ainsi que leur fonctionnement. Ces experts reçoivent des demandes de recherches, souvent par courrier, et après avoir réalisé le travail demandé, expédient les réponses par la même voie.

- CHAUMIER Jacques. *Les Techniques documentaires au fil de l'histoire : 1950-2000*. Paris : ADBS, 2002 (Sciences de l'information ; Série Études et techniques).
- COMBEROUSSE Martine. *Histoire de l'information scientifique et technique*. Paris : Nathan, 1999 (Collection 128 ; Information-documentation).
- FAYET-SCRIBE Sylvie. *Histoire de la documentation en France. Culture, science et technologie de l'information : 1895-1937*. Paris : CNRS Éditions, 2000.



Les années soixante : réseaux et bases de données, les utilisateurs “intermédiaires”

La fin des **années soixante** voit la situation évoluer considérablement. Du fait de la simplification de l'utilisation des machines, le nombre des personnes formées pour réaliser des recherches documentaires s'accroît. Ces intermédiaires travaillent pour répondre à la demande d'utilisateurs finaux, qui souhaitent exploiter les données résultant des recherches. Ainsi est né le métier de documentaliste.

En parallèle, la possibilité de se connecter à un ordinateur distant est devenue réalité avec l'utilisation d'abord du réseau téléphonique existant, puis de réseaux

dédiés, souvent mis en place pour des objectifs gouvernementaux et publics (militaire, recherche, santé). Les requêtes s'effectuent sur des terminaux affectés à certaines bases pour la consultation, et sont lancées par lot afin d'optimiser le temps d'utilisation. Cette méthode reste également très rigide, car il n'est guère possible de modifier et de relancer une requête en fonction des premiers résultats obtenus. Par ailleurs, le documentaliste doit utiliser et maîtriser des langages de requête spécifiques pour consulter les bases de données.

Les années soixante-dix : la recherche “en ligne”, les disques durs, les serveurs commerciaux

Dans les **années soixante-dix**, la consultation de bases de données en mode « conversationnel » se développe, avec des échanges d'information de plus en plus rapides entre le documentaliste et la machine ; les requêtes peuvent alors être traitées de façon individuelle. Ce gain de temps a été rendu possible grâce au transfert des informations des bandes magnétiques vers les disques de stockage de données, disques qui permettent un

accès libre et plus direct aux informations : il n'est plus nécessaire, comme avant, de parcourir l'ensemble des données lors de chaque requête.

Les bases de données deviennent alors de plus en plus nombreuses, et se commercialisent *via* des accès par l'intermédiaire de serveurs commerciaux payants. Si environ 20 bases étaient disponibles au milieu des années 1965, elles sont près de 200-300 dix ans plus tard.

Les années quatre-vingt : les interfaces graphiques, la consultation en simultanée, l'utilisateur “final”, le cédérom

Dans les **années quatre-vingt**, une multitude d'améliorations et de nouvelles possibilités sont apportées dans la

recherche documentaire informatisée. Des interfaces graphiques facilitent tous les aspects de la recherche : système de

recherche adapté au niveau de compétence du chercheur (recherche guidée / experte), accès aisé à des fonctions d'aide. La recherche peut alors être enfin effectuée directement par l'« utilisateur final », sans plus passer par un intermédiaire. Les développements en informatique rendent aussi possible la consultation de plusieurs bases de données en simultané, permettant une amélioration en qualité de tous les aspects de la recherche.

Enfin, lors de cette période, l'utilisation de cédéroms permet l'interrogation de bases de données sur des postes informatiques en local, sans connexion à un réseau, avec un confort graphique amélioré, une simplicité apparente d'utilisation, etc.

C'est à cette période que la formation à la recherche documentaire informatisée s'est développée dans les bibliothèques.

Les années quatre-vingt dix : Internet, cédérom en réseau, revues électroniques

Le développement d'Internet, dans les années quatre-vingt dix, a facilité l'accès à des ressources documentaires électroniques, et le nombre d'« utilisateurs finaux » de ces ressources s'est alors multiplié. La recherche documentaire informatisée est une activité qui semble maintenant pratiquée par le grand public, même si elle nécessite un minimum de formation et d'encadrement.

Les bases de données disponibles sur cédérom sont aussi désormais interrogeables en réseau et connaissent une utilisation très importante dans les lieux d'étude et de travail.

Parallèlement à cela, les revues auparavant disponibles uniquement sur papier sont elles aussi consultables en ligne dans leur version électronique.

Les années deux mille : vers le collaboratif

Internet évolue encore, avec le Web 2.0. Les applications informatiques deviennent « tout Web » avec des suites bureautiques complètes en ligne, l'interactivité avec les utilisateurs se développe, à travers des réseaux sociaux, des possibilités de commentaires et la création d'interfaces enrichies.

Internet apparaît comme un énorme gisement d'information, dont la porte d'entrée est l'entonnoir Google. Avec l'accès simplifié à une multitude d'information, la recherche s'est appauvrie.

- **SERRES Alexandre. *Histoire des outils et réseaux d'information : de Paul Otlet à Internet* [en ligne].**

http://www.sites.univ-rennes2.fr/urfist/histoire_internet



L'offre

L'économie des ressources électroniques

Auteur :

Benoît EPRON

Maitre de conférences – enssib

- 30** Les revues scientifiques : des modèles économiques stabilisés
- 31** Les ouvrages : une offre encore en devenir



Après une première phase de développement important de l'offre de ressources électroniques pour l'enseignement supérieur, ces dernières années ont montré une relative stabilisation des modèles proposés. Cette maturité touche essentiellement le secteur des revues scientifiques, celui des ouvrages ou des monographies restant encore, dans une large mesure, le terrain d'expérimentations.

La structuration des offres de ressources électroniques pour l'enseignement et la recherche se fait principalement selon deux modèles économiques.

Les revues scientifiques : des modèles économiques stabilisés

Le modèle de l'accès libre

Qu'il s'agisse de dépôts institutionnels, d'archives ouvertes ou de publications sous licence *Creative Commons*, une part importante de la production scientifique est aujourd'hui disponible en accès libre. Les modèles économiques qui sous-tendent ces offres sont divers.

Des publications scientifiques collaboratives ou prises en charge par une com-

munauté de chercheurs utilisent les outils de publication d'Internet pour développer une version numérique. Cette version numérique est dans ce cas accessible librement (voir le projet SCOAP).

Dans d'autres cas, l'équilibre économique de la revue est maintenu par une logique d'embargo, c'est-à-dire un délai limité dans le temps entre la publication dans la revue et la disponibilité finale dans un dépôt en accès libre (voir, pour l'offre fran-

- CHARTRON Ghislaine (dir.). *Les chercheurs et la documentation numérique : nouveaux services et usages*. Paris : Éditions du Cercle de la Librairie, 2002.
- CHARTRON Ghislaine. *L'édition scientifique face à Internet*. In *Bibliothèques numériques : cours INRIA, 9-13 octobre 2000, La Bresse*. Paris : ADBS, 2000 (Études et techniques).
- SALAÛN Jean-Michel (dir.). *Économie et bibliothèques*. Paris : Éditions du Cercle de la librairie, 1997.
- SHAPIRO Carl, VARIAN HAL R. *Économie de l'information : guide stratégique de l'économie des réseaux*. Bruxelles : De Boeck Université, 1999.

cophone, la construction de l'offre Cairn). Enfin, en suivant la logique des dépôts institutionnels, il existe aujourd'hui un nombre croissant de modèles « auteur/institution – payeur » dans lesquels le chercheur ou son institution paye la revue pour bénéficier du droit de diffusion en accès libre (voir PLOS).

L'accès par licence

Le modèle de référence pour la diffusion des revues scientifiques numériques reste aujourd'hui encore le modèle de la licence, qui donne accès le plus souvent à un bouquet de plusieurs revues. Le montant de la licence est dans certains cas calculé en fonction du nombre de personnes pou-

vant y accéder simultanément en équivalent temps-plein. Ce sont donc des licences institutionnelles, proposant un accès sur site ou hors site (avec mots de passe) aux utilisateurs. Ces licences sont aujourd'hui négociées en partie dans le cadre de consortium (comme COUPERIN ou CAREL). Les bibliothèques fournissent ainsi l'accès, en général depuis l'espace physique de leur institution (par reconnaissance d'adresses IP) à un ensemble fini de revues en ligne.

Ce glissement de la propriété vers le seul accès aux revues fait émerger des interrogations sur la garantie d'un accès pérenne aux ressources ainsi acquises. Des solutions, centralisées ou distribuées, commencent à être mises en place.

Les ouvrages : une offre encore en devenir

Les acteurs du secteur des ouvrages universitaires ont intégré ce mouvement de développement d'offres numériques plus tardivement que les revues. Ils ont ainsi pu bénéficier de l'expérience déjà acquise sur le secteur adjacent des revues, notamment l'accès par licences, mais ont également été confrontés à des problé-

matiques propres, liées pour partie à la singularité du livre dans sa transposition numérique.

En effet, bien que les nouveaux supports de lecture de contenus éditoriaux numériques soient encore relativement rares, des projets de plus en plus nombreux pro-



- Cairn [en ligne].

<http://www.cairn.info/>

- CAREL : *Consortium pour l'Acquisition des Ressources Electroniques en Ligne* [en ligne].

http://www.bpi.fr/fr/professionnels/collections_et_services2/carel_ressources_electroniques.html

- Couperin - *CONsortium Universitaire de Publications Numériques* [en ligne].

<http://www.couperin.org/>

- Creative Commons France [en ligne].

<http://fr.creativecommons.org/>

posent une offre de livres numériques et s'accompagnent de nouveaux modèles d'affaires.

Historiquement, les premiers modèles s'inscrivent dans une transposition fidèle du modèle de la bibliothèque : le nombre d'« exemplaires » numériques d'un ouvrage correspond au nombre de consultations ou d'emprunts simultanés possibles. Dans cette option, le coût des ouvrages est souvent similaire à celui des versions papier existantes.

Aujourd'hui de nouvelles approches sont proposées qui intègrent les logiques de licences issues du secteur des revues. Elles offrent l'accès à une collection d'ouvrages numériques, éventuellement de façon illimitée et à nouveau en utilisant le nombre d'utilisateurs potentiels comme base de calcul. Ces nouveaux modèles modifient profondément les logiques d'acquisition et de gestion de collections à l'œuvre en bibliothèques universitaires. Ces modifi-

cations, couplées à une mise à disposition dépassant les frontières des BU, touchent d'ailleurs les logiques macroéconomiques régulant les relations de coopérations entre les établissements.

Ces modèles, en grande partie construits sur la base des objets et des logiques économiques liés au papier, seront vraisemblablement amenés à évoluer. Un des points clés de cette évolution réside dans l'émergence d'une offre de dispositifs de lecture réellement satisfaisants (objets sans nom fixé pour l'instant : *e-book*, *e-reader*, liseuses, etc.), adaptés à la lecture d'ouvrages longs dans leur version numérique, *offre qui se développe avec le Kindle aux USA, l'Iliad de Irex ou le CyBook en Europe.*

Enfin, en se numérisant, les ouvrages universitaires vont intégrer un marché de dimension plus internationale, entraînant une modification importante des logiques concurrentielles entre les éditeurs des différents bassins linguistiques.



- **MANUM - *Projet expérimental de bibliothèque numérique pour les étudiants en sciences humaines et sociales*, mai 2003 [en ligne].**
<http://manum.enssib.fr/>
- **OpenDOAR - *The Directory of Open Access Repositories* [en ligne].**
<http://www.openoar.org/index.html>
- **PLOS [en ligne].**
<http://www.plos.org>
- **SCOAP [en ligne].**
<http://scoap3.org/>
- **WEENINK Kasja, WAAIJERS Leo, VAN GODTSENHOVEN Karen. *À DRIVER's Guide to European Repositories*. Amsterdam : Amsterdam University Press, January 2008 [en ligne].**
<http://dare.uva.nl/document/93898>

Techniques d'accès aux
ressources électroniques
et à Internet

Auteurs :

Marie-Françoise DEFOSSE
Professeur associée – enssib

Élisabeth NOËL
Conservateur des bibliothèques – enssib



La connaissance des techniques de l'Internet est utile pour mieux comprendre les résultats des recherches d'informations, pour l'évaluation des sources d'informations et pour leur utilisation.

Dispositifs techniques d'accès à l'information

Internet, qu'est-ce-que c'est ?

INTERNET (INTERconnected NETwork) est né en 1969 du besoin conjoint des chercheurs américains et de l'armée de communiquer *via* des ordinateurs. Au début des années 90, l'arrivée des ordinateurs individuels a démocratisé l'accès au réseau sur le lieu de travail, la maison et les lieux publics.

Internet est devenu un instrument de communication ouvert à tous pour communiquer, échanger, consulter des documents et aussi en publier. Tout le monde peut être à la fois lecteur et fournisseur d'informations, d'autant que le Web 2.0 facilite la production. Mais rien n'authentifie les informations, qui peuvent être falsifiées (problèmes de date, de mise à jour...) ou erronées (**hoax** : fausses informations, canulars), d'où l'importance de l'évaluation des sites Internet.

Personne ne connaît exactement l'étendue d'Internet, son contenu demeure encore éphémère et volatil. L'archivage du Web est partiellement assuré via des initiatives intéressantes ([Internet Archive](#),

<http://www.archive.org/> ou le dépôt légal de l'Internet, à la Bibliothèque nationale de France, http://www.bnf.fr/fr/professionnels/depot_legal/a.dl_sites_web_mod.html).

Aussi, il est toujours possible d'intercepter les paquets d'information qui circulent dans le réseau. D'où aussi des problèmes de sécurité des transactions commerciales ou de confidentialité sur les sites non sécurisés. **Un site sécurisé se reconnaît par la présence d'un cadenas fermé ou d'une clé intacte (non brisée) dans le coin inférieur droit ou gauche de l'écran, selon le navigateur utilisé. De plus, l'adresse URL d'un site sécurisé commence habituellement par « https » ou « shttp » au lieu de « http ».**

Enfin, le téléchargement de logiciels entraîne toujours le risque de rapatrier des virus ! Il est donc conseillé de vérifier chaque logiciel ou document récupéré par un anti-virus à jour, les meilleurs étant gratuits. Il est aussi utile de se prémunir contre les *spywares*, logiciels espions qui fournissent aux sociétés qui les diffusent des informations sur les habitudes des internautes, sans que ceux-ci en soient informés.

- LEVINE John R. *Internet pour les nuls*. Paris : Éditions Générales First , 2008.
- VIRIEUX Françoise. *Comment marche Internet ?* Paris : Le pommier, 2004 (Les petites pommes du savoir).

Connexion à un réseau

Le protocole de base : TCP-IP

Du point de vue technique, Internet est un terme générique qui désigne la multitude des réseaux d'ordinateurs qui communiquent entre eux, au niveau mondial, grâce à un protocole d'échange de données standard : TCP/IP (Transmission Control Protocol/Internet Protocol – se prononce T-C-P-I-P), indépendamment des types d'ordinateurs utilisés (Mac, PC).

TCP/IP est basé sur le repérage de chaque ordinateur par une adresse appelée **adresse IP**, qui permet d'acheminer les données à la bonne adresse. Par exemple, 194.153.205.26 est une adresse TCP/IP donnée sous sa forme technique. Ce sont ces adresses que connaissent les ordinateurs pour communiquer entre eux.

Les noms de domaines

Les utilisateurs ne veulent pas travailler avec des adresses numériques mais avec des noms plus explicites. Le protocole TCP/IP a donc permis d'associer des noms en langage courant aux adresses numériques, grâce à un système appelé DNS (*Domain Name System*). Le **nom de domaine** est le nom à deux composantes, dont la première partie correspond au nom de l'organisation ou de l'entreprise, la deuxième à la classification de domaine (.fr, .com, ...) – c'est l'extension.

Exemple : dans l'adresse <http://www.enssib.fr>, « enssib.fr » est le nom de domaine : enssib est le nom de l'organisation, .fr correspond à la classification. Les

« www » ne correspondent à rien de significatif, ils peuvent être remplacés par d'autres lettres, si elles sont déclarées. « signets.bnf.fr » correspond à un site particulier dépendant du domaine bnf.fr

Chaque machine d'un même domaine est appelée **hôte**. Le nom d'hôte qui lui est attribué doit être unique dans le domaine considéré.

Syntaxe d'une URL

Une URL (*Uniform Resource Locator*) est un format de nommage universel pour désigner une ressource sur Internet. Il s'agit d'une chaîne de caractères qui se compose de quatre parties :

- 1- Le nom du **protocole** : le langage utilisé pour communiquer sur le réseau. Ce n'est pas toujours « http ».
- 2- L'adresse du serveur : le **nom de domaine** de l'ordinateur hébergeant la ressource demandée. Il est possible d'utiliser l'adresse IP du serveur, ce qui est moins pratique.
- 3- Le numéro de port : facultatif, ce numéro associé à un service permet au serveur de savoir quel type de ressource est demandé.
- 4- Le **chemin d'accès** à la ressource : cette dernière partie indique au serveur l'emplacement où la ressource est située, c'est-à-dire le chemin (répertoire et sous-répertoires) et le nom du fichier demandé.

Les réseaux de télécommunications

En France, le réseau **Renater** (REseau NATional de télécommunications pour la

Une URL a donc la structure suivante :

Protocole	Adresse du serveur	Port (facultatif)	Chemin /fichier
http://	www.enssib.fr		/glossaire/glossaire.htm

Exemples : <http://www.bium.univ-paris5.fr/medecine/formation/zotero.pdf>
<ftp://ftp.adobe.com>

Technologie, l'Enseignement et la Recherche), géré par un groupement d'intérêt public, interconnecte les différentes universités et organismes de recherche français au niveau national, avec les autres réseaux de la recherche au niveau européen et avec l'Internet, en général.

D'autres réseaux de communication permettent également une connexion à Internet. Les particuliers avec un ordinateur équipé d'un modem peuvent ainsi être connectés à Internet *via* le réseau téléphonique commuté (ou Numéris), par l'intermédiaire d'un prestataire qui leur fournit le point d'accès. Le **Fournisseur d'accès à Internet** (FAI, *provider* en anglais) permet de se connecter à Internet *via* un modem, par une ligne téléphonique, par l'ADSL (*Asymmetric Digital Subscriber Line*), par le câble ou par le Wi-Fi (*Wireless Fidelity* - réseau local sans fil). Il offre différents services : adresses de messagerie, hébergement de pages personnelles, service technique (*hot-line*)... Son choix doit prendre en compte les besoins, les services offerts, leur qualité et leur coût.

Protocoles de connexion

Le réseau offre différents types de services :

SMTP - Le courrier électronique (courriel, mél ou e-mail)

Le protocole **SMTP** (*Simple Mail Transfer Protocol*) est le protocole standard du courrier électronique, de la messagerie individuelle, des listes de discussion, qui permet de transférer le courrier d'un serveur à un autre.

Des logiciels installés sur l'ordinateur (comme **Thunderbird**, **Eudora** ou **Outlook**) aident à gérer facilement son courrier grâce à des interfaces conviviales. Les messages reçus restent stockés sur le serveur jusqu'à leur transfert sur l'ordinateur client pour

consultation. La messagerie est aussi consultable directement en ligne, sur le site du fournisseur d'accès, à travers des services Webmail qui permettent de gérer tous les messages directement sur le serveur. Si la consultation nécessite d'être connecté, elle permet cependant de lire et d'envoyer des mails depuis n'importe où, avec un même compte mail.

