

L'évaluation des sources

Lorsque vous avez sélectionné des documents en fonction de leur pertinence et que vous les avez identifiés, il reste à évaluer leur fiabilité.

- 1° L'éditeur: à partir des éléments d'identification, le site paraît-il fiable?
- 2° L'auteur: de même, cet auteur est-il compétent sur le sujet? (profession, réputation, bibliographie,...)
- 3° Le contenu: l'évaluation de la fiabilité du contenu sous-entend que vous lisiez attentivement le document pour vérifier s'il répond aux critères suivants:
 - L'actualité de l'information. Voir la page d'accueil où figure le copyright et la date de création du site et surtout vérifier que celui-ci est mis à jour. Si plusieurs liens sont inactifs, cela indique que les mises à jour ne sont pas faites régulièrement.
 - La date du document lui-même est souvent indiquée à la fin de celui-ci. Si elle ne figure pas, on peut essayer de le dater en relevant les événements évoqués dans le texte.
 - La forme: les documents sont-ils bien structurés? quelle est la qualité du style rédactionnel? L'orthographe est-elle soignée?
 - La précision des informations
 - La distinction claire entre le contenu informatif et le contenu commercial
 - La bibliographie

Cette évaluation sera évidemment d'autant plus aisée si vous avez déjà une bonne connaissance du sujet.

En tout cas, que vous utilisiez des sources traditionnelles ou des sources Internet, il y a un grand principe à ne jamais oublier: toujours se référer à plusieurs documents et les comparer.

1. Guide de méthodologie documentaire destiné aux étudiants de 1re année du Coursus de licence de l'Université de Rennes 2, p.5 http://www.uhb.fr/scd/Methodoc_accueil.html
2. Ibidem page 6
3. Ibidem page 7

Fiches pédagogiques

Internet et la recherche documentaire

PAR MICHÈLE STESSEL, ANIMATRICE CIEP-ISCO

Le développement d'Internet a révolutionné la recherche documentaire. Aujourd'hui, nous pouvons accéder à des milliers de documents, en restant à la maison, sur notre lieu de travail ou à l'école. Auparavant, nous devions pour consulter ces mêmes documents, nous rendre dans une bibliothèque, un centre de documentation ou un autre lieu de conservation. Après un trajet plus ou moins long pour y parvenir, il fallait consulter les catalogues sur fiches ou mieux informatisés, remplir les bulletins de commande lorsque les documents n'étaient pas en accès direct, puis attendre que le magasinier les apportent, en espérant qu'ils ne soient pas égarés ou en prêt à l'extérieur. Dans ce cas, il fallait revenir un autre jour. Il fallait aussi composer avec les heures d'ouverture de ces lieux de conservation. Lorsque nous avions besoin d'un document conservé à l'étranger, il fallait entreprendre diverses démarches: soit passer par le prêt «interbibliothèques», soit commander des copies payantes, dans la majorité des cas.

Internet nous permet donc, en principe, de gagner du temps et même dans certains cas de l'argent. C'est le côté positif qui n'est pas à négliger. On pourrait dire aussi qu'il permet à un plus grand nombre de personnes d'avoir accès à l'informa-

tion: beaucoup d'entre elles qui n'auraient jamais mis un pied dans une bibliothèque surfent sur le WEB.

Cependant il y a un revers à la médaille. Les informations que nous y trouvons, sont-elles fiables? Qui les produit? Qui les diffuse? Répondent-elles à nos objectifs de recherche, autrement dit sont-elles pertinentes? Et puis, comment retrouver ces documents fiables et pertinents dans la masse des informations disponibles sur la toile?

Pour parvenir, il est nécessaire, d'une part de connaître les outils de recherche liés à Internet et d'autre part, de procéder à une évaluation critique de ces sources.

REMARQUE IMPORTANTE

Internet met à notre disposition une énorme quantité de documents, mais tous les documents produits de par le monde ne sont pas sur Internet!

Une quantité importante de livres, d'articles de revue, de documents d'archives n'ont pas (encore) fait l'objet d'une numérisation. Les grandes bibliothèques (nationales et universitaires) ainsi que les archives nationales ont commencé ce travail de longue haleine et coûteux.

Cette remarque vaut aussi pour les supports audiovisuels.

Dès lors, lorsque nous entreprenons une recherche, nous devons encore et toujours, consulter les catalogues des bibliothèques. Heureusement, un grand nombre d'entre

elles proposent des catalogues en ligne, exemples: les bibliothèques nationales, les bibliothèques universitaires, celles des ministères, les réseaux de bibliothèques publiques. Ces outils offrent la possibilité de faire des recherches par nom d'auteur, par mots du titre, par thèmes. Cela permet de gagner une étape, en préparant la liste des documents que l'on voudrait consulter sur place.