Pour correspondre avec une personne où qu'elle se trouve dans le monde, il suffit qu'expéditeur et destinataire possèdent une adresse électronique qui se présente sous la forme :

nom_usager@adresse_organisation.extension

Il n'existe pas d'annuaire mondial des adresses Internet ; même si plusieurs services permettent des recherches, le résultat est rarement performant.

FTP - Le transfert de fichier

Le protocole **FTP** (*File Transfer Protocol*) est un protocole de transfert de fichier, qui définit la façon selon laquelle des données doivent être transférées sur un réseau TCP/IP. Il permet :

- un partage de fichiers entre machine distante ;
- une indépendance aux systèmes de fichiers des machines clientes et serveur ;
- le transfert efficace des données.

Ce protocole permet d'accéder à un serveur distant et de transférer des fichiers contenant des textes, des images, du son ou des logiciels. Il nécessite de disposer d'un programme de décompression (winzip). Il peut être utilisé *via* un navigateur ou par un logiciel spécifique.

La syntaxe est la suivante :

ftp://adresse_serveur/répertoire/
sous-répertoire/fichier

Ex : ftp://ftp.inria.fr/

NNTP - Forums de discussion

Le protocole des **forums** est le *Network News Transfer Protocol*. Les **news** ou **newsgroups** sont consultables sur les serveurs Usenet (news : nom_du_groupe), à travers un logiciel spécifique ou intégré, et divisés en grandes catégories selon des thèmes définis comme :

Préfixe	Contenu
Alt	discussions « alternatives » abordant les sujets les plus variés
Bionet	recherche en sciences de la vie, biologie
Comp	tous les domaines de l'informatique
Fr	groupes francophones, messages en français
Rec	loisirs, groupes des passionnés

Sci	recherche en sciences exactes
Soc	recherche en sciences humaines, sociales

Les adresses sont structurées dans le sens inverse des URL.

Exemple : [fr.bio.medecine](mailto:fr.bio.medecine@france.fr) correspond au domaine France, au thème biologie, au groupe Médecine.

HTTP - L'accès aux serveurs d'informations WWW

Le protocole HTTP (*Hypertext Transfer Protocol*) est le protocole de communication entre un serveur WWW, offrant des contenus multimédias reliés par des liens hypertexte, et un logiciel client, le navigateur (Internet Explorer ou Firefox par exemple). Ce service Internet, qui permet de naviguer à travers des pages Web, est généralement appelé "Web".

Nature du Web

Le **world wide web** est un service récent d'Internet. C'est en 1990 que des informaticiens du CERN (Centre Européen de Recherche Nucléaire) proposent cette application destinée à la publication et la diffusion de documents multimédia sur le réseau, au sein de la communauté internationale des physiciens. Le Web s'est imposé grâce à ses qualités :

- le recours au multimédia, qui mixe le texte, les images fixes et animées, le son ;
- l'hypertexte, qui permet une navigation intra et inter-documents et a permis de tisser une « toile » mondiale.

Modalités de connexion au Web

Internet privilégie l'environnement **client-serveur** : l'utilisateur travaille donc depuis son propre ordinateur, sur lequel ont été transférés les fichiers demandés au site consulté. Quand nécessaire, des échanges

se font entre l'ordinateur individuel (le client) et le serveur. L'affichage se fait grâce au langage HTML (*HyperText Markup Language*) qui permet d'indiquer au client (ici, le navigateur) comment interpréter les données. L'option Affichage/Source dans le menu du navigateur permet de voir une page telle qu'elle est codée en HTML.

Certains documents nécessitent l'activation de **plugins**, logiciels qui s'exécutent à l'intérieur du navigateur : Realplayer, Shockwave... pour le son et les animations. Ces logiciels sont généralement gratuits.

Le Web en 2011

La facilité de production de pages HTML et la popularité croissante des blogs, l'existence de nombreux sites d'hébergement gratuits, les faibles coûts des serveurs ont conduit à un développement quantitatif

énorme du nombre de pages HTML depuis 1995.

Le Web est souvent divisé en trois parties : le Web public, libre d'accès, le Web privé, sur mot de passe et le Web provisoire, en construction. Les moteurs de recherche indexent en grande partie le Web public, mais ignorent le reste, qui demeure invisible à leurs yeux.

Web statique, Web dynamique

Le standard établi pour la diffusion de documents sur le Web est le langage HTML, qui permet de définir la présentation du document ainsi que les liens hypertextes vers d'autres documents à l'aide de balises de formatage. La description du contenu a été prévue dans les balises de métadonnées (metadata : données sur les données), étendues avec le Dublin Core (*Dublin Core Metadata Initiative*, <http://dublincore.org/>). Cependant pour diverses raisons, depuis 2001 les moteurs de recherche prennent peu en compte ces éléments, qui restent utiles dans le cadre d'un travail en réseau documentaire ou en intranet avec utilisation d'un moteur d'indexation.

Les documents fixes produits à l'aide d'un éditeur de texte ou d'un éditeur HTML sont stockés en tant que fichiers texte (extension HTM, HTML, SHTML...) sur le disque dur du serveur Web qui les diffuse. Ils constituent des pages HTML statiques.

Cependant, il est « plus facile » de gérer une base d'informations que de modifier des centaines de textes HTML. D'où le développement de bases de données, qui diffusent l'information à la demande, suite à l'interrogation de la base de données par l'intermédiaire d'un formulaire : il s'agit alors de **pages HTML dynamiques**.

Le Web invisible

Le Web invisible est défini comme l'ensemble des documents diffusés sur le Web, sans être indexés par les moteurs de recherche traditionnels. Plusieurs raisons dues aux méthodes de référencement dans les outils de recherche l'expliquent. Les moteurs de recherche envoient des robots balayer le Web pour alimenter en données leur index, mais certaines pages Web ne peuvent pas être indexées : mot de passe demandé (sites commerciaux permettant le paiement en ligne par carte bancaire, Intranet), niveau de profondeur d'indexation ou de profondeur de contenu indexé, pages vers lesquelles ne pointe aucun lien, formats de fichiers non reconnus, indexation difficile des pages Web dynamiques, mauvais référencement ou volonté des auteurs et responsables de sites, etc.

Le Web invisible regroupe des informations en général intéressantes, spécialisées, appartenant à la littérature grise, aux archives ouvertes et aux bases de données.

- **AFNIC. Chartes de nommage [en ligne].**

<http://www.afnic.fr/obtenir/chartes>

- **CHAIMBAULT Thomas. L'archivage du web. Dossiers documentaires de l'enssib, mars 2008 [en ligne].**

<http://www.enssib.fr/bibliotheque-numerique/document-1730>

- **Comment ça marche. Les adresses IP [en ligne].**

<http://www.commentcamarche.net/internet/ip.php3>

- **USENET fr. Liste des groupes FR [en ligne].**

<http://www.usenet-fr.net/liste-groupes.html>



Création et utilisation de ressources électroniques : quelques notions juridiques

Auteur :

Yves ALIX

Conservateur de bibliothèques
Bibliothèque nationale de France

- 40** Le cadre de référence juridique
- 40** Le droit d'auteur
- 44** Les déclarations de site Web



Sur Internet, une infinité d'œuvres susceptibles d'être protégées par le droit d'auteur circulent en permanence et peuvent être copiées instantanément et à l'identique, dans des conditions de qualité égale à l'original numérique. Les lois du 1^{er} août 2006 sur le droit d'auteur dans la société de l'information (dite loi Dadvsi) et du 21 juin 2004 sur la confiance dans l'économie numérique adaptent la protection à l'environnement numérique.

Le cadre de référence juridique

Le cadre juridique a été codifié en France dans le *Code de la propriété intellectuelle* (CPI), adopté le 1^{er} juillet 1992, qui a intégré toutes les dispositions antérieures. Depuis 1992, d'autres textes ont été adoptés, modifiant le CPI. Outre les deux lois mentionnées plus haut, il s'agit de la loi de 1995 sur la reprographie (photocopie), celle de 1998 sur le droit des bases de données ou encore la loi du 18 juin 2003 sur le droit de prêt des livres en bibliothèque.

L'application des règles du CPI à Internet

L'existence récente de dispositions spécifiques à Internet ou aux ressources numériques ne signifie pas pour autant que les principes fondamentaux du droit d'auteur, non plus que les principes de responsabilité liés à la publication – contenus dans les lois sur la presse (1881) ou l'audiovisuel (1986) – ne s'appliquent pas sur le réseau. Les utilisateurs doivent ainsi constamment garder à l'esprit que les œuvres de toute nature sont protégées sur Internet au même titre que partout ailleurs, quel que soit leur support.

Le droit d'auteur

Il protège l'œuvre, et donc son auteur, dès la création, en dehors de toute formalité préalable. Aujourd'hui, presque toutes les formes de créations (bande dessinée, photos, œuvres graphiques, mais

aussi créations en 3D, logiciels, bases de données, blogs, œuvre d'architecture) peuvent être protégées par le droit d'auteur. Pour accéder à la protection, il suffit que la création, même inachevée, soit



concrétisée par une mise en forme (les idées ne peuvent être protégées) et que celle-ci soit originale.

En principe, l'auteur est une personne physique, et seul titulaire des droits d'auteur.

Mais de nombreuses exceptions existent ainsi que des cas plus complexes (auteurs multiples, adaptateurs, traducteurs, etc.) :

- l'auteur peut avoir cédé ses droits à un éditeur ou un producteur ;

- l'auteur peut avoir confié la gestion de ses droits à une société d'auteurs, chargée de contrôler l'utilisation des œuvres, de percevoir et de répartir les rémunérations (gestion collective) ;

- si l'auteur d'un logiciel est salarié et que le produit a été réalisé dans l'exercice de ses fonctions, l'employeur est investi des droits de l'auteur ;

- journaux et revues, dictionnaires et encyclopédies, dans leur version papier comme en ligne, sont des œuvres dites « collectives » dont la propriété appartient aux personnes physiques ou morales sous le nom desquelles elles sont divulguées.

Les prérogatives de l'auteur comportent deux volets.

Les droits moraux

Perpétuels, transmissibles aux héritiers, ils protègent l'œuvre contre les mutilations et les atteintes susceptibles de la dénaturer ou d'en altérer le contenu. Le respect de l'intégrité de l'œuvre est donc impératif. **Par exemple, sur un site, recadrer une photo ou couper un texte, etc., nécessitent d'avoir obtenu l'autorisation de le faire.** Les œuvres utilisées doivent être clairement identifiées, correctement légendées (droit au respect du nom de l'auteur et de la paternité de l'œuvre).

Les droits patrimoniaux

Destinés à assurer la subsistance de l'auteur, ils sont limités dans le temps, cessant en principe 70 ans après la mort du ou des auteurs (ou 70 ans après première publication pour les œuvres collectives) :

- **le droit de reproduction** soumet à autorisation toute copie, intégrale ou partielle, d'une œuvre, sur tout support. Seules sont autorisées les copies à usage privé du copiste, non destinées à utilisation collective. Le titulaire de droits peut aujourd'hui, avec la loi Dadvsi, limiter même les copies privées par des mesures techniques, toute copie pouvant aussi donner lieu à compensation financière ;

- **le droit de représentation** (ou de communication publique) couvre toutes les formes de présentation directe des œuvres au public, **par exemple dans les cinémas, les théâtres, sur Internet, dans une exposition, par la diffusion de musique enregistrée**, autant de représentations soumises à autorisation des ayants droit.

Les exceptions

Des exceptions aux droits patrimoniaux de l'auteur sont tolérées :

- **représentations privées** et gratuites dans le cercle de la famille ;

- **copies ou reproductions strictement réservées à l'usage du copiste** (pour le logiciel, seule la copie de sauvegarde est autorisée) ;

- **usage de courtes citations**, si elles sont incorporées à une œuvre nouvelle à caractère critique, pédagogique, scientifique ou d'information. Le nom de l'auteur et la source doivent toujours être indiqués, la citation doit être brève par rapport à l'œuvre citée. La jurisprudence refuse cette exception aux œuvres musicales, audiovisuelles ou aux arts plastiques.

La loi du 1^{er} août 2006 a élargi le régime des exceptions, entre autres :

- pour la **représentation ou la reproduction** d'extraits d'œuvres (sauf œuvres conçues à des fins pédagogiques, partitions de musique et œuvres réalisées pour une édition numérique de l'écrit) à **des fins exclusives d'illustration** dans le cadre de l'enseignement et de la recherche ;

- pour la reproduction et la représentation d'œuvres, par des personnes morales

et par les établissements ouverts au public, en vue d'une **consultation strictement personnelle de l'œuvre par des personnes handicapées** ;

- pour la reproduction d'une œuvre, effectuée à des fins de conservation ou de préservation des conditions de sa consultation sur place par des bibliothèques, des musées ou des archives.

Œuvres et ressources libres d'accès et d'utilisation

Les actes officiels (**constitution, lois, règlements, travaux parlementaires, décisions de justice, rapports officiels...**), les données brutes (**liste d'adresses et de numéros de téléphone, tableaux scientifiques, langues, données historiques**) et les œuvres tombées dans le domaine public sont libres d'utilisation, sous réserve du droit moral pour la dernière catégorie citée.

Les **revues de sommaires** sont libres. Les **revues de presse**, qui présentent conjointement divers commentaires sur un même événement, sont possibles, sous réserve que soient clairement indiqués le nom de l'auteur et la source de l'œuvre. Le **panorama de presse**, lui, reproduit simplement les articles. Le Centre français du droit de copie (CFC) gère collectivement les panoramas de presse papier ; pour les panoramas électroniques, il faut demander les autorisations à tous les ayants droit, ou passer par des agrégateurs de presse (PressEd, France Actu, CFC, etc.).

Les droits voisins

Introduits dans la loi en 1985, ils confèrent aux **interprètes** un droit exclusif d'autoriser ou non la fixation de leur interprétation (**un chanteur ou un acteur ne peut être enregistré ni filmé sans son autorisation, sauf pour un usage exclusivement privé**). Les producteurs de phonogrammes

(CD, fichier MP3) et de vidéogrammes (vidéocassettes, DVD, tout enregistrement audiovisuel, éventuellement diffusé sur Internet) ont le droit exclusif d'autoriser ou non toute forme de communication de l'enregistrement, y compris le prêt et la diffusion en ligne. Ces droits durent 50 ans à compter de la première mise à disposition du public.

Les droits des producteurs de bases de données

Quelle que soit la nature des données qui la constituent, l'organisation de l'ensemble d'une base de donnée et sa valeur en tant qu'investissements économiques sont protégées contre toute extraction substantielle ou systématique, pendant quinze ans à compter de la mise à disposition. L'autorisation du producteur de la base doit donc être sollicitée pour tout usage excédant le cadre strictement privé.

Les licences Creative Commons

Très répandues aujourd'hui sur Internet et dans le monde de la recherche, transposées en français en 2004, les *Creative Commons* sont une alternative à la gestion par le droit d'auteur, particulièrement intéressante pour les contenus scientifiques, et garantissent à la fois protection des droits de l'auteur et libre circulation de cette œuvre. Les ressources placées sous licence CC ne sont pas « libres de droits », mais leur utilisation est facilitée, dans un cadre non commercial essentiellement : **l'auteur choisit, à travers les six licences existantes, les libertés qu'il donne a priori au public.**

Les données et ressources disponibles sur Internet sont aussi soumises à d'autres droits.

Le droit à l'image

L'article 9 du Code civil protège la vie privée et donne à chacun le droit à la protection de son image. Il faut donc obtenir l'autorisation (écrite) d'une personne ou celle des parents pour les mineurs, pour utiliser son image. L'article 226 du Code pénal punit toute personne qui, sans consentement, fixe, enregistre ou transmet l'image d'une personne se trouvant dans un lieu privé, ainsi que les montages qui pourraient être réalisés. Ces dispositions s'appliquent à l'ensemble de la sphère Internet et à toute mise à disposition en ligne, le courrier privé y échappant seul.

Le droit de la presse et de l'audiovisuel

Il s'applique en particulier aux blogs et à tous sites à caractère éditorial. Outre l'éventuelle déclaration à la CNIL, les blog-

geurs et éditeurs de sites doivent s'appliquer à respecter le droit à la vie privée et ne pas mettre en cause des personnes dans un sens qui pourrait être considéré comme diffamatoire ou injurieux. S'appliquent également les dispositions des lois réprimant la diffusion de messages pornographiques ou violents susceptibles d'être vus par des mineurs, racistes, xénophobes ou homophobes, ou encore niant la réalité des génocides.

Les liens ou les hyperliens

S'il est admis que chaque site présent sur le Web accepte implicitement l'établissement de liens hypertextes simples, cela ne vaut pas pour les hyperliens profonds (qui atteignent directement une page sans passer par l'accueil du site) lorsqu'ils dénaturent le contenu ou l'image du site cible. Il est prudent de demander systématiquement l'autorisation du webmaster du site vers lequel un lien est établi.



- **Association des professionnels de la documentation : Accéder à la doc professionnelle > Droit de l'information > Actualités [en ligne].**

<http://www.adbs.fr>

- **Commission Nationale de l'Informatique et des Libertés [en ligne].**

<http://www.cnil.fr>

- **Creative Commons [en ligne].**

<http://fr.creativecommons.org>

- **Educnet. Guides juridiques Legamedia [en ligne].**

<http://www.educnet.education.fr/legamedia/>

- **Forum des droits sur l'Internet [en ligne].**

<http://www.foruminternet.org>

- **Forum des droits sur l'Internet. Hyperliens : statut juridique [en ligne].**

<http://www.foruminternet.org/specialistes/concertation/recommandations/recommandation-du-forum-des-droits-sur-l-internet-hyperliens-statut-juridique.html>

Le téléchargement

En soi, le téléchargement de fichiers (textes, images, sons ou vidéo) est un acte technique sur lequel le droit n'a pas prise, dès lors que les données récupérées le sont pour un usage privé. Cependant, la DADVSI considère que le téléchargement

avec remise à disposition (par exemple *via* des sites *peer to peer* ou sur des plateformes communautaires) est un délit. En outre, la question de la licéité de la source téléchargée se pose. De même, numériser des CD audio en MP3 pour distribuer librement les fichiers sur Internet est illégal.

Les déclarations de site web

La création d'un site Web ou d'un blog nécessite, pour être dans la plus stricte légalité, une déclaration à la CNIL (Commission Nationale Informatique et Libertés), quand le site comprend des « fichiers nominatifs », c'est-à-dire incluant ou demandant des informations sur des personnes. Dans les autres cas,

pour la plupart des sites éducatifs ou pédagogiques, mais aussi les blogs professionnels ou institutionnels, un formulaire a été spécialement conçu par la CNIL (disponible en ligne sur le site <http://www.cnil.fr>). Les blogs à caractère personnel sont dispensés de déclaration.

- **Legalis.net [en ligne].**

<http://www.legalis.net>

- **Legifrance [en ligne].**

<http://www.legifrance.gouv.fr>

- **Ministère de la Culture et la Communication. *Droit d'auteur, droits voisins dans la société de l'information - DADVSI [en ligne].***

<http://www.culture.gouv.fr/culture/actualites/index-droits05.html>

- **Organisation mondiale de la propriété intellectuelle. *Traité de l'OMPI sur le droit d'auteur [en ligne].***

<http://www.wipo.int/treaties/fr/ip/wct/index.html>

- **Organisation mondiale de la propriété intellectuelle. *Traité de l'OMPI sur les interprétations et exécutions et les phonogrammes [en ligne].***

http://www.wipo.int/treaties/fr/ip/wppt/trtdocs_wo034.html

- **Search by creative commons [en ligne].**

<http://search.creativecommons.org>



Typologie et usage des informations et documents scientifiques

Auteur :

David AYMONIN

Conservateur des bibliothèques
École Polytechnique Fédérale de Lausanne, Suisse

- 46** Nature des documents scientifiques
- 46** Identification des documents scientifiques
- 47** Usage des documents, selon leur type



L'information scientifique est l'ensemble des données produites, testées, acceptées, reconnues par la communauté scientifique internationale : c'est le substrat sur lequel s'appuient les nouvelles recherches pour avancer sans avoir à redémontrer. C'est le fait de matérialiser l'information en la déposant sur un support – même électronique – qui la fait devenir document.

Nature des documents scientifiques

Ce document, fruit de l'information produite par les chercheurs au moment où ils exploitent les résultats de leurs expériences et de leurs recherches, est destiné à d'autres chercheurs. Ces derniers le liront et en utiliseront les principales informations pour mener de nouvelles recherches et... publier de nouveaux résultats dans de nouveaux documents.

La publication de ces documents est souvent assistée par des partenaires chargés de la collecte et de la diffusion. Ce sont des éditeurs commerciaux (*publishers*) – comme Elsevier, Wiley, Hermès – ou des sociétés savantes (*Learning societies*), comme par exemple l'*American Chemical Society*, l'IEEE ou l'Académie des sciences.

Les documents de recherche subissent généralement un processus d'évaluation

par les pairs (*peer review*), généralisé dans le milieu académique. Ainsi, les revues scientifiques sont dotées d'un comité éditorial, composé d'experts, qui sélectionne et améliore les articles soumis grâce à l'avis de rapporteurs.

Parallèlement à ce système, le facteur d'impact (*impact factor*) des revues est censé mesurer l'importance d'une revue scientifique. Calculé chaque année par l'Institute for Scientific Information (ISI, devenu Thomson Reuters) et bien que souvent remis en cause, il sert souvent de critère d'évaluation des revues et des institutions. De nouveaux indicateurs permettent d'évaluer plus finement les chercheurs eux-mêmes, comme l'indice de Hirsch, ou H-Index.

Identification des documents scientifiques

La majeure partie des documents scientifiques est indexée par des bases de données bibliographiques ou par des moteurs de recherche, appelées sources d'information secondaires,

accessibles gratuitement ou sur abonnement, selon les cas, mais désormais presque toujours *via* le Web, leur contenu est très variable d'un domaine à l'autre et d'un produit à l'autre. La

- **POCHET Bernard. *Méthodologie documentaire : rechercher, consulter, rédiger à l'heure d'internet ?* 2^e éd. Bruxelles : De Boeck, 2005 (Sciences et méthodes).**

recherche sur le terme « virus » dans PubMed (base biomédicale), Google Scholar (indexation d'une partie du Web « scientifique »), Francis (base de sciences

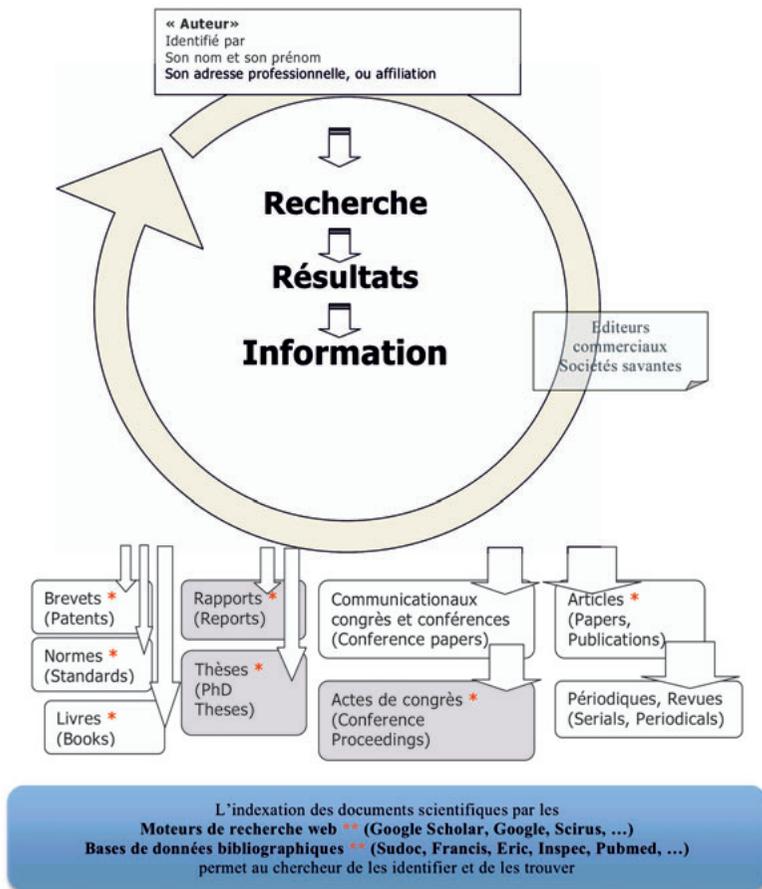
humaines et sociales) ou Derwent (base de brevets) donnera donc des résultats très différents, en nombre comme en types de publications signalées.

Usage des documents, selon leur type

Tous les types de documents scientifiques n'ont pas le même intérêt pour un étudiant de premier cycle ou pour un doctorant, car leur contenu et leur public cible sont très différenciés.

Schématiquement, on peut les classer selon qu'ils servent :

- A. à vérifier ou développer une connaissance ;
- B. à suivre l'actualité scientifique ;
- C. à faire le point le plus complet possible sur un sujet de recherche.



Littérature grise : documents diffusés hors des circuits commerciaux.

* source d'information primaire : contient le texte et les illustrations des auteurs

** Source d'information secondaire : donne les références de documents primaires

A - Sources d'information et documents servant à connaître ou vérifier une définition, une valeur, un terme

Chaque jour je dois trouver rapidement un ou une	Type de document utile (papier ou électronique)
Fait, date, donnée, définition, traduction	dictionnaire, encyclopédie, dictionnaire technique, dictionnaire de langue, table de données ou constantes, chronologie, etc.
Méthode ou technique	manuel pratique, série technique
Théorie ou information de base	manuel d'enseignement

B - Sources d'information et documents servant à suivre l'actualité scientifique

Régulièrement je m'informe sur les développements	Type de document utile (papier ou électronique)
De la science et de la technologie	Je lis régulièrement les articles de <i>La recherche</i> , <i>Pour la science</i> , <i>Scientific american</i>
De la recherche	Je parcours les revues <i>Nature et Science</i>
De ma discipline de recherche	Je lis régulièrement quelques revues scientifiques spécialisées
De mon sujet de Master ou de thèse	Je recherche les articles de revues et les papiers de conférence récemment publiés. Pour ne pas perdre de temps, je les cherche dans des bases de données bibliographiques multidisciplinaires rapidement mises à jour, comme <i>Web of Science</i> ou <i>Scopus</i>
Du travail de mes collègues ou concurrents (et je les rencontre)	En participant aux congrès et colloques de ma discipline. Je m'intègre progressivement au « collège invisible » composé des chercheurs et experts de mon domaine

C - Sources d'information et documents de recherche scientifique
servant à mener des recherches d'information exhaustives

À intervalles réguliers ou à des moments particuliers de mes études	Type de document utile (papier ou électronique)
Je veux faire l'état de l'art sur un sujet de recherche nouveau pour moi	Les livres scientifiques, trouvés dans les catalogues de bibliothèques, vont me permettre d'accéder à l'essentiel des résultats de recherche et des connaissances, sous une forme assez synthétique
Je veux être sûr-e de n'avoir manqué aucune publication importante sur mon sujet de recherche - avant de soutenir ma thèse ou mon mémoire - pour mettre ma bibliographie à jour avant de soumettre un article	Les bases de données spécialisées par domaine, comme PubMed (biomédical), INSPEC (physique, mécanique, informatique), Francis (sciences humaines et sociales), Psyclit (psychologie), indexent la plupart des types de documents de recherche, issus de nombreux pays, sur une longue période



- **DOC'INSA. SAPRISTI ! Sentiers d'Accès et Pistes de Recherche d'Informations Scientifiques et Techniques sur l'Internet ! [en ligne].**

<http://docinsa.insa-lyon.fr/sapristi/>

- **Infosphère [en ligne].**

<http://www.bibliotheques.uqam.ca/InfoSphere/>

- **Intute [en ligne].**

<http://www.intute.ac.uk/>

- **SUDOC - Système universitaire de documentation [en ligne].**

<http://www.sudoc.abes.fr/>

- **WIKIPEDIA. Bibliométrie et scientométrie [en ligne].**

<http://fr.wikipedia.org/wiki/Bibliométrie>

Les supports de l'information

Auteur :

Élisabeth NOEL
Conservateur des bibliothèques – enssib

- 52** Du papyrus au numérique
- 53** Vers une dématérialisation de l'information



Des tablettes d'argiles en Mésopotamie aux supports numériques actuels, les hommes ont toujours trouvé des moyens pour conserver les informations sur des supports, ceux-ci évoluant en parallèle avec les technologies.

Du papyrus au numérique

Les supports de l'information se différencient par :

- leur mode de lecture (directe – livre, parchemin – ou indirecte, avec une machine – supports optiques ou magnétiques) ;
- leur mode de consultation (en local ou à distance) ;
- leur alimentation ;
- leur gestion (centralisée ou éclatée).

Ainsi, les premiers supports (**papyrus, papiers**) se lisaient rapidement et facilement, mais leur mise à jour nécessitaient une réécriture des données, d'où les nombreuses rééditions d'ouvrages.

L'apparition de la **photographie** a permis de conserver des clichés des supports écrits sur microfiche, ce qui permettait une meilleure diffusion de l'information et un gain de place. La consultation de ces **microformes** rendait nécessaire l'utilisa-

tion d'appareils de lecture spécifiques, peu conviviaux et peu répandus, disponibles sur les lieux de stockages des microformes (bibliothèques).

Les **supports magnétiques** ont favorisé le développement et l'accès à l'information. L'enregistrement du son, des images ou de données numériques s'est d'abord effectué sur bandes magnétiques (bobines, cassettes ou cartouches), supports qui permettent toujours de nos jours une conservation fiable, économique et à faible risque.

Les **disques magnétiques**, durs ou souples (disquettes), servent au stockage des données informatiques, avec des puissances de plus en plus importantes. L'accès aux données est sectoriel, le disque tournant autour de son centre, et le temps d'accès à l'information est assez rapide.

- COMBEROUSSE Martine. *Histoire de l'information scientifique et technique*. Paris : Nathan, 1999 (Collection 128 ; Information-documentation).
- FAYET-SCRIBE Sylvie. *Chronologie des supports, des dispositifs spatiaux, des outils de repérage de l'information*. Solaris, 1997, n° 3 [en ligne]. http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d04/4fayet_0intro.html
- MCLUHAN Marshall. *La galaxie Gutenberg face à l'ère électronique*. Paris : Gallimard, 1977.
- SCHAER Roland. *Tous les savoirs du monde, encyclopédies et bibliothèques de Sumer au xx^e siècle*. Paris : Bibliothèque nationale de France ; Flammarion, 1996.



La **clé USB** remplace maintenant la disquette : petite et maniable, elle peut stocker de 256 Mo à plusieurs Go selon les modèles. Elle nécessite un simple port USB pour être utilisable.

À ces supports magnétiques s'ajoutent des **supports optiques**. Les disques optiques font en général 12 centimètres de diamètre. Les informations, enregistrées sous forme numérique, sont consultées par un rayon laser, sans contact

avec la surface, donc sans usure. Les CD (Compact Disc) et les DVD (Digital Versatile Disc) sont les deux principaux formats de supports optiques.

Classiquement, la capacité de stockage de l'information s'accroît en même temps que diminue la taille des supports. Cependant, ces supports restent fragiles et nécessitent toujours des outils de lecture adaptés, qui eux-aussi évoluent sans cesse.

Vers une dématérialisation de l'information

Le développement des réseaux de communication, d'abord avec le Minitel en France, puis spécialement avec l'essor d'Internet, a transformé la recherche, la consultation et le stockage de l'information. En effet, le couplage des réseaux informatiques des universités, des grandes écoles, des entreprises avec les infrastructures de télécommunications permet aujourd'hui d'utiliser à distance les ressources locales produites par ces établissements.

Le succès mondial d'Internet ne doit pas faire oublier que des réseaux de ce type ont

commencé à se développer dans l'enseignement et la recherche dès les années 80.

Ces réseaux de communications établissent une distance de fait entre le support de l'information et l'utilisateur, distance qui induit un sentiment de dématérialisation. Si l'information, stockée sur le disque dur de l'ordinateur serveur consulté, peut être consultée ou recopiée un nombre infini de fois à distance, elle peut aussi être modifiée, voire disparaître du serveur. C'est dans ce sens que l'on parle d'informations virtuelles.



- **Archives-ouvertes.fr [en ligne].**

<http://hal.archives-ouvertes.fr/>

- **Conservation préventive du patrimoine documentaire. Documents sur supports magnétiques [en ligne].**

http://www.culture.gouv.fr/culture/conservation/dswmedia/fr/all_magn.htm

- **Dataligence. Dossier spécial stockage des données [en ligne].**

<http://www.dataligence.com/site/dossier-special-stockage-donnees.html>

- **École nationale des ponts et chaussées - Service documentaire. L'Open Access [en ligne].**

http://www.enpc.fr/fr/documentation/doc_electronique/dossier_openaccess.htm

Ainsi, des publications électroniques en ligne (comme des revues en ligne, des pages d'accueil de site Web, des échanges dans le cadre d'un forum de discussion) ne sont consultables que par ordinateur connecté à Internet, et ne renvoient pas toujours à un équivalent sous forme imprimée (certaines revues n'existent qu'en ligne).

Le *cloud computing* se développe maintenant, d'autant plus que les équipements informatiques et les débits d'accès le facilitent : les données et logiciels, auparavant stockés dans les ordinateurs des utilisateurs, sont maintenant répartis sur des serveurs du réseau. C'est ce qui se passe en définitive avec le stockage de nombreuses données en ligne via les blog ou

les réseaux sociaux, et avec l'offre de logiciels utilisables en ligne, comme la suite logicielle proposée par Google Doc, <http://docs.google.com>.

Après une période où l'enjeu était d'être connecté au réseau Internet, puis ensuite d'être présent, visible sur Internet, la problématique actuelle réside plus dans la question de l'intérêt des services offerts, mais aussi dans la mise à disposition des ressources, tout particulièrement dans le cadre de l'**Open Access Initiative**, qui veut permettre de constituer des réservoirs d'archives électroniques des communications scientifiques autour de protocoles techniques facilitant la recherche et l'échange.



- **INIST. Libre accès à l'information scientifique et technique [en ligne].**

<http://openaccess.inist.fr/>

- **Initiative de Budapest pour l'Accès Ouvert [en ligne].**

<http://www.soros.org/openaccess/fr/index.shtml>

- **MORELLE Marc-Alexis. Les médias, mémoire de l'humanité [en ligne].**

<http://www.malexism.org/medias/introduction.html>

- **Open Archives Initiative [en ligne].**

<http://www.openarchives.org/>

Savoir rechercher et interroger :

les repères méthodologiques

Auteur :

Bernard POCHET

Bibliothécaire en chef, maître de conférence
Faculté universitaire de Gembloux, Belgique

- 56** Analyser les besoins
- 57** Le choix des sources et des outils documentaires
- 57** La préparation de la question documentaire
- 58** L'interrogation proprement dite
- 59** L'évaluation des réponses



La littérature scientifique peut prendre de multiples formes et présenter des niveaux de complexité variables, mais répond à un besoin unique, celui de partager l'information à des fins de simple information, de formation ou d'échange.

La recherche documentaire donne des moyens multiples et parfois complexes pour accéder à ces écrits. Depuis l'apparition du phénomène du libre accès, cette complexité s'est encore accrue, avec l'apparition de nouveaux outils. L'attitude critique est d'autant plus nécessaire que l'accessibilité à une partie de cette littérature, parfois noyée dans des ensembles totalement hétérogènes, s'est aussi accrue. La recherche documentaire, pour être efficace, efficiente et exhaustive doit être systématique et méthodique. C'est essentiellement un travail étape par étape, partant du sens pour retourner au sens : elle doit permettre de donner une réponse à une question.

Analyser les besoins

L'étape du questionnement est essentielle pour bien cerner le sujet de la recherche. Il convient :

- de bien analyser le sujet de la recherche ;
- de consulter des ouvrages de référence (dictionnaires, encyclopédies...);
- de consulter d'autres documents sur le même sujet.

De l'objectif du travail à réaliser ou du besoin d'information identifié dépendra le choix des documents à trouver... d'un livre pour réaliser un travail de premier cycle à plusieurs articles édités dans des revues scientifiques pour un mémoire de maîtrise.

56



- DARROBERS Martine, LE POTTIER Nicole. *La Recherche documentaire*. Paris : Nathan, 2005.
- MESGUICH Véronique, THOMAS Armelle. *Net recherche : le guide pratique pour mieux trouver l'information utile*. Paris : ADBS édition, 2010.
- MORIZIO Claude. *La Recherche d'information*. Paris : ADBS, 2004 (Collection 128 ; Information-documentation).
- PIOLAT Annie. *La Recherche documentaire : manuel à l'usage des étudiants, doctorants et jeunes chercheurs*. Marseille : Solal, 2002.
- POCHE Bernard. *Méthodologie documentaire : comment accéder à la littérature scientifique à l'heure d'Internet ? 2^e éd.* Bruxelles : De Boeck, 2005 (Sciences et méthodes).

Le choix des sources et des outils documentaires

À côté de la littérature conventionnelle – documents disponibles dans le commerce – et de la littérature non conventionnelle – littérature grise (mémoires, thèses, rapports...) –, il existe aussi des documents en accès libre, de tous types.

Sélectionner les bons outils de recherche est alors essentiel. Pour la recherche d'un livre, la consultation d'un catalogue suffira. Mais alors, tout comme pour la consultation de moteurs de recherches généralistes (Google, Yahoo!...), il ne s'agit pas à proprement parler de recherche documentaire.

Les outils de recherche documentaire, qui fournissent des références de documents (avec éventuellement des liens vers les

documents au format électronique), peuvent être classés ainsi :

- les bibliographies analytiques spécialisées, domaine par domaine, qui utilisent un vocabulaire spécialisé (Medline...);
- les bases de sommaires, bâties sur les tables des matières d'un grand nombre de titres de périodiques scientifiques (Current Contents, article@INIST, Ingenta...);
- les outils de « récolte » de (méta)données dans les dépôts institutionnels ou les sites d'archives de documents en accès libre (OAIster.org...);
- les portails d'accès à la littérature scientifique qui intègrent plusieurs sources et bases de données différentes (ScientificCommons, Google Scholar...).

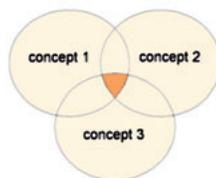
La préparation de la question documentaire

Le point de départ consiste à transcrire la question initiale en une phrase simple reprenant tous les concepts. Dans cette phrase, il est nécessaire de repérer les champs conceptuels, les thématiques présentes : **quel est l'objet, l'organisme, l'individu, le groupe étudié ? Quel est le sujet de la recherche ? Quelle est la technique, la méthode de recherche utilisée ? Quel est l'espace géographique concerné ?...** Pour chaque champ conceptuel, il faut ensuite trouver tous les termes, mots-clés ou descripteurs susceptibles d'être utilisés. Les thésaurus attachés aux bibliographies doivent nécessairement être utilisés, ainsi que, pour les outils qui ne proposent pas de vocabulaire spécifique, des dictionnaires, pour envisager tous les synonymes possibles, au moins en anglais et en français.

Tableau des concepts et termes associés

Concept 1	Concept 2	Concept 3
terme1_1	terme2_1	terme3_1
terme1_2	terme2_2	terme3_2
terme1_3	...	

Une fois tous les termes rassemblés, concept par concept, la rédaction de la question documentaire



consiste à lier ces termes avec des opérateurs booléens. Les termes relatifs à un même concept sont reliés avec l'opérateur « OU » afin de retrouver un maximum de documents traitant de ce

concept. Les concepts sont ensuite croisés entre eux avec l'opérateur « ET » afin de déterminer le « point focal », le sous-ensemble contenant en principe les documents qui répondent à la question posée.

La question documentaire prendra alors la forme : (terme1_1 OU terme1_2 OU terme1_3) ET (terme2_1 OU terme2_2) ET (terme3_1 OU terme3_2). Lorsque c'est possible, la recherche étape par étape, concept par concept, est préférable.

L'interrogation proprement dite

D'une manière générale, les termes sont écrits sans accent et en lettres minuscules. Les systèmes documentaires sont en général insensibles aux lettres accentuées et à la casse (minuscules/majuscules).

Le « mode simple » avec une seule zone de saisie, sans combinaison possible, est en général trop frustrant pour permettre une réelle recherche documentaire ; le « mode avancé », avec plusieurs zones de saisie et plusieurs combinaisons possibles, et le « mode expert », où une question documentaire complexe peut être introduite en mode direct, proposent plusieurs outils :

Les opérateurs booléens sont au nombre de 3. Avec le « ET » (« AND » en anglais) comme opérateur, tous les termes liés doivent être présents dans la réponse. Avec le « OU » (« OR » en anglais), au moins un terme doit être présent. Le « SAUF » (« NOT » en anglais) élimine tous les documents qui contiennent le terme qui suit l'opérateur.

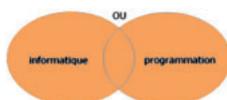
Certains outils remplacent l'espace entre 2 termes par le ET (opérateur par défaut), d'autres acceptent les signes « + » pour

le « ET », le « - » pour le « SAUF » et le « pipe », « | » pour le OU. Il faut à tout moment vérifier les options adoptées par le système documentaire.

Les opérateurs de proximité permettent d'imposer que 2 termes soient contigus ou proches. L'opérateur le plus fréquent est « NEAR » ; il peut être complété par un nombre qui indique le nombre de mots qui séparent les 2 termes (exemple : pomme NEAR1 terre). Une solution alternative pour forcer la recherche d'une expression exacte est d'utiliser les guillemets (exemple : « pomme de terre »).

La troncature permet de réduire la frappe à la racine du terme. Par exemple « energ* » donnera « énergie », « énergétique » ou « energy ». Il faut cependant être attentif à ne pas utiliser de racine trop courte, au risque d'obtenir des réponses inattendues (dans cet exemple : « éner-gumène »). Certains systèmes acceptent les troncatures gauche (« *ution ») ou centrale (« poly*ation »).

Les parenthèses sont utilisées pour organiser les termes et les combinaisons.



L'**historique** permet de contrôler le déroulement de la recherche et la combinaison de questions. Cela donnerait par exemple :

\$1	terme1_1
\$2	terme1_2
\$3	terme1_3
\$4	\$1 OR \$2 OR \$3
\$5	terme2_1
\$6	terme2_2
\$7	\$5 OR \$6
\$8	terme3_1
\$9	terme3_2
\$10	\$8 OR \$9
\$11	\$4 AND \$7 AND \$10

La **consultation des index et des listes d'autorité** permet de vérifier la présence et l'orthographe d'un terme ou d'un nom d'auteur dans le système documentaire.

Les **liens hypertextes** proposent, à tout moment, des rebonds pour poser une nouvelle question (toutes les notices comprenant un nom d'auteur, toutes les notices d'articles édités dans un périodique...).

La **fonction « documents connexes »** (« related records ») permet de trouver d'autres documents considérés comme proches par le système documentaire.

L'évaluation des réponses

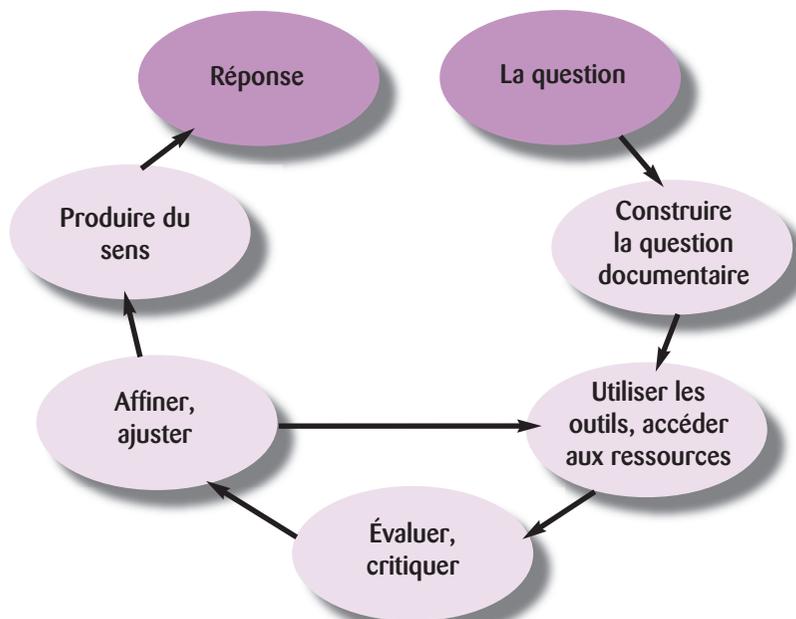
Après avoir interrogé un système documentaire, il faut enfin vérifier la pertinence des réponses (répondent-elles à la question posée ?) et adapter la question si nécessaire. Deux problèmes peuvent être rencontrés :

- **le bruit** : un trop grand nombre de réponses à cause d'une question trop large (termes trop génériques, troncatures trop courtes, trop de termes...);
- **le silence**, beaucoup plus difficile à détecter : la recherche documentaire apporte un nombre de réponses inférieur à ce qui aurait dû être trouvé, si la question avait été correctement posée. La base de données contient donc d'autres réponses pertinentes. Il faut ajouter des termes et des synonymes (langue anglaise, noms locaux ou latins pour les organismes...), utiliser des tron-

catures (pour envisager le pluriel des termes...)...

La qualité des réponses doit aussi être évaluée avec une attitude critique indispensable. Le fond et la forme des documents permettent de déterminer l'objectif, et le niveau scientifique, afin de connaître la crédibilité à y porter. Un article publié dans une revue de vulgarisation, écrit par un journaliste, sans référence bibliographique, aura nécessairement moins de crédit qu'un article écrit par un scientifique affilié à une université et publié dans une revue scientifique réputée.

Une fois les références des documents (notices) sélectionnées, elles doivent impérativement être sauvegardées (logiciel de gestion de références par exemple), pour obtenir ensuite (bibliothèque, Internet, auteur...) les documents sélectionnés.



© Bernard Pochet, 2008

Boucle de recherche d'information



- Bibliothèque nationale de France. *Guide de recherche en bibliothèque : définir le domaine de recherche* [en ligne].

http://grebib.bnf.fr/html/definir_domaine.html

- Faculté universitaire des sciences agronomiques de Gembloux. *Métho.DOC. Chercher pour trouver* [en ligne].

<http://www.bib.fsagx.ac.be/edudoc/Metho.DOC/>

Les langages documentaires

Auteur :

Muriel AMAR

Conservateur de bibliothèques
Bibliothèque nationale de France

- 62** Qu'est-ce qu'un vocabulaire contrôlé ? À quoi sert-il ?
- 63** Types de vocabulaires contrôlés et fonctions
- 64** Comment utiliser les vocabulaires contrôlés ?



Comment trouver des documents sur un sujet ?
 Comment distinguer les documents écrits par un auteur des documents écrits sur cet auteur ?
 Comment identifier tous les documents portant sur un même thème, quels que soient les mots utilisés pour exprimer ce thème ?
 Pour cela, les catalogues et les bases de données proposent des rubriques, aux noms variés (sujets, mots-clés, matières, descripteurs, thèmes).

À l'inverse de l'index des moteurs de recherche sur le Web, constitué de mots extraits des pages (texte intégral), ces index sont créés avec des mots choisis, dans un lexique prédéfini, par des professionnels de l'information, à la

suite d'une analyse du contenu du document. Cette **opération d'indexation** a pour objectif de réunir des documents proches par leurs thématiques, au-delà des variations linguistiques ou conceptuelles.

Qu'est-ce qu'un vocabulaire contrôlé ?

À quoi sert-il ?

Le vocabulaire utilisé dans l'indexation est contrôlé, ce qui permet de gérer lors des recherches :

- la **synonymie** : le choix d'un terme, retenu parmi plusieurs portant sur un même concept, *comme logement, pour maison, immeuble, logis, habitation* pour décrire les contenus lors de l'indexation,

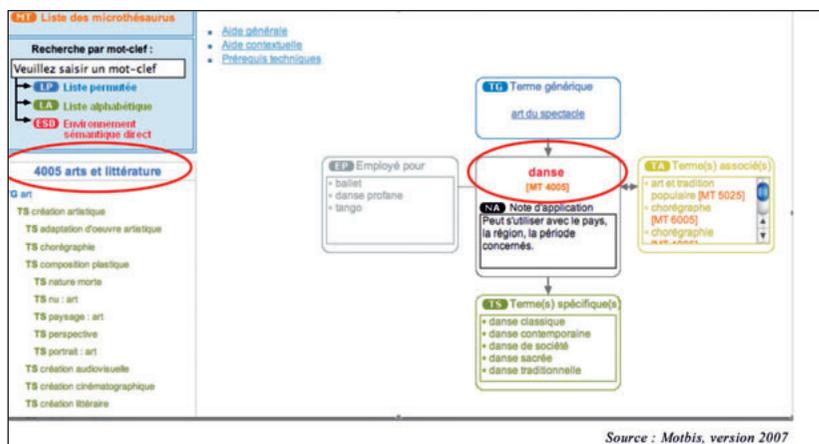
garantira de retrouver toutes les occurrences de la notion lors de la recherche ;
 - l'**homonymie** : inversement, certains mots identiques de la langue désignent des réalités différentes, *par exemple, le mot souris peut renvoyer soit au rongeur soit au périphérique informatique*. Des formes distinctes seront alors choisies.

- DEGEZ Danièle, MENILLET Dominique. *Le Thésauroglossaire des langages documentaires : un outil de contrôle sémantique*. Paris : ADBS, 2001.
- LEFÈVRE Philippe. *La recherche d'informations. Du texte intégral au thesaurus*. Paris : Hermès, 2000.
- MANIEZ Jacques. *Actualité des langages documentaires : fondements théoriques de la recherche d'information*. Paris : ADBS, 2002.



Le vocabulaire contrôlé peut aussi aider à circonscrire l'étendue conceptuelle d'un sujet de recherche documentaire. Ainsi,

pour le terme *danse*, le vocabulaire contrôlé Motbis propose l'environnement sémantique suivant :



Environnement sémantique du terme Danse dans Motbis

La recherche peut être précisée par des **termes spécifiques** (*danse contemporaine par exemple*), ou élargie par un **terme générique** (*art du spectacle*). Motbis indique aussi les **termes associés**, comme *chorégraphie*. D'autres indications peuvent s'ajouter, comme l'équivalence

(contrôle de la synonymie, avec la formule *Employé pour*), la définition ou la catégorie. Ces relations, ajoutées à la règle d'univocité des termes (ni synonymes ni homonymes), distinguent cruciallement le mot de la langue du **descripteur**, **mot du vocabulaire contrôlé**.

Types de vocabulaires contrôlés et fonctions

Il existe deux grands types de vocabulaires contrôlés, appelés aussi **langages documentaires** :

- Les langages documentaires **analytiques**, pour *décrire* les différentes composantes thématiques d'un document, dont font partie les **thésaurus** ;
- Les langages documentaires **synthétiques**, qui permettent de **situer** les différents thèmes d'un document, avec par exemple les **listes de mots matières** comme **Rameau** et les **classifications**, comme la **classification universelle Dewey**.

Complémentaires dans leur approche du contenu d'un document, l'utilisation concomitante de ces types de langages multiplie, lors de la recherche, les possibilités d'accès à l'information. Ainsi, un ouvrage sur le *dépistage des enfants sourdoués à l'école élémentaire* pourra-t-il être trouvé à la fois :

- dans une hiérarchie disciplinaire qui mènera à l'indice **371.952** de la classification Dewey, indice qui **situe** l'ouvrage dans le contexte suivant :

3 Sciences sociales

37 Éducation

371 Organisation de l'éducation

371.9 Enseignement spécialisé

371.95 Élèves surdoués

371.952 Dépistage des élèves surdoués

• par le sujet principal, à travers la vedette-matière Rameau **Enfants surdoués-Education-France**, composée selon un ordre prédéfini et obligatoire de différents éléments : une tête de vedette exprimant l'essentiel du sujet (ici *Enfants surdoués*) et des sous-vedettes apportant des précisions de points de vue, de lieu, éventuellement de temps et de forme. L'ensemble de ces éléments constitue une

vedette matière qui permet, elle aussi, de **situer** le document par rapport à un sujet plus général (ici *Enfants surdoués*).

• par un ensemble de composants thématiques autonomes, choisis dans un thésaurus (ici Motbis) : *élève surdoué, école élémentaire, test, apprentissage précoce*. Chacun de ces descripteurs décrit un aspect du sujet traité, décomposé en concepts indépendants.

Comment utiliser les vocabulaires contrôlés ?

Les langages documentaires, bien que puissants pour la recherche documentaire, restent globalement peu visibles et par conséquent assez largement méconnus. Pourtant de nombreux catalogues de bibliothèques et bases de données mettent leurs vocabulaires contrôlés en ligne, *via* un accès « thésaurus » ou « lexique » proposé dans les rubriques dédiées à la recherche-sujet.

En outre, les notices de résultats de catalogues et de bases de données proposent

des termes « cliquables », issus du vocabulaire contrôlé, qui permettent de relancer la recherche automatiquement. Ainsi, une recherche commencée par la rubrique *Mots du titre* pourra être affinée ensuite par les éléments « cliquables » du vocabulaire contrôlé dans les premières réponses, permettant une **navigation « sémantique »** de notices en notices en attendant... les ontologies et le Web sémantique, ce Web de demain qui devrait aider l'utilisateur à exploiter les affinités thématiques entre documents...

- **Bibliothèque nationale de France. RAMEAU (Répertoire d'autorité-matière encyclopédique et alphabétique unifié) [en ligne].**

<http://rameau.bnf.fr/>

- **DALBIN Sylvie. Descripteurs [en ligne].**

<http://descripteurs.net>

- **DUVAL Marc. Le Mot-clé [en ligne].**

<http://www.dsi-info.ca/mot-cle.html>

- **OCLC Deweybrowser [en ligne].**

<http://deweyresearch.oclc.org/ddcbrowser/wcat>



Évaluer l'information sur Internet

Auteur :

Alexandre SERRES

Maître de conférences – URFIST de Rennes

- 66** Enjeux et difficultés
- 67** Démarche et questionnements



Évaluer un document a toujours été une opération complexe par nature, mettant en jeu de nombreuses connaissances et compétences : connaissance minimale de l'édition, des acteurs et des circuits de l'information scientifique pour juger de la fiabilité, de la crédibilité d'une source ou d'un auteur, connaissances disciplinaires pour évaluer la validité même d'une information.

Internet accentue encore ces difficultés.

Enjeux et difficultés

Mesurer les enjeux

L'évaluation de l'information est devenue une question cruciale pour plusieurs raisons :

- **l'absence de filtre éditorial** : Internet est un système de publication totalement ouvert, où n'importe qui peut publier et diffuser n'importe quel type d'information, contrairement aux revues scientifiques ou au monde de l'édition.

- en raison de ce caractère ouvert, **les info-pollutions** explosent sur Internet : rumeurs, manipulations, information médiocre, invasion du spam (sur les messageries, les blogs, les réseaux sociaux), prolifération des sites racistes, sectaires, négationnistes, etc.

- **la remise en cause des médiateurs de l'information** (éditeurs, biblio-

thèques...), c'est à l'utilisateur désormais d'évaluer et de valider l'information, de vérifier sa fiabilité.

- **le manque de formation** : savoir évaluer l'information implique, non seulement une bonne culture générale mais aussi l'acquisition d'une véritable culture informationnelle, loin d'être développée actuellement.

Percevoir les difficultés de l'évaluation

Internet accentue les difficultés propres à l'évaluation, par différents « brouillages » de frontières : la distinction entre support, contenu, mise en forme, graphisme n'est pas évidente ; les rôles et compétences des acteurs ne sont pas toujours identifiés (certains textes peuvent être enrichis

- SERRES Alexandre. *Évaluation de l'information sur Internet*. Bulletin des bibliothèques de France, 2005 - t. 50, n° 6 [en ligne]. <http://bbf.enssib.fr/consulter/bbf-2005-06-0038-006>

à la fois par l'auteur et les lecteurs, etc.) ; enfin, tous les genres documentaires existent en ligne et se ressemblent, le numérique aplatit les différences entre les documents, l'identification du type de ressource devient une véritable difficulté. Aussi, l'é-

valuation de l'information n'est-elle pas une opération « technique » : il n'existe pas de recette toute prête pour évaluer la fiabilité d'un site ou d'un document.

Démarche et questionnements

Évaluer l'information sur Internet revient à questionner les résultats trouvés, dans une démarche en trois temps.

Bien savoir ce qu'on cherche

Tout d'abord, pour juger si une ressource est pertinente ou non, il faut cerner précisément ses besoins : quel type d'informations (textes, images, données) est recherché ?, sur quel sujet précis, dans quel contexte (universitaire ou non), pour quel objectif (faire un dossier ou un exposé) ?

Chercher à identifier la ressource

Ensuite, identifier soigneusement la ressource en se posant **trois questions simples** :

- **par qui ce document est-il publié ?**
quel est l'éditeur, le site hôte, sa nationalité ? quelle est la nature du site, ses objectifs, le public visé ? le décodage des URL et des noms de domaine s'avère ici une compétence précieuse.
- **par qui a-t-il été écrit ?**
qui est l'auteur, quel est son statut, son domaine de compétence, sa notoriété ? à qui s'adresse-t-il, pourquoi publie-t-il sur le Web ?



- **GINGRAS François-Pierre.** *L'utilisation des ressources d'Internet et les travaux scientifiques : questions d'appréciation.* Université d'Ottawa, 1997 [en ligne].

<http://aix1.uottawa.ca/~fgingras/metho/sources.html>

- **HoaxBuster.** *Première ressource francophone sur les canulars du Web* [en ligne].

<http://www.hoaxbuster.com/>

- **HUYGHE François-Bernard.** *Évaluer l'information sur Internet* [en ligne].

http://www.huyghe.fr/actu_519.htm

- **MOREAU Antoine.** *Scio : la Liste Noire* [en ligne].

<http://www.e-scio.net/noire.html>

- quel est le type de document ?

quel est son genre documentaire, sa nature (document scientifique ou de vulgarisation), ses limites (un site ou une partie de site) ? quand a-t-il été créé, mis à jour ?

Distinguer les objets de l'évaluation

Après l'identification, l'évaluation distinguera nettement ces **cinq objets différents** :

- **la source, l'auteur** : chercher à mesurer sa fiabilité, son accessibilité, sa notoriété (par qui est-il cité ?), son degré d'expertise dans le domaine. Des recherches peuvent permettre d'identifier son institution d'appartenance, ses réseaux de travail et d'affiliation.
- **le contenu** : quel est le degré de fiabilité de l'information, sa « fraîcheur », son actualisation (dont l'importance peut-être relative au thème), son caractère d'exactitude, la qualité de l'expression, la richesse des sources, la clarté d'indication des références bibliographiques, etc. ? Cette partie difficile demande des connaissances minimales sur le sujet.

- **l'organisation du site ou du document** : il convient de distinguer ce qui relève de l'organisation de l'information et de la navigation dans le site (structuration plus ou moins réussie) du contenu informationnel.

- **la mise en forme, le graphisme, le design, la lisibilité** : un très beau site au plan visuel peut être mal structuré, difficile à consulter (et l'inverse).

- **l'accessibilité, les aspects techniques** : enfin il faut distinguer les aspects purement techniques : rapidité de chargement des pages, choix des formats graphiques, présence ou non d'outils de recherche, logiciels nécessaires pour la consultation du site.

L'évaluation de l'information reste toujours une opération délicate et « non technique », il ne faut pas hésiter à solliciter le recours d'autres personnes expertes du domaine, enseignants, autres étudiants, mais aussi autres sites, autres documents. C'est la consultation et la confrontation d'avis et de sources différentes qui permettent la véritable évaluation d'une ressource.



- **PRUDHOMME Brigitte. *SAPRISTI ! Évaluation de l'information présente sur Internet* [en ligne].**

<http://docinsa.insa-lyon.fr/sapristi/index.php?rub=1004>

- **SERRES Alexandre. *Évaluation de l'information sur Internet* [en ligne].**

http://www.sites.univ-rennes2.fr/urfist/evaluation_information_ressources

- **UNIVERSITE DE MONTREAL (Québec), Bibliothèques des sciences de la santé. *L'évaluation d'un site Web* [en ligne].**

<http://guides.bib.umontreal.ca/disciplines/74-D-terminer-la-fiabilit-de-l-information?tab=276>

- **UNIVERSITE LAVAL (Québec), Faculté de Lettres. *ICARISH (Initiation critique aux ressources Internet en sciences humaines). Module 3 : Savoir évaluer et sélectionner l'information trouvée* [en ligne].**

http://www.icarish.ulaval.ca/guide/module_3/pre/pre_m3.html

Lutter contre le plagiat

Auteur :

Dominique MANIEZ
Responsable TICE – Université Lyon 2

- 70** Qu'est-ce que le plagiat ?
- 70** Le droit de citation
- 71** Règles pour une écriture éthique
- 72** Sanctions encourues



Internet a bouleversé les usages en matière de recherche d'information, mais a aussi accru et facilité les possibilités de copier-coller dans la production des écrits scientifiques. Diverses études récentes démontrent une méconnaissance assez générale des règles en matière de plagiat chez les étudiants, ainsi qu'une absence d'information claire des enseignants sur ce sujet.

Qu'est-ce que le plagiat ?

Définir précisément le plagiat n'est pas une chose facile. C'est un vol littéraire dont la faute est double : d'abord intellectuelle – il est malhonnête de s'attribuer le mérite d'un texte écrit par un autre – ; ensuite, morale, en spoliant les droits moraux et patrimoniaux d'un auteur dont l'œuvre est pillée. Le plagiat prend de nombreuses formes, de la simple cita-

tion sans guillemets à la recopie servile de paragraphes entiers, en passant par l'emprunt de la trame d'une histoire.

Les travaux de recherche scientifique s'appuient sur les articles déjà produits, à partir d'une étude de l'état de l'art, il est par conséquent parfaitement autorisé de citer un extrait d'œuvre sous réserve du respect du droit de citation.

Le droit de citation

La loi exige tout d'abord que l'œuvre ait été divulguée au public et que le droit de paternité de l'auteur soit respecté. Le nom de l'auteur et la référence du document doivent donc être clairement identifiés. La citation apparaîtra très clairement au sein de l'œuvre qui cite, par des signes distinctifs, par exemple des guillemets ou l'italique dans le cas d'un texte.

La citation ne doit pas porter atteinte au droit moral de l'auteur : elle doit être exactement reproduite en respectant la visée originale de l'auteur : ainsi, une citation tronquée ou sortie de son contexte afin de lui donner un sens autre que celui voulu par l'auteur serait condamnable.

Une autre exigence concerne la brièveté. La citation doit être proportionnellement

- MARTINEAU Yzabelle. *Le Faux Littéraire, plagiat littéraire, intertextualité et dialogisme*. Québec : Nota bene, 2002.
- MAUREL-INDART Hélène. *Plagiats, les coulisses de l'écriture*. Paris : La Différence, 2007.

courte par rapport à l'œuvre citée, mais également par rapport à l'œuvre citante. Ainsi, le fait de citer la moitié d'une œuvre, même extrêmement courte (un poème, une chanson, etc.) n'est pas convenable.

De même, le poids des citations à l'intérieur du travail ne doit pas être trop important. Une citation ne sert qu'à illustrer un propos, une thèse et ne doit pas sortir de cette finalité.

Règles pour une écriture éthique

La lutte contre le plagiat a commencé plus tôt dans les pays anglo-saxons qui ont été plus prompts à prendre conscience du phénomène du *plagiarism*. Il convient donc pour tout auteur :

- de toujours reconnaître les contributions des autres et les sources utilisées ;
- tout texte d'un autre auteur, recopié à

l'identique, doit être encadré par des guillemets ;

- toute source utilisée pour la rédaction, qu'elle soit paraphrasée, résumée ou citée entre guillemets, doit être clairement identifiée (citation de l'auteur, de la source...);
- la signification exacte de la pensée de l'auteur doit être restituée.



- **Inspection générale de l'administration de l'Éducation nationale et de la Recherche. *L'évaluation des étudiants à l'Université : point aveugle ou point d'appui ?* Juillet 2007 [en ligne].**

<http://www.education.gouv.fr/cid5592/l-evaluation-des-etudiants-a-l-universite-point-aveugle-ou-point-d-appui.html>

- **Internet : fraude et déontologie selon les acteurs universitaires [en ligne].**

<http://responsable.unige.ch>

- **ISAAC Henri. *L'université numérique* [en ligne].**

http://www.nouvelleuniversite.gouv.fr/IMG/pdf/Rapport_univ_num.pdf

- **MAUREL-INDART Hélène. *Le plagiat* [en ligne].**

<http://www.leplagiat.net/>

- **PERREAULT Nicole. *Le plagiat et autres types de triche scolaire à l'aide des technologies : une réalité, des solutions.* Profweb, 21 janvier 2007 [en ligne].**

<http://www.profweb.qc.ca/?id=2300>

Sanctions encourues

L'auteur d'un plagiat avéré risque trois types de condamnations : morale, pénale et disciplinaire, ce qui permet d'apprécier la gravité symbolique d'un tel comportement. Surtout, localement, l'auteur d'un plagiat risque des sanctions disciplinaires au sein de l'établissement où le délit a été commis. Les sanctions disciplinaires applicables aux usagers des établissements publics d'enseignement supérieur qui ont commis une fraude ou une tentative de fraude sont l'avertissement, le blâme, l'exclusion de l'établissement, de manière temporaire ou définitive, peines qui vont

jusqu'à l'exclusion définitive de tout établissement public d'enseignement supérieur.

De plus en plus d'établissements commencent à sanctionner les plagiats avérés, allant jusqu'à mettre en ligne sur leur site Web les décisions adoptées par les commissions de discipline. Il est facile de constater que les risques encourus sont très importants. Il est par conséquent **beaucoup plus intelligent de citer ses sources et de tenter de produire une pensée originale...**



- Université d'Ottawa. *Attention au plagiat !* [en ligne].

<http://www.uottawa.ca/plagiat.pdf>

- Université de Lyon. *De la documentation... au plagiat, septembre 2007* [en ligne].

http://www.compilatio.net/files/sixdegres-univ-lyon_enquete-plagiat_sept07.pdf

- Université de Montréal. *Intégrité fraude et plagiat* [en ligne].

<http://www.integrite.umontreal.ca/>

- Wiki de l'université Descartes sur le plagiat [en ligne].

<http://wiki.univ-paris5.fr/wiki/Plagiat>

Les logiciels de gestion de références bibliographiques

Auteur :

Marie-Laure MALINGRE

Conservateur des bibliothèques – URFIST de Rennes

- 74** Qu'est-ce qu'un logiciel de gestion de références bibliographiques ?
- 74** Comment travailler avec un logiciel de gestion de références ?
- 76** Comment choisir un logiciel de bibliographie personnelle ?



Tout travail universitaire et scientifique s'appuie sur un ensemble de documents (articles de revues, ouvrages, aujourd'hui ressources Web...) permettant de le situer et de le légitimer par rapport à la littérature déjà produite sur le sujet ; il mentionne les références de ces documents selon une présentation normalisée.

Les logiciels de gestion de références bibliographiques ont été développés pour en faciliter la recherche, l'organisation et l'exploitation.

Qu'est-ce qu'un logiciel de gestion de références bibliographiques ?

Les logiciels de gestion de références bibliographiques ou logiciels de bibliographie personnelle sont des outils spécialisés permettant principalement :

- de créer une ou plusieurs bases de données personnelles ;
- d'y stocker, par saisie manuelle ou import automatisé, les références bibliographiques issues d'une recherche documentaire et/ou les métadonnées décrivant une page Web ;
- d'organiser et de gérer ces références au sein de la base constituée ;
- de formater les références selon un

style donné (conforme aux normes de présentation en vigueur) ;

- de générer automatiquement dans une publication la bibliographie ainsi formattée et les citations dans le texte correspondantes ;
- de réutiliser aisément le corpus de références dans différents contextes de publication ;
- de l'exporter sous divers formats et de le publier sur le Web ;
- de partager les références, pour certains d'entre eux.

Comment travailler avec un logiciel de gestion de références ?

Sauvegarder et archiver des références dans une base personnelle

La saisie de références se fait par des champs de saisie permettant d'indiquer,

pour chaque type de document, les informations qui vont constituer la référence, en respectant certaines règles minimales. L'insertion d'un lien (adresse Web ou chemin vers un fichier local) vers le **texte intégral des documents** permet l'accès au document lui-même. La présentation

finale dans une publication dépendra ultérieurement du choix d'un style d'édition. La **récupération automatique d'un lot de références** depuis une base de données distante ou un catalogue de bibliothèque constitue un des principaux atouts d'un outil bibliographique. Trois possibilités existent :

- la page de résultats affiche une **fonction d'export direct** et autorise le transfert automatique des références sélectionnées, au format compatible avec le logiciel de bibliographie ;
- l'import des références passe par des **filtres d'import** configurés dans l'outil bibliographique (si la base de données ne propose pas de récupération directe) : les résultats sont enregistrés sous forme de fichier texte, il faut ensuite choisir le filtre correspondant à la base de données et à son fournisseur pour transférer correctement le fichier ;
- le logiciel propose la **connexion directe à des bases de données** distantes : la recherche et le rapatriement de références se font sans quitter l'application, avec une grande facilité de procédure, mais moins

de souplesse et de complexité dans le mode d'interrogation.

Certains logiciels permettent aussi de récupérer les données descriptives de la page Web affichée dans le navigateur, de capturer la page Web ou d'enregistrer le lien vers la page.

Organiser et gérer ses références

Chaque logiciel comporte un module de recherche plus ou moins développé dans la base de références. Les références peuvent être triées et organisées en dossiers ou collections spécifiques.

La base permet de modifier les références une à une ou par lots (modification du contenu d'un champ sur plusieurs notices), de détecter et éliminer les doublons, de gérer les index ou les noms d'auteurs retenus. Enfin, la fonction d'export propose la conversion de la base de références sous différents formats pour l'édition (texte, Web) ou pour le traitement par d'autres outils bibliographiques.



- **Bibliographic/Software and Standards Information [en ligne].**

http://wiki.services.openoffice.org/wiki/Bibliographic_Software_and_Standards_Information

- **DELL'ORSO Francesco. *Bibliography Management Software: with a Detailed Analysis of Some Packages* [en ligne].**

<http://www.burioni.it/forum/dellorso/bms-dasp/>

- **MALINGRE Marie-Laure. *Introduction aux outils de gestion bibliographique* [en ligne].**

http://www.sites.univ-rennes2.fr/urfist/gestion_bibliographique

- **URFIST de Lyon. *Gestion des références bibliographiques* [en ligne].**

http://urfist.univ-lyon1.fr/38131655/0/fiche___pagelibre/&RH=1214984515166

- **URFIST de Paris. *FOURMI : Bibliographies (rédaction, gestion, manuels d'utilisation, citation des documents)* [en ligne].**

<http://urfist.enc.sorbonne.fr/anciensite/fourmi/fourmibiblio.htm>

Éditer ses références dans le cadre d'une publication

Les logiciels bibliographiques apportent une aide essentielle lors de l'écriture de travaux scientifiques, en proposant des modèles d'édition des références conformes aux **normes (instructions aux auteurs spécifiées par les revues scientifiques, normes d'écriture des thèses)**, ainsi qu'un interfaçage avec un ou plusieurs traitements de texte. Ce dispositif vise une utilisation conjointe du logiciel bibliographique et du traitement de texte, permettant la recherche de références en cours de rédaction et leur formatage selon un style fourni par le logiciel bibliographique (éventuellement modifié ou créé par l'utilisateur), pour l'insertion automatique des citations dans le texte, des

notes de bas de page et de la bibliographie ; l'avantage est une saisie unique des données de la référence, reformatées à volonté en cas de publication dans d'autres revues.

Partager ses références

La fonction de partage, particulièrement intéressante dans un environnement où la recherche s'élabore de plus en plus souvent de manière collective, permet de déposer et mettre en commun les références de documents ou de pages Web sur des plateformes dédiées : plateformes de partage de références bibliographiques et/ou de signets Web (**comme CiteULike, Connotea**), ou modules de partage implantés sur des logiciels bibliographiques en ligne (**par exemple sur RefWorks**).

Comment choisir un logiciel de bibliographie personnelle ?

Le paysage des logiciels bibliographiques s'est beaucoup diversifié, leurs fonctionnalités se sont développées ; un grand nombre d'outils gratuits (Zotero, CiteULike, Connotea, Bibsobomy...) ont fait leur apparition face aux grands produits commerciaux (EndNote,

Reference Manager ou RefWorks). Le choix prendra en compte les types de références proposés, les filtres d'import et les styles disponibles, l'interfaçage avec le traitement de texte, les capacités de gestion, l'accès (en ligne ou en local), les coûts.

• Biblioexpress

<http://www.bibsonomy.org/>

• CiteULike

<http://www.citeulike.org/>

• Connotea

<http://www.connotea.org/>

• EndNote

<http://www.endnote.com/>

• RefWorks

<http://www.refworks.com/>

• Zotero

<http://www.zotero.org/>



La veille informationnelle

Auteur :

Thomas CHAIMBAULT

Bibliothécaire - ensib

Vagabondages : <http://www.vagabondages.org/>

- 78** Qu'est-ce que la veille informationnelle ?
- 79** Des outils spécifiques



Pour disposer d'une information récente et efficace, ou pour continuer à être informé des nouvelles parutions pertinentes après une recherche documentaire, il peut être nécessaire de mettre en place un processus de veille informationnelle.

Qu'est-ce que la veille informationnelle ?

L'activité de veille est une activité continue et itérative, inscrite dans le temps, au contraire de la recherche d'information, ponctuelle et finie. Il s'agit plus précisément, par le biais de stratégies spécifiques, de surveiller de manière active et efficace le domaine étudié.

La mise en place d'une stratégie de veille s'effectue en plusieurs étapes. Le **ciblage** permet de définir précisément les thèmes sur lesquels va porter la veille et de choisir les sources à utiliser. C'est l'occasion de formaliser la veille par un plan de recherche, de viser à une plus grande efficacité afin d'éviter toute perte d'information. La **récolte des informations** à proprement parler s'effectue en utilisant les outils de recherche documentaire, afin de récupérer les informations. L'étape d'**analyse** des données, essentielle, permet d'écarter les données non fiables, non pertinentes, ou au contraire de collecter, classer les informations utiles pour, *in fine*, en préparer la diffusion. La **diffusion** des résultats de la veille, propre

aux institutions, peut se transposer pour l'étudiant au niveau de sa production dans le cadre de son travail universitaire.

La récolte d'information s'effectue selon deux méthodes. La première, le *pull* (signifiant tirer), consiste à aller chercher les informations sur les différents sites et bases de données intéressantes. Elle s'oppose au *push* (pousser), consistant à pousser, envoyer l'information vers l'utilisateur de manière automatique, en fonction de critères et de préférences préétablis par celui-ci.

Le *pull* est un processus lourd et fastidieux, qui exige de faire le tour des sites et des bases sélectionnés. Si les résultats sont précis (les résultats non pertinents peuvent être écartés rapidement), la veille apparaît irrégulière et exige qu'on y passe du temps.

Le *push* au contraire propose une économie de temps. Peu d'efforts sont à fournir une fois les alertes créées, qui avertissent automatiquement des nouveautés

- JAKOBIAK François. *Intelligence économique : la comprendre, l'implanter, l'utiliser*. Paris : Les Éditions d'Organisation, 2004.
- NOËL Élisabeth. *Veille et nouveaux outils d'information*. In *Usages, usagers et compétences informationnelles au xx^e siècle*. Paris : Hermès, 2008.

pertinentes, selon la périodicité choisie. Attention cependant au risque de surabondance des résultats, quand la stratégie de recherche n'est pas assez précise, à l'avalanche de mails venant d'abonnements à des listes de discussion, ou de spams.

Chaque méthode est valable et effective, en fonction de la stratégie de recherche adoptée. **La veille doit cependant répondre à un besoin identifié et être menée selon un objectif précis, afin d'être la plus efficace possible.**

Des outils spécifiques

La veille sur Internet peut se faire via plusieurs procédés. Le plus simple consiste en la répétition plus ou moins régulière de la même requête dans un outil de recherche, afin d'y découvrir des nouveautés. Des outils spécialisés, comme *Copernic Agent*, proposent d'automatiser la recherche selon une périodicité définie. D'autres outils plus complexes proposent d'aspirer à intervalle défini les sites Web sélectionnés et de signaler les changements intervenus. Ce type de veille est efficace mais nécessite une bonne maîtrise des outils.

La méthode *push* connaît un récent regain d'intérêt depuis l'utilisation de plus en plus importante de la technologie RSS (*Rich Site Summary* – sommaire de site

enrichi, ou *Really Simple Syndication* – syndication vraiment simple). Un flux, ou fils, RSS est un fichier au format XML, qui signale régulièrement le nouveau contenu textuel et multimédia de pages Web en les décrivant sommairement. Une fois abonné, le fil peut être lu à l'aide d'un logiciel approprié appelé « **agrégateur** », sans avoir à visiter le site Web d'où provient l'information. L'agrégateur signale automatiquement les nouveautés selon une périodicité paramétrable et permet de gérer la multitude de flux en les classant dans des catégories. Ainsi, l'abonnement à des flux RSS permet d'optimiser la veille, qui se concentre en un même endroit, accélérant l'accès au contenu, tout en évitant de laisser traîner son adresse de courrier électronique.



- JDEY Aref. *E-réputation. Cartographie des outils de veille*. 2009 [en ligne]. <http://www.demainlaveille.fr/2009/02/14/e-reputation-cartographie-des-outils-de-veille/>
- Educnet. *Utiliser les flux RSS et les podcasts* [en ligne]. <http://www.educnet.education.fr/dossier/rechercher/guides/flux>
- Manue. *Bibliothécaire et blogueuse...* Biblioacid, octobre 2005 [en ligne]. http://www.nicolasmorin.com/BiblioAcid_revue/BAv2n3.pdf
- ROLAND Michel. *RSS et syndication* [en ligne]. <http://www.slideshare.net/jdeyaref/rss-nouvelle-technologie-de-veille?src=embed>
- ROLAND Michel. *Automatiser la veille* [en ligne]. http://wiki-urfist.unice.fr/wiki_urfist/index.php/Automatiser_la_veille

Il existe des agrégateurs en ligne comme *Bloglines*, sous forme de page personnalisables comme Netvibes, d'autres intégrés aux navigateurs (les signets dynamiques), au logiciel de messagerie, ou qui s'installent sur l'ordinateur, en local.

Les outils de partage de données du Web 2.0 permettent aussi d'assurer une certaine veille. Ainsi, une application comme *Del.icio.us*, qui sert à enregistrer des signets en ligne, propose de voir qui a enregistré le même lien et de parcourir les favoris publics d'autres utilisateurs. Des fils RSS sont associés à *Del.icio.us*, et permettent de s'abonner aux signets d'un autre internaute ou spécifiques à un thème.

Enfin, des outils dits de *microblogging* comme Twitter, peuvent aussi aider à la veille. Cet outil permet de publier une information, en 140 caractères maximum, mais surtout en temps réel, depuis un téléphone portable par exemple. En choisissant avec soin les comptes twitter que l'on veut suivre ou en menant une veille sur certains mots (Search Twitter propose de créer un flux rss sur une requête), il est ainsi possible de se tenir au courant très rapidement d'évènements d'actualités, mais aussi de publications scientifiques, d'expériences, d'annonces lors de colloques... tout en restant critique face à la nécessité et à la qualité d'une information en temps réel.

- Copernic Agent Professional [en ligne].

<http://www.copernic.com/fr/products/agent/professional.html>

- Del.icio.us [en ligne].

<http://del.icio.us/>

- Google Reader [en ligne].

<http://www.google.fr/reader/>

- Netvibes [en ligne].

<http://www.netvibes.com/>

- Search twitter [en ligne].

<http://search.twitter.com/>

- Twitter [en ligne].

<http://twitter.com/>



Rédaction et citation des références bibliographiques

Auteurs :

Fatima NADJI

Doc'INSA Lyon

Dalila BOUDIA

Doc'INSA Lyon

- 82** Les références bibliographiques
- 83** L'appel de citation aux références bibliographiques
- 84** Abréviations des titres de périodiques



Le travail de recherche et l'écriture d'un texte scientifique (rapport, mémoire d'étude, thèse, etc.) suppose une recherche d'information approfondie. Cette recherche, inscrite dans la démarche scientifique, prend directement appui sur les travaux antérieurs. L'information choisie et exploitée permet à l'auteur de développer une réflexion personnelle. Ainsi, chacun des documents, retenus et analysés, contribue à la crédibilité scientifique du travail de recherche. Afin de faciliter la réflexion des lecteurs, qui à leur tour vont vouloir croiser leurs informations, il convient de référencer correctement les travaux cités en rédigeant une partie intitulée « Références bibliographiques ».

Les références bibliographiques doivent permettre, sans aucune ambiguïté, l'identification des documents, d'où découlera ensuite leur localisation. Les documents

listés dans la partie des « Références bibliographiques » du travail rédigé sont appelés dans le texte : ces citations établissent un lien avec la référence bibliographique.

Les références bibliographiques

82

La rédaction des références bibliographiques permet de référencer les documents retenus, selon des règles précises et une présentation cohérente et logique.

Les références bibliographiques sont rédigées conformément à la norme AFNOR NF Z-44-005 de décembre 1987, qui concerne les publications imprimées, livres et publications en série, leurs parties composantes (chapitres, articles) et les brevets ; la norme AFNOR NF ISO 690-2 (Z 44-005-2) de février 1998 correspond aux documents électroniques, documents complets ou parties de documents.

D'une manière générale, il est impératif d'identifier le **type** du document cité (ouvrage, article, congrès...), de préciser s'il est fait référence au document complet ou à une partie de ce document (chapitre d'un ouvrage, communication dans une conférence...), et de déterminer son **support** (papier, électronique...).

Une référence bibliographique :

- varie selon le type (ouvrage, articles, congrès...) et le support (papier, en ligne, cédérom...) du document auquel elle se rapporte ;

- est constituée d'éléments (auteur, titre, édition...) indispensables au bon signalement. Ces éléments doivent respecter un ordre précis. Pour retrouver ces éléments, les principales sources sont : les pages de titre, de couverture des documents... et parfois le dos de la page de titre. Les bases de données peuvent servir aussi à compléter les références bibliographiques ;

- doit être présentée de façon homogène, dans un style choisi pour chaque élément (taille des caractères, police et graisse).

Le *Guide de gestion des références bibliographiques* présente de manière détaillée la rédaction des références bibliographiques des documents sur support papier et électronique : *Guide de gestion des références bibliographiques* <http://docinsa.insa-lyon.fr/refbibli/index.php>.

L'appel de citation aux références bibliographiques

L'appel de citation dans le texte peut être identifié par le nom du premier auteur ou par un numéro. Les références bibliographiques seront classées dans le premier cas par ordre alphabétique d'auteurs, et dans le deuxième cas par ordre numérique.

Exemples d'appels de citation et de références bibliographiques

L'appel de citation par auteur :

- Dans le texte, les citations sont placées entre crochets (ou parenthèses) et sont formées du nom du premier auteur (ou de ses 3 premières lettres) suivi de l'année de publication. Pour différencier plusieurs documents d'un même auteur publiés la même année, les citations sont suivies d'une lettre distincte (a, b, c...).

Exemple : « ... le changement d'une propriété physique due à une réaction au niveau de sa surface avec le gaz à examiner [Esaki 80]. Une grande partie des dispositifs à état solide... »

- Dans la partie « Références bibliographiques », la référence est précédée de l'appel de citation telle qu'elle figure dans le texte. Les références sont classées par ordre alphabétique du premier auteur, puis chronologiquement.

Exemple : [Esaki 80] Esaki, Microwave infrared detector with semiconductor superlattice region, USA, Brevet n° 4348686, 28/07/1994...

L'appel de citation par numéro de citation :

- Dans le texte, les citations sont numérotées par ordre d'apparition. Les numéros sont placés entre crochets, entre parenthèses ou en exposants.

Exemple : « Pour une réalisation d'un certain nombre de composants du circuit primaire dont les traversées du couvercle de cuve [6]. ... »

- Dans la partie « Références bibliographiques », les références sont classées par numéro.

Exemple : [6] Pinard Legry G. Fiabilité des matériaux de structure dans l'industrie nucléaire. *Matériaux et techniques*, 1996, n° 7-8, pp. 11-18.

L'appel de citation par numéro de référence bibliographique :

- Dans le texte, les citations reprennent les numéros attribués dans la partie « Références bibliographiques ».

Exemple : « ... [27], à partir d'études sur des modèles expérimentaux établissent le schéma suivant : au cours de l'inspiration

... Là se situe la partie la plus rapide et la plus résistive de toute la voie aérienne : la valve nasale [2], [25]. La vitesse du flux... »

• Dans la partie « Références bibliographiques », les références sont classées par ordre alphabétique du nom du premier auteur et numérotées (dans cet ordre).

Exemple :

[25] O'NEILL G., TOLLEY N.S., The dynamics of nasal airflow, *Facial plastic surgery*, 1990, vol. 7, n°4, p. 215-220.

[26] PHILIP, COLE P., Review, Stability of nasal airflow resistance, *Clin. Otolaryngol*, 1989, vol. 14, p.177-182.

[27] PROCTOR D.F., ANDERSON I.B., The nose, Amsterdam : Elsevier Biomedical, 1982, 501 p.

Abréviations des titres de périodiques

Les titres de périodiques peuvent être abrégés selon la norme internationale ISO 4 de 1997, **règles pour l'abréviation des mots dans les titres et des titres des publications.**

Certains sites présentent des tables de correspondance entre les titres de périodiques complets et abrégés ; ils permettent de rédiger correctement les abréviations des titres de périodiques sans avoir à consulter la norme.

- BOUDIA Dalila, NADJI Fatima. *Guide de gestion des références bibliographiques : rédaction et citation* [en ligne].

<http://docinsa.insa-lyon.fr/refbibli/index.php>

- BU-Lettres / BU-Sciences (Université catholique de l'Ouest - Angers). *Les normes de rédaction bibliographique* [en ligne].

http://australie.uco.fr/info/biblio-info/menu3/menu3_1/bus_normes_bibliographie.php

- Filippi Dominique. *Comment présenter des références bibliographiques.* Bibliothèque Universitaire, Paris 8 [en ligne].

http://www.bu.univ-paris8.fr/web/bibliotheque/supports_formation/RefBiblio.html

- Infosphère - sciences humaines et sciences de la gestion. *Citer ses sources* [en ligne].

http://www.bibliotheques.uqam.ca/InfoSphere/sciences_humaines/module7/citer.html

- URFIST de Paris. *CERISE : citer un document* [en ligne].

<http://urfist.enc.sorbonne.fr/anciensite/cerise/p85.htm>



Gestion des signets, favoris ou bookmarks

Auteurs :

Gabriel GALLEZOT

Maître de conférence – URFIST de Nice

Michel ROLAND

Conservateur des bibliothèques – URFIST de Nice

- 86** Utilisation d'outil intégré au navigateur
- 86** Les gestionnaires de favoris en ligne
- 88** La gestion des références en ligne



Les gestionnaires de favoris permettent à l'utilisateur du Web de mémoriser ses "trouvailles", soit dans son navigateur, soit en ligne, pour mieux y retourner par la suite.

Utilisation d'outil intégré au navigateur

L'URL de la page Web peut être simplement mémorisée dans le navigateur (fonction : ajouter un signet, favori, *bookmark*, marque-page), éventuellement dans des dossiers personnalisables, et rendue accessible *via* le menu déroulant. Cependant, il devient vite difficile

de retrouver facilement des adresses, en raison de leur grand nombre et du manque de possibilité de description fine. Cette fonction reste utile pour garder « sous la main », dans la barre personnelle, les signets les plus fréquemment utilisés, les outils en particulier.

Les gestionnaires de favoris en ligne

Les outils sociaux du Web 2.0 permettent d'accéder de partout (nomadisme) aux liens enregistrés, d'étiqueter (*tagging*) les signets dans des dossiers et de les partager, en profitant du travail de sélection fait par d'autres utilisateurs de l'outil partageant des intérêts semblables.

Del.icio.us

Del.icio.us est le plus populaire des outils de gestion de signets en ligne, mais pas le seul. Il nécessite de se créer d'abord un compte, pour accéder par la suite au service après identification. *Del.icio.us* peut ensuite être intégré dans le navigateur, soit par extension ou module complémentaire, soit par l'ajout d'un book-

marklet, qui lance une mini-application pour effectuer certaines opérations. Deux boutons apparaissent alors dans la barre de navigation, l'un pour mémoriser l'URL de la page Web, l'autre pour ouvrir la page Web du compte *Del.icio.us* de l'utilisateur. Enfin, il reste toujours possible de mémoriser l'URL d'une page directement sur *Del.icio.us*.

L'enregistrement d'une URL permet d'ajouter des éléments d'information, zone de note pour décrire le contenu de la page, et **étiquette (ou tag)**, qui ressemble à un mot-clé, mais donne en priorité un sens parlant à l'utilisateur avant de décrire objectivement le contenu.



Save a Bookmark

delicious Save a Bookmark Signed in as **urfirst_pac**

URL <http://www.reversegraffitiproject.com/> do not share

TITLE Reverse Graffiti Project

NOTES 255 chars

TAGS ?

Separate Tags with spaces. E.g., hotels bargains newyork (not newyork)

Recommended Tags: click to add from your existing tags
art videos

Network Tags: click to share this link with your network
for:bibliothecaire for:cercamon for:chimist for:comedd for:gaga68 for:n_anna for:wiki_ao

Popular Tags: click to add from popular tags on
art graffiti culture design streetart video graphics green

Formulaire de saisie des informations sur Del.icio.us

La liste d'étiquettes déjà attribuées par d'autres à la page Web est proposée, ainsi que la possibilité de rendre le signet public ou privé (*do not share*). Cette possibilité doit cependant être activée. La page d'accueil d'un compte liste les signets mémorisés en ordre chronolo-

gique inverse (les derniers arrivent en haut de page), ainsi que les étiquettes associées (en liste ou « nuage de tags »), le nombre de personnes qui ont sauvegardé ce signet sur del.icio.us (indice de popularité) et la date de sauvegarde.

del.icio.us / bibliothecaire /

your bookmarks | your network | subscriptions | links for you | post

logged in as **bibliothecaire** | settings | logout | help

All your items (2396)

earlier | later » page 2 of 24

ATHINAIS HOTEL edit / share / delete
to hater Athens for cercamon ... saved by 1 other person ... not shared ... on april 06

title affordance.info: Faut-il stabiliser Wikipédia ?

url http://affordance.typepad.com/mon_weblog/2008/03/faut-il-stabili.html

notes

tags wikipedia do not share

affordance.info: Mon historique de navigation m'appartient (et le reste aussi) edit / delete
to illrestraining privacy ... saved by 1 other person ... on april 06

affordance.info: Le plug-in de la redocumentation. edit / delete
to wordpress fully featured installation plugin ... saved by 1 other person ... on april 06

Word macros - Meta edit / delete
to wordmacro word ... saved by 13 other people ... on april 05

Wiki Markup :: start edit / delete
to wiki markup ... on april 05

MDC:Wiki Markup Reference - MDC edit / delete
to wiki wiki markup ... saved by 13 other people ... on april 05

Column Two: Wiki markup has no future edit / delete
to wiki wiki markup wysiwyg ... saved by 7 other people ... on april 05

- collecting telegraphy bookmarking del.icio.us feeds reference_manager referencement referencings reference social source
- creating_content bibliotlog bibliowiki sau blog blogging blogs bloglist CMS content_management EP engines libraries microblogging platform todaywiki wiki wikis wikifon Wikimada Wikis_wiki_engine ecrispublisher
- open_access archive archiving CNRS communication e-open CPU creative_commons digital_libraries digital_repositories e-journal e-journals e-print Elsevier EPST EurOpenHolar evaluation institutional_repositories interoperability jolic: au CA-CHRE DA-PHE DAJ open_access open_archives open_science open_source production_scientifique publication_scientifique ras ranking repository research scholars scholeware-science_2.0 sites Shanghai
- people Bernard-Lee borges Christer Jacob daniel_bayd Diana_Lokenhoff Dunstai Erikakrast Harriet huguette_figot Huang lich jerry_walsh Jurek_Ragotz Levenez_Jeanne Martine Michael Olga_Lalina oliveras Olivier_LeDeuff Steve_Jobs Steven_Coburn
- reading armande_renaud books e-book e-books e-paper e-readers litch literacy reading
- signetsBPFirefox bkmk C/L Cerca cercachiere chercheRe DBP div communication dresardis dresblog dresbookmarking dresWiki

Page d'accueil d'un compte Del.icio.us

Le fait de voir qui a enregistré le même signet, puis les signets de cet internaute, de trouver de nouveaux signets associés à un tag utilisé correspond à du *social bookmarking*. D'autres fonctions, de recherche dans *Del.icio.us*, d'abonnements aux signets d'un autre internaute ou relatifs à un tag défini, de création de fils de syndication (RSS), de partage des signets recueillis enrichissent sans cesse l'utilisation de *Del.icio.us*.

Diigo

Diigo est un autre outil de gestion de signets en ligne, plus complexe et plus puissant, qui offre des fonctions

avancées permettant d'interagir avec d'autres applications Web. Il permet, outre la mémorisation de base de type *Del.icio.us*, de garder des extraits et de rajouter des annotations, comme lors d'une lecture « studieuse », crayon à la main.

Ces annotations peuvent aussi être vues par d'autres, si elles sont rendues publiques.

Les outils de gestion de signets, en ligne ou non, possèdent généralement des fonctions d'import et d'export permettant des échanges ou des transferts faciles.

The screenshot shows the Diigo web interface. At the top, there is a navigation bar with links for Home, Message (5), Friends, Meet People, Communities, Discover, and a search bar. Below this is a menu for 'My Dashboard', 'My Bookmarks', 'My Lists', 'My Tags', 'My Sites', and 'My Groups'. The main content area is titled 'My Bookmarks' and includes a filter box 'filter by tags, space-separated'. There are several bookmark entries, each with a title, tags, and a date. For example, one entry is 'Aide de Windows Vista : Que sont les favoris Internet Explorer ?' with tags 'IE, bookmarking' and a date of 'about 22 hours ago'. Another entry is 'OPACs à la loupe : Marlene's corner' with tags 'opac, library_2.0, web' and a date of 'on 2008-06-11'. A third entry is 'The Library in the New Age - The New York Review of Books' with tags 'book, readings, library, damian' and a date of 'on 2008-06-10'. On the right side, there is a sidebar titled 'My Recent Tags (50)' with a list of tags including 'Steve_Jobs, USA, api, architecture, attention, bibliography, blogging, book, bookmarking, censorship, copyright, decline, desktop_tools, digital_natives, droit_d_auteur, droits, e_readers, epst, facebook, flickr, flickr_export, folksonomy, fonctionnariat, france, identity, intellectual_property, libraries, livestreaming, literacy, metadata, networks, non-modern, opensource, reading, reading_2.0, reference'.

Signets sur *Diigo*

La gestion des références en ligne

La popularité et l'efficacité d'un outil comme *Del.icio.us* a inspiré le développement d'outils de même type dans les outils de gestion des références bibliographiques. En effet, l'importance acquise

par Internet pour le monde scientifique a fait de l'accès en ligne une qualité déterminante pour l'utilisateur. Cependant la gestion des références scientifiques suppose l'importation dans l'outil de méta-

données plus riches ; la capacité de l'outil à les récolter automatiquement est un critère de choix essentiel. *CiteULike* est un bon exemple

Ces deux types d'outils croisent certains critères : la nature de l'objet mémorisé

(page/signets, ou référence bibliographique), la localisation de l'outil (local en mode client, hébergé en ligne par un fournisseur ou encore installé sur un serveur local), enfin le coût, les logiciels bibliographiques pouvant être payants.

objets mémorisés ⇨ accès à l'outil ☞		SIGNETS (Pages web)	REFERENCES (Documents)	
		Gratuit		Onéreux
LOCAL		Favoris IE Marques-pages Firefox	Bibus Zotero BibDesk	EndNote
EN LIGNE	hébergé	Del.icio.us ... Google bookmarks Diigo	CiteULike Zotero 2.0	Refworks EndNote Web
		Bibsonomy Connotea 2collab		
	serveur local	Scuttle	Wikindx (Bibus)	
		Connotea		

Type d'outils de gestion de signets/références

Ces outils ne sont pas équivalents mais procèdent de la même logique : mémoriser en ligne des références (bibliographiques ou pages Web). Certains sont clairement orientés vers les références bibliographiques (*CiteULike*), d'autres « signets » mais permettent, au détour d'un menu dédié, d'intégrer des références bibliographiques

(*Connotea*), d'autres encore affichent simplement leur dualité (Bibsonomy).

CiteULike

Le fonctionnement de *CiteULike* propose trois grands principes d'intégration des références bibliographiques :

depuis un *Bookmarklet* ajouté au navigateur, depuis une liste de sommaires présentée au sein de *CiteULike*, ou encore une intégration manuelle. Des tags peuvent être associés à la des-

cription du document, ainsi que le document source en PDF et la possibilité de partager la référence avec d'autres utilisateurs.



Exemple d'interface de partage de références d'articles (CiteULike)



- **URFIST de Nice. *Mémoriser : signets, favoris et références* [en ligne].**
http://wiki-urfist.unice.fr/wiki_urfist/index.php/Mémoriser:_signets,_favoris_et_références
- **Bibsonomy [en ligne].**
<http://www.bibsonomy.org/>
- **CiteULike [en ligne].**
<http://www.citeulike.org/>
- **Connotea [en ligne].**
<http://www.connotea.org/>
- **Del.icio.us [en ligne].**
<http://del.icio.us/>
- **Diigo [en ligne].**
<http://www.diigo.com/>
- **Google Bookmarks [en ligne].**
<http://www.google.com/bookmarks/>
- **Iterasi [en ligne].**
<http://iterasi.com/>
- **Zotero [en ligne].**
<http://www.zotero.org>

Publier sur le Web

Auteur :

Élisabeth NOËL
Conservateur des bibliothèques – enssib

- 92** Les blogs
- 92** Les wikis
- 93** Les CMS (*Content management systems*)



Il devient nécessaire aux étudiants de savoir publier sur le Web, que ce soit dans un cadre pédagogique, pour la recherche d'emploi, ou pour des questions personnelles. De nombreux outils de publications sont maintenant disponibles, qui répondent à des besoins différents.

Les blogs

Les blogs se sont répandus récemment. Ce sont des sites Web qui proposent des billets (*posts*, notes ou encore articles) assez courts, parfois avec du multimédia, classés par ordre anté-chronologique (les plus récents en premier). En général, les internautes peuvent laisser des commentaires sur les billets ; le blogueur propose aussi une liste de sites associés, le *blogroll*.

Ces sites sont très faciles à créer, avec des plates-formes de blogs en ligne comme *Wordpress*, un des outils les plus intéressants actuellement, ou en installant un logiciel de blog sur un serveur. Les outils d'administration proposés maintenant permettent d'en faire de véritables sites Web, riches en fonctionnalités et personnalisés.

Les blogs servent souvent à donner au jour le jour des nouvelles d'une personne (journaux intimes), mais peuvent aussi refléter des découvertes, des résultats de recherche (sur Internet). Ils sont de plus en plus utilisés par les scientifiques (comme *Affordance*, tenu par Olivier Ertzscheid), renouvelant ainsi leur « collègue invisible », ou par des associations professionnelles, car ils sont très utiles pour diffuser les résultats de la veille. Certaines universités proposent des solutions institutionnelles à leurs membres et à leurs étudiants, comme l'Université de Paris 5 avec son réseau social.

Les wikis

Les wikis sont des sites dynamiques sur lesquels l'écriture est possible aux visiteurs autorisés, à travers un applicatif localisé directement sur le serveur. L'archivage des différentes versions successives permet de revenir à une version précédente, si nécessaire.

Un wiki peut être créé facilement dans une « ferme à wikis », en ligne, ou avec le logiciel adéquat sur un serveur. L'écriture se fait à travers un éditeur de texte ou *via* une syntaxe particulière, plus complexe, la syntaxe wiki. Il est toujours possible de s'entraîner dans les « bacs à sable ».

- DELACROIX Jérôme. *Les wikis*. Paris : M21 Éditions, 2005.
- LAPOINTE Pascal, DROUIN Josée Nadia. *Science, on blogue ! Le nouveau monde d'Internet*. Paris : Multimondes, 2007.

Surtout connu à travers l'exemple de Wikipedia, le wiki est idéal pour l'écriture collaborative de documents et pour la gestion de projet. Néanmoins, son

utilisation reste relativement difficile, ce qui fait qu'il n'est pas systématiquement exploité par les utilisateurs potentiels.

Les CMS (*Content management systems*)

Les CMS (*Content management systems*, Systèmes de gestion de contenus) permettent de concevoir et utiliser facilement des sites Web. Assez élaborés, ils simplifient la chaîne de publication, en séparant le contenu et sa présentation, et proposent une écriture simplifiée en ligne.

Les CMS sont adaptés à la création de sites de petites associations, d'institutions. Les étudiants peuvent être amenés à les utiliser dans le cadre de leurs études. SPIP est très répandu en France, et utilisé par de nombreuses universités.



- **Café des sciences [en ligne].**

<http://www.cafe-sciences.org/>

- **CHAIMBAULT Thomas. *Ouvrir son blog*. Fiche pratique ensib [en ligne].**

<http://www.ensib.fr/bibliotheque-numerique/document-1691>

- **ERTZSCHEID Olivier. *Blogs et RSS dans les pratiques scientifiques en IST*. Affordance [en ligne].**

http://affordance.typepad.com/mon_weblog/files/iexposciencesblog.pdf

- **ERTZSCHEID Olivier. *Science on blogue ! Affordance* [en ligne].**

http://affordance.typepad.com/mon_weblog/2008/04/science-on-blog.html

- **Université Paris Descartes. *Carnets 2* [en ligne].**

<http://carnets.parisdescartes.fr/>

- **URFIST de Paris. *Les wikis* [en ligne].**

<http://urfist.enc.sorbonne.fr/anciensite/rss/wiki.htm>

- **SPIP [en ligne].**

<http://www.spip.net/>

- **Wordpress [en ligne].**

<http://www.wordpress-fr.net/>

- **ZohoWiki [en ligne].**

<http://wiki.zoho.com/>

Les outils de recherche d'information sur le Web

Auteurs :

Marie-Françoise DEFOSSE
Professeur associée – enssib

Élisabeth NOËL
Conservateur des bibliothèques – enssib

- 96** Outils de recherche
- 97** Quel outil choisir ?
- 97** Méthode de tri des résultats
- 98** Un peu de syntaxe



Une multitude d'outils différents composent le paysage de la recherche d'informations sur le Web, paysage mouvant (rachats, disparitions temporaires ou définitives, créations incessantes de nouveaux outils) dominé par Google. Seule une bonne connaissance de la nature et du fonctionnement de ces outils, de leur particularité et de la syntaxe des équations de recherche permet d'arriver à des résultats satisfaisants sans y consacrer trop de temps.

Outils de recherche

Les annuaires de recherche

Il s'agit de répertoires (catalogues, guides ou *directories*) de sites classés manuellement en grandes catégories organisées hiérarchiquement. Construits « à la main », après sélection plus ou moins sommaire, ils sont constitués de rubriques (ou catégories), elles-mêmes subdivisées en sous-rubriques. Un annuaire s'utilise par navigation dans les catégories ou par une recherche mots-clés dans les descriptifs des sites. Actuellement, ce modèle à tendance à disparaître, sauf pour les annuaires spécialisés.

Annuaire spécialisés

- CISMef [en ligne]
<http://www.chu-rouen.fr/cismef/>
- Intute [en ligne]
<http://www.intute.ac.uk>
- Signets de la Bibliothèque nationale de France [en ligne]
<http://signets.bnf.fr>
- The Virtual Library [en ligne]
<http://vlib.org>

Les moteurs de recherche

Ce sont les outils d'accès au Web les plus utilisés. Des robots (appelés crawlers ou spiders) parcourent le Web de lien en lien et indexent automatiquement les documents trouvés. Les moteurs proposent des formulaires de recherche simple et avancée, par mots-clés.

Moteurs de recherche généralistes

- Bing [en ligne]
<http://www.bing.fr>
- Google [en ligne]
<http://www.google.fr>
- Yahoo ! [en ligne]
<http://fr.yahoo.com>

Les métamoteurs

Les métamoteurs interrogent plusieurs outils de recherche simultanément, en dédoublonnant les réponses. Sous forme logicielle, installés sur le poste, ils facilitent la veille. Plus exhaustive que sur un moteur, la recherche sera néanmoins moins pointue. Ils propo-

- GALLEZOT Gabriel, SIMONNOT Brigitte. *L'entonnoir*. Caen : C&F éditions, 2009.
- MANIEZ Dominique. *Les 10 plaies d'Internet*. Paris : Dunod, 2008.



sent des spécificités comme la catégorisation automatique des résultats, c'est-à-dire la création à la volée de dossiers thématiques contenant les résultats trouvés (avec Clusty, <http://clusty.com>) ou la recherche associée dans le web et les réseaux sociaux (avec PolyMeta, <http://www.polymeta.com>).

Métamoteurs

- Copernic (logiciel)
<http://www.copernic.com/fr/>

- Ixquick [en ligne]
<http://eu.ixquick.com/fra/>
- Metacrawler [en ligne]
<http://www.metacrawler.com>

À noter : les méta-moteurs, qui sont souvent des produits d'appels proposés gratuitement par des éditeurs de logiciels de veille, permettent ainsi de mener sa veille.

Quel outil choisir ?

Le choix de l'outil dépend étroitement du type de recherche à effectuer.

Un annuaire permet de retrouver un ensemble de sites (et non de pages web) centrés sur une thématique spécifique. Il permet de débiter une recherche, d'identifier rapidement et précisément des sites de référence et de repérer des ressources du Web invisible.

Un moteur de recherche, fonctionnant automatiquement, proposera des résultats plus nombreux mais avec un bruit plus prononcé. La recherche porte sur les pages elles-mêmes : le moteur est à uti-

liser pour la taille de son index, la rapidité de mise à jour de l'index. Enfin, la recherche avancée propose des paramètres très intéressants, permettant de mener des recherches pointues sur un sujet.

Le métamoteur permet d'obtenir un premier aperçu des informations dites « de référence ». Ses résultats dépendent des outils de recherche qu'il utilise.

Les outils de recherche spécialisés s'avèrent souvent être plus efficaces que les outils généralistes.

Méthode de tri des résultats

Les moteurs de recherche complètent la recherche par des techniques de tri des résultats. Plusieurs solutions ont été développées, malheureusement souvent détournées.

Tri par pertinence

Les résultats d'une requête sont affichés selon un ordre déterminé par le calcul d'un score pour chaque réponse. La pertinence est basée sur les 5 facteurs suivants, appliqués aux termes de la question :

- Le poids d'un mot dans un document est déterminé par sa place dans le document :

il est maximum pour le titre et le début du texte ; à l'intérieur du texte, il est plus important si le mot est en majuscule.

- La densité est basée sur la fréquence d'occurrence du terme dans un document, par rapport à la taille du document. Si deux documents contiennent le même nombre d'occurrences, le document le plus petit sera donc favorisé.

- Le poids d'un mot dans la base est basé sur la fréquence d'occurrence du terme dans toute la base de données. Les mots peu fréquents dans le corpus sont favorisés. Les mots vides sont soit éliminés, soit sous-évalués.

- La correspondance d'expression est basée sur la similarité entre l'expression de la question et l'expression correspondante dans un document. Un document contenant une expression identique à celle de la question reçoit le poids le plus élevé.
- La relation de proximité est basée sur la proximité des termes de la question entre eux dans le document. Les termes proches sont favorisés.

Tri par popularité

Google a lancé cette méthode, basée sur le nombre de liens pointant sur les pages (PageRank). Si ce critère, remis en question

par les professionnels du référencement, prend maintenant moins d'importance, la popularité d'une page demeure fondamentale... Plus elle est demandée, plus elle est connue, plus elle est consultée, plus elle est « populaire » et mieux elle est placée dans la liste de résultats !

Tri par calcul dynamique de catégories

Cette méthode effectue un classement des documents trouvés dans des dossiers (*clustering*) constitués automatiquement en fonction des réponses. Un dossier peut lui-même être constitué de sous-dossiers.

Un peu de syntaxe

La plupart des outils de recherche sur Internet utilisent la syntaxe suivante, qui pourra être employée sans risque grave. D'autres possibilités de recherche existent, mais varient d'un outil à l'autre.

La recherche avancée permet souvent de poser la question de manière guidée. Le site Abondance propose des fiches pour chaque outil de recherche, présentant les spécificités de syntaxe de chacun.

Opération	Opérateurs	Exemple
Rechercher une expression	" "	"vache folle"
Imposer un terme	+	+vache +folle
Exclure un terme	-	+bretagne -grande

Attention, les signes + et - doivent être collés à gauche du terme concerné. Cette syntaxe correspond en général aux formulaires de recherche simple.

- **ADBS. Outils de recherche web [en ligne].**

<http://www.adbs.fr/outils-de-recherche-web-828.htm?RH=1214314880901>

- **ANDRIEU Olivier. Abondance – outils de recherche [en ligne].**

<http://outils.abondance.com/>

- **URFIST de Lyon. DADI : répertoire des bases de Données gratuites Disponibles sur Internet [en ligne].**

<http://dadi.univ-lyon1.fr/>

- **URFIST de Lyon. Vite tous les outils [en ligne].**

http://urfist.univ-lyon1.fr/28536467/0/fiche___pagelibre/&RH=1215024972516



Les outils de recherche sur Internet spécifiques au domaine scientifique

Auteur :

Christophe BOUDRY

Maître de conférences – URFIST de Paris

- 100** Principaux outils de recherche spécifiques aux domaines scientifiques
- 102** De nouveaux outils, basés sur l'indexation sociale...



Dans un domaine donné, les outils de recherche généralistes du Web sont parfois à l'origine de résultats bruités (non pertinents car hors sujet), ce qui représente une gêne considérable lors d'une recherche d'informations spécialisées. Pour limiter ces difficultés, des outils ont été développés pour répondre aux attentes des scientifiques en matière de recherche d'informations, en permettant de cibler spécifiquement le domaine scientifique.

Pour gagner en efficacité, ces outils sélectionnent des sites Internet appartenant uniquement au domaine scientifique, mais permettent également de localiser des

ressources documentaires variées utiles aux scientifiques, comme des articles scientifiques, des livres, des comptes rendus de congrès ou bien encore des brevets.

Principaux outils de recherche spécifiques aux domaines scientifiques

Il existe un grand nombre d'outils de recherche dédiés aux différents domaines scientifiques. Si il y a encore quelques années, la plupart se contentaient de référencer uniquement des pages Web, ils permettent aujourd'hui, dans leur quasi-totalité, de repérer non seulement des sites Web du domaine scientifique, mais également des articles scientifiques de revues (grâce à des accords avec certaines bases de données), des brevets, des comptes rendus de congrès, des livres

scientifiques. Ces sites tendent donc à devenir de véritables portes d'entrées vers les ressources documentaires scientifiques.

Ces outils peuvent couvrir tous les domaines scientifiques (*Scirus - for scientific information ; Google Scholar*) ou être spécialisés dans des domaines disciplinaires plus restreints : par exemple *Chemdex* dans le domaine de la chimie, ou bien encore *Techxtra* dans le domaine de l'informatique et des mathématiques.

- FOENIX-RIOU Béatrice. *Recherche et veille sur le Web visible et invisible : agents intelligents, annuaires sélectifs, interfaces des grands serveurs, portails thématiques*. Paris : Tec & Doc, 2001.
- LENORMAND Patrick. *Internet : techniques de recherche pour les professionnels*. Paris : ENI, 2007.

Advanced search

All of the words in The complete document

AND

All of the words in The complete document

Search tips

author:smith find results that have "smith" in the author field
 DNA -sequencing find results that have "DNA" but not "sequencing" in the text
 car* finds "car" as well as "carbon", etc. [View all search tips](#)

Dates	Only show results published between <input type="text" value="before 1900"/> and <input type="text" value="2009"/>
Information types	Only show results that are <input checked="" type="checkbox"/> Any information type <input type="checkbox"/> Abstracts <input type="checkbox"/> Articles <input type="checkbox"/> Books <input type="checkbox"/> Company homepages <input type="checkbox"/> Conferences <input type="checkbox"/> Patents <input type="checkbox"/> Preprints <input type="checkbox"/> Scientist homepages <input type="checkbox"/> Theses and Dissertations
File formats	Only show results that are

Exemple d'interface de moteur de recherche spécifique du domaine scientifique (Scirus)

Le grand nombre de disciplines scientifiques et d'outils de recherche disponibles dans chacune de ces disciplines rend impossible une présentation

exhaustive de ces outils. Aussi, le tableau suivant propose une sélection d'outils dans quelques disciplines scientifiques :

Discipline	Adresse
Toutes disciplines	Scirus - for scientific information (moteur) http://www.scirus.com/ Google Scholar (moteur) http://scholar.google.com IPL2 (annuaire) http://www.ipl.org/ SciNet Science Search http://www.scinet.cc/
Biologie/Médecine	Intute http://www.intute.ac.uk/biologicalsciences/
Mathématiques, informatique	Techxtra http://www.techxtra.ac.uk/ Intute http://www.intute.ac.uk/mathcompsci/
Chimie	Chemdex http://www.chemdex.org/
Sciences humaines et sociales	http://www.intute.ac.uk/socialsciences/ http://www.intute.ac.uk/humanities/
Histoire	History on-line http://www.history.ac.uk/ihr/Resources/index.html

Pour découvrir d'autres outils de recherche du domaine scientifique, un certain nombre de sites proposent des listes par disciplines scientifiques. En voici quelques-unes :

- Open Directory - Science: Directories [en ligne]
<http://dmoz.org/Science/Directories/>

- Science Search Engines [en ligne]
<http://www.searchengineguide.com/pages/Science/>
- Yahoo! Science - Web Directories [en ligne]
<http://dir.yahoo.com/>

De nouveaux outils, basés sur l'indexation sociale...

Dans la mouvance du Web 2.0, sont apparus un certain nombre de sites Web « de partage de ressources » basés sur l'indexation sociale (*folksonomie*) et destinés aux scientifiques. Le principe de ces outils est de permettre aux usagers, après indexation par des mots-clés personnels (*tags*), de partager leurs ressources documentaires (signets, références d'articles scientifiques, références de livres, diaporamas...) avec d'autres usagers sur des plateformes spécifiques. Le terme d'indexation sociale indique ici que, contrairement aux outils de recherche traditionnels, ce sont les usagers – et non des professionnels de la documentation – qui

pratiquent l'indexation des ressources qu'ils souhaitent partager avec les autres internautes.

Le partage de ces ressources a pour but de mutualiser des connaissances individuelles, afin de permettre aux usagers d'accéder à de véritables bases de connaissance dans lesquelles ils peuvent, bien entendu, pratiquer des recherches dans la totalité des ressources partagées ; ces recherches s'avèrent être le plus souvent redoutablement efficaces et complémentaires de celles pratiquées avec les outils de recherche disciplinaires plus traditionnels.

Type de ressources partagés	Nom et adresse du service
Signets généralistes et spécialisés	Del.icio.us http://delicious.com
	Diigo http://www.diigo.com
Signets/références d'articles scientifiques	Citeulike http://www.citeulike.org
	Connotea http://www.connotea.org
Livres	Librarything http://www.librarything.com
Diaporama	Slideshare http://www.slideshare.net



Page d'un diaporama de conférence

Les ressources partagées vont des diaporamas (Slideshare) aux bibliothèques personnelles (Librarything). Slideshare peut ainsi servir d'espace de stockage et de

partage pour les diaporamas d'un enseignant, mais aussi d'un colloque, avec l'utilisation judicieuse et partagée d'un tag commun.



- GINOUVES Véronique, PEYSSARD Jean-Christophe. *Bibline* : annuaire de liens spécialisé en sciences humaines et sociales dans le domaine des études méditerranéennes. Bulletin des bibliothèques de France, 2006, n° 3, p. 49-52 [en ligne].

<http://bbf.enssib.fr/consulter/bbf-2006-03-0049-008/>

**La recherche des documents
multimedia** sur Internet
(images, vidéo, sons)

Auteur :

Christophe BOUDRY

Maître de conférences – URFIST de Paris



Internet offre principalement des ressources textuelles, mais aussi des ressources de type multimédia en grand nombre. Cependant, la méthodologie de recherche de ces données est moins connue et présente certaines caractéristiques propres à ces types de documents.

Méthodes de recherche des documents multimédia

La recherche par mot-clé sur le texte associé aux documents multimédia

La méthode de recherche de documents multimédia s'effectue, en saisissant un ou plusieurs mots-clés dans le formulaire proposé, à l'intérieur des données textuelles associées au document multimédia, par exemple les données de la page Web où est présenté le document multimédia (notion décrite par le terme *concept-based image retrieval* dans la littérature anglo-saxonne, en ce qui concerne les images).

Cette méthode de recherche est la plus utilisée par les outils de recherche traditionnels sur Internet. Les méthodes d'indexation et de recherche mises en œuvre s'apparentent alors tout à fait à celles utilisées pour la recherche de pages Web, les données indexées sont alors des informations textuelles qui se rapportent aux

documents multimédias à l'intérieur de la page Web, et qui sont localisées à deux emplacements :

- dans le code HTML des balises et attributs spécifiques aux documents multimédias. Pour les images il s'agit de la balise ``, qui indique l'insertion d'une image dans une page HTML, de la balise ``, qui indique un lien hypertexte vers un document multimédia, ou encore de l'attribut `ALT="..."` qui représente le texte alternatif d'une image (l'inconvénient est qu'il n'est pas toujours renseigné par les créateurs de pages HTML) ;
- dans le texte de la page associée ou dans le titre de la page « support » de l'image.

Ce type de méthode de recherche fait abstraction des données intrinsèques aux documents multimédia recherchés, et dépend de la « pertinence » du texte de la page Web associé aux documents.

La recherche sur le contenu des documents multimédia

Développées relativement récemment et encore peu utilisées sur Internet, les méthodes de recherche sur le contenu des documents multimédia (notion de *content-based image retrieval* dans la littérature anglo-saxonne, en ce qui concerne les images) ont pour principe **d'identifier ces documents à partir de leur contenu**, c'est-à-dire directement à partir des données du fichier, et non plus à partir du texte associé.

Ce processus de recherche est divisé en deux phases :

- une phase d'indexation, qui extrait des informations à partir des fichiers (par exemple : informations sur la couleur pour un fichier image, ou informations sur la composition en terme de fréquence pour un fichier son), afin de représenter chaque document par un ensemble de paramètres permettant de décrire son contenu le plus précisément possible. Les valeurs de ces paramètres sont indexées, et représentent la « signature » de chaque fichier dans la base de données,
- une phase de recherche, où l'utilisateur va choisir un document multimédia de départ, afin que l'outil de recherche compare sa « signature » à toutes celles des

documents présents dans l'index. L'outil propose alors la liste des documents ayant une « signature paramétrique » similaire à celle du document de départ.

Les méthodes de traitement des données utilisées pour implémenter ces techniques de recherche peuvent varier, une des plus classiques étant l'analyse de type multi-paramétrique. Dans le cas des images, l'indexation des données extrait des images certains paramètres physiques qui rendent compte de la couleur, de la texture, de l'intensité ou bien encore des formes contenues dans chaque image. À chaque image sont associées, dans l'index, les valeurs des paramètres utilisés pour la décrire. Le point de départ de la recherche est alors une image, ce qui permet de localiser dans l'index des images similaires, dont les paramètres de description sont les plus proches de ceux de l'image initiale. Il est possible de sélectionner certaines parties pour la recherche, ou de privilégier des caractéristiques comme la couleur, la texture, l'intensité ou la forme.

Ce type de méthode est encore peu utilisé par les outils de recherche de documents multimédia, car l'extrême variabilité des documents de ce type rencontrés sur Internet représente un frein majeur à l'obtention de performances satisfaisantes.

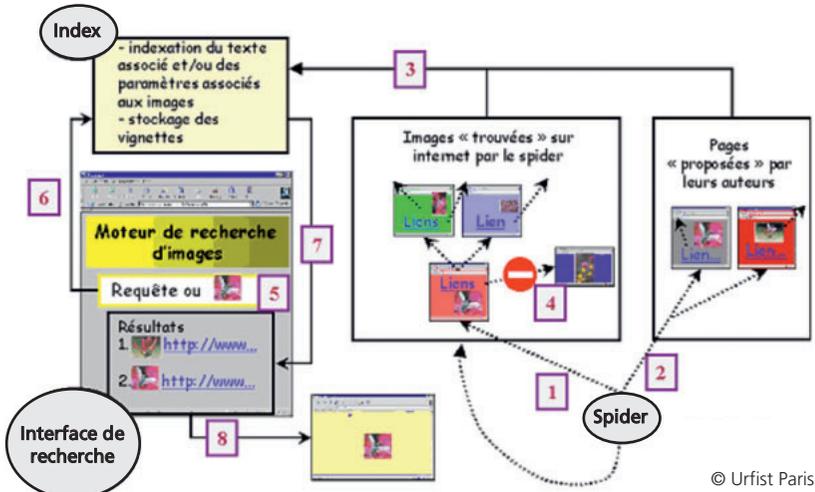
Outils de recherche multimédia sur Internet

Les moteurs de recherche d'images

Ces outils peuvent théoriquement utiliser indifféremment la recherche par le texte associé et/ou sur le contenu des images, mais rares sont ceux qui proposent réellement une recherche sur le contenu.

Le *spider* (ou robot) repère les images sur les pages Web (à l'aide des extensions de fichiers) qu'il parcourt par suivi récur-

sif des liens des pages Web visitées [1] ou après soumission par leur auteur [2], et qu'il intègre dans l'index [3]. Certaines images sont inaccessibles au *spider* [4] : images situées sur des pages web en accès réservé ou orphelines... L'index constitue le lieu de stockage des images (ou plus exactement des vignettes des images), du texte associé aux images et/ou des paramètres associés à chaque image.



© Urfirst Paris

Principes de fonctionnement des moteurs de recherche d'images

L'interface de recherche permet à l'utilisateur de saisir sa requête textuelle ou de sélectionner une image comme point de départ pour la recherche [5]. Les termes recherchés ou les paramètres décrivant l'image de départ sont alors « confrontés » à ceux présents dans l'index [6]. L'utilisateur peut consulter [7] en retour une liste d'images, proposées selon un ordre de pertinence donné. Il peut accéder aux images d'origine présentes sur le site Web via un lien hypertexte [8].

Certains de ces outils ont été spécifiquement développés pour la recherche d'images (exemple de *Picsearch*), tandis que d'autres sont des adaptations de moteurs de recherche de pages Web aux images (cas de *Google* et de *Alta Vista*).

Principaux moteurs de recherche d'images

Sur les textes associés

- Corbis [en ligne]
<http://www.corbisimages.com/>

- Google Image [en ligne]
<http://images.google.com>
- Lycos [en ligne]
<http://search.lycos.com/?tab=multi&cat=images>
- Picsearch - the search engine for pictures and images
<http://www.picsearch.com/>

Par le contenu

- Gazopa [en ligne]
<http://www.gazopa.com>
- Idée Labs [en ligne]
<http://labs.ideeinc.com>
- Retriever [en ligne]
<http://labs.systemone.at/retriever>
- Tiltomo [en ligne]
<http://www.tiltomo.com>

Les sites de partage d'images

Directement issus de la mouvance Web 2.0, les sites de partage d'images sont apparus récemment. Leur principe est le partage des images numériques des

usagers du Web, en utilisant des *tags* (ou mots-clés) librement choisis pour les décrire et les rechercher. Les plus connus sont *Flickr* et *Photobucket*, qui constituent ainsi une base énorme de photos, souvent de très bonne qualité, et qui sont réutilisées dans de nombreuses applications du Web 2.0, comme *Retriever*, pour des recherches d'images similaires.

Principales plates-formes de partage d'images

- Flickr [en ligne]
<http://flickr.com>
- Photobucket [en ligne]
<http://photobucket.com>

Outils de recherche d'images animées (vidéos)

Le domaine des outils de recherches de vidéo a subi un changement de paysage radical ces dernières années. Il y a encore peu de temps, des outils au fonctionnement similaire à celui des moteurs de recherche d'images présentés plus haut permettaient de repérer les vidéos hébergées sur des pages Web, *via* les onglets « Vidéo » des outils de recherche. Si ces outils existent encore aujourd'hui, ils fonctionnent quasiment tous avec des vidéos provenant de sites de partage de vidéos, non plus provenant de pages Web « quelconques ».

Les sites de partage de vidéos sont très nombreux. *Youtube*, fondé en février 2005 et racheté en octobre 2007 par Google, est le leader incontesté en termes de trafic. La plupart de ces sites proposent les vidéos au format « Flash Vidéo ».

Principales plates-formes de partage de vidéos

- Dailymotion [en ligne]
<http://www.dailymotion.com>
- Metacafe [en ligne]
<http://www.metacafe.com>

- Vimeo [en ligne]
<http://www.vimeo.com>
- Youtube [en ligne]
<http://www.youtube.com>

Outils de recherche de sons et musique

Sur Internet, les outils disponibles pour rechercher des sons – fichiers dont le format le plus connu est le MP3 – sont principalement des moteurs de recherche. Certains proposent un choix au niveau du type de format de fichier recherché (comme *Altavista*, *Findsound* ou encore *Alltheweb*).

Il faut également signaler les systèmes de recherche basés sur le partage de fichiers (*peer to peer*), dont le représentant emblématique a été *eMule*. Leur fonctionnement est basé sur le principe que tout utilisateur connecté à ces outils doit donner accès aux autres utilisateurs à certains fichiers (les fichiers MP3 notamment) présents sur son ordinateur. En échange, chaque utilisateur peut accéder à tous les fichiers mis à disposition par les autres utilisateurs du réseau. Ces outils, accusés par les producteurs et distributeurs officiels de musique de faciliter le piratage, sont en constante évolution.

Les sites d'écoute en ligne de musique se développent : ils permettent de chercher des titres, de les écouter, de créer sa propre sélection et d'obtenir les fichiers, tout cela légalement.

Principaux outils de recherche de sons et musique

- Deezer [en ligne]
<http://www.deezer.com>
- Findsound FindSounds - Search the Web for Sounds [en ligne]
<http://www.findsounds.com>

Centre de ressources de formations pour la recherche de l'information spécialisée

Les **URFIST** (Unités Régionales de Formation à l'Information Scientifique et Technique) ont une double mission : former le personnel des bibliothèques ainsi que le public universitaire (étudiants de 3^e cycle et enseignants) à l'information scientifique et technique, et plus généralement les sensibiliser aux nouvelles technologies de l'information.

URFIST de Bordeaux

Département Documentation URFIST
4 avenue Denis-Diderot
33607 Pessac
Tél. 05 56 84 29 19
Fax 05 56 84 86 96

Responsable : Marie-France ANDRAL

Site Web : <http://www.univ-bordeaux.fr/ddoc/urfist/>

Académies : Bordeaux - Limoges - Poitiers

URFIST de Lyon

Université Claude Bernard - Lyon I
Campus de la Doua
20 Avenue Gaston Berger - BP 72215
69603 Villeurbanne Cedex
Tél. 04 72 44 80 86
Fax 04 26 23 45 28

*Responsables : Claire DENECKER,
Jean-Pierre LARDY*

Site Web : <http://urfist.univ-lyon1.fr>

Académies : Clermont-Ferrand - Dijon - Lyon - Grenoble

URFIST de Nice

Petit Valrose - Avenue Joseph Vallot
06108 Nice Cedex
Tél. 04 92 07 67 29
Fax 04 92 07 67 00

*Responsables : Gabriel GALLEZOT,
Michel ROLLAND*

Site Web : <http://urfist.unice.fr/>

Académies : Aix-Marseille - Corte - Nice

URFIST de Paris

17, rue des Bernardins
75005 Paris Cedex
Tél. 01 43 26 85 22 / 34
Fax 01 56 24 97 33

*Responsables : Aline BOUCHARD,
Christophe BOUDRY,
Manuel DURAND-BARTEZ,
Annaïg MAHE*

Site Web :

<http://urfist.enc.sorbonne.fr/>

Académies : Caen - Créteil - Orléans - Tours - Paris - Reims - Rouen - Versailles - Antilles-Guyane

URFIST de Rennes

Université de Rennes II - SCD/Section Lettres

Place du recteur Le Moal - CS 64302
35043 Rennes Cedex
Tél. 02 99 14 14 46
Fax 02 99 14 14 47

Responsables : Marie-Laure MALINGRE, Alexandre SERRES

Site Web :

<http://www.sites.univ-rennes2.fr/urfist/>

Académies : Nantes - Rennes

URFIST de Strasbourg

Université Louis Pasteur (Strasbourg I)
BU de Sciences

34, boulevard de la Victoire - BP 10
67070 Strasbourg Cedex
Tél. 03 68 85 08 00
Fax 03 68 85 08 19

*Responsables : Elsa POUPARDIN,
David VIVARES, Sandrine WOLFF*

Site Web : <http://urfist.u-strasbg.fr>

Académies : Besançon - Nancy-Metz - Strasbourg

Centre de ressources de formations pour la recherche de l'information spécialisée

URFIST de Toulouse

Université de Toulouse I - BIU
11, rue des Puits Creuses
31070 Toulouse
Tél. 05 34 45 61 80 / 81 / 82
Fax 05 34 45 61 85

*Responsables : Sylvie FAYET,
Muriel LEFEBVRE, François MAGNAN*
Site Web : <http://www.urfist.cict.fr>
Académies : Bordeaux - Limoges -
Montpellier - Toulouse

L'enssib (École Nationale Supérieure des Sciences de l'Information et des Bibliothèques) a pour mission de former les conservateurs et les bibliothécaires de l'État et des collectivités territoriales, les cadres des services de documentation et d'information scientifique et technique.

enssib

17-21, boulevard du 11 novembre 1918
69623 Villeurbanne Cedex - France
Tél. 04 72 44 43 43
Fax 04 72 44 43 44
Site Web : <http://www.enssib.fr>

Le service FORMIST (FORMation à l'Information Scientifique et Technique) y a été mis en place pour répondre aux besoins des étudiants, enseignants et bibliothécaires en matière d'éducation à la maîtrise de l'information, de veille et de travail en réseau.

Les ressources repérées par le service FORMIST sont désormais accessibles via le site de l'enssib :
<http://www.enssib.fr/formist/>

Des ouvrages sur la culture de l'information et la recherche d'information sont prévus aux Presses de l'enssib (Collection Papiers, www.enssib.fr/presses/papiers).

Enfin, chaque année à la mi-juin ont lieu les Rencontres FORMIST :
www.enssib.fr/formist/rencontres

FORMIST - enssib

Tél. 04 72 44 43 17
Fax 04 72 44 43 44
*Responsables : Élisabeth NOËL,
Pierre MOISON*
Site Web : <http://www.enssib.fr/formist>

Ont collaboré à cet ouvrage :

Coordination

Élisabeth Noël (enssib)

Comité scientifique

Christophe Boudry (URFIST de Paris)

Ghislaine Chartron (URFIST de Paris)

Lise Herzhaft (URFIST de Lyon)

Jean-Pierre Lardy (URFIST de Lyon)

Élisabeth Noël (enssib)

Coordination technique

Catriona Bergia (enssib)

Émilie Romand-Monnier (enssib)

Développements techniques

Ramzi Abbes (enssib)

Leila Majlal (stagiaire enssib)

Société Araxe

Graphismes

Estelle Mongredien

Maquette

Charvet Imprimeurs (Lyon)

Mise en page

Alexandre Bocquier (enssib)

Le guide REPERE, destiné aux étudiants de Licence essentiellement, propose une approche de la recherche documentaire dans les ressources électroniques. Il présente la démarche de recherche d'information – de la définition du besoin à l'exploitation des résultats – en insistant sur les divers aspects des ressources proposées (accès, usages, coûts, acteurs, législation, typologie, technique). Riche de nombreuses références et adresses Internet, cette brochure résume l'indispensable aux étudiants qui pourront s'initier à la recherche d'information.

Septième édition : septembre 2011
Mise à jour des chapitres
Tous liens Internet vérifiés au 1^{er} juin 2011

ISSN : 1777-9855

Coordination et mise à jour
par Elisabeth Noël
École Nationale Supérieure
des Sciences de l'Information et des Bibliothèques

enssib
17-21, boulevard du 11 novembre 1918
69623 Villeurbanne Cedex
<http://www.enssib.fr>

Cette brochure est distribuée à titre gratuit par l'enssib. Vente interdite.