Il y a aussi tous les documents produits des particuliers ou des petites associations (exemple, un comité de quartier): tous et toutes n'ont pas un site Internet. Prendre contact par téléphone et se rendre sur place restent un bon réflexe.

CARACTÉRISTIQUES DE L'INFORMATION SUR INTERNET

Pour caractériser Internet, en tant que source d'information, on utilise souvent fort à propos des métaphores aquatiques. On le qualifie «d'océan» sur lequel nous devons naviguer tant bien que mal en essayant de ne pas perdre le cap et en évitant les récifs, ou de «déluge informationnel» dans lequel nous risquons de nous noyer, ou encore de «marée noire» pour désigner «l'infopollution» que peut générer un trop grand flux d'informations disparates.

L'information sur le web se caractérise par:

- **une hétérogénéité:** on y trouve toutes sortes de supports: textes, photographies, cartes, documents audiovisuels,... dans toutes sortes de formats informa-

que pour un document imprimé. L'adresse URL fournit des informations intéressantes qui permettent d'identifier le domaine: le nom et le type d'organisation «productrice» du site ou «hôte» des pages consultées et le code du pays.

On peut distinguer ainsi, les sites reconnus, les sites anonymes et les blogues, forums et listes de discussion.

Les sites reconnus garantissent une certaine qualité et fiabilité de l'information. Il s'agit par exemple des sites gouvernementaux, des sites d'institutions supranationales, des sites d'universités, de centres de recherches scientifiques, d'organisations ou d'associations reconnues, de sites d'entreprises.

Par contre, il faut se montrer prudent lorsque l'on utilise des informations provenant de pages personnels: de qui s'agit-il? Cette personne est-elle compétente par rapport au sujet traité?

Les données provenant des blogues et des forums doivent être utilisées avec encore plus de précaution: n'importe qui peut y intervenir et donner son opinion, sans pour autant connaître la question traitée.

Identifier l'auteur

Toutes les questions que l'on doit se poser par rapport à l'auteur d'un document imprimé restent pertinentes par rapport à l'auteur d'un document publié sur le WEB. L'auteur est-il bien identifié? Y-a-t-il une adresse courriel qui permette de le contac-

ter? Y-a-t-il des données relatives à son curriculum qui permettent de vérifier ses compétences, soit dans le document, soit sous forme de liens Internet? Est ce qu'il y a une bibliographie dans le document ou sur une page WEB liée à celui-ci qui permette de vérifier que cet auteur est un spécialiste de la question traitée?

Il est aussi important d'identifier l' (les) intention (s) de l'auteur: informer, expliquer, former? convaincre, faire adhérer à une opinion? vendre un produit? Autrement dit, l'information a-t-elle un but d'information, un but commercial ou un but de propagande?

Identifier le type de document

Trois critères déterminent le type de document:

- 1° Ses limites: s'agit-il de tout un site? d'une page?
- 2° Le genre documentaire: s'agit-il d'un ouvrage? d'un article de revue? d'un article de presse? d'un document officiel?...
- 3° Sa nature: scientifique, destiné au grand public, ..La nature de l'hôte du document donne des indications. Par exemple concernant une recherche sur l'environnement, un site gouvernemental fournira des informations de type informatif et éducatif; un site universitaire donnera des infos de type scientifique, mais aussi formatif; un site d'une organisation non gouvernementale fournira des infos de type informatif, éducatif, mais elles seront l'expression d'une opinion donc plus orientées; un site commercial dont le contenu sera aussi orienté.

La surabondance: elle entraîne une forme d'aveuglement: ce n'est pas l'arbre qui cache la forêt, mais la forêt qui cache l'arbre! Comment dans cette masse d'informations, discerner les informations valables. Beaucoup de personnes face à cette situation se découragent, perdent patience et finalement se limitent aux premiers résultats obtenus sans se préoccuper de leur fiabilité.

La désinformation: celle-ci n'est pas un phénomène nouveau: les «on dit», les colportages, la rumeur existent depuis longtemps. De même, la manipulation de l'information à des fins commerciales ou politiques ne datent pas d'hier.

Mais avec Internet elle prend une dimension mondiale et elle prend aussi de la «vitesse». A titre d'exemple, on peut citer les hoax qui «polluent» les courriels.

Cette désinformation ne résulte pas toujours de mauvaises intentions, mais provient parfois du manque de vérification de la qualité des informations qui circulent. Comme nous l'avons dit au début, ces informations sont très inégales: les études scientifiques sont lors d'une requête mises sur le même pied que des publications de seconde ou de troisième main, des billets d'humeur, des pages personnelles, des publicités commerciales, ect...

La contamination: Contrairement aux sources traditionnelles, il n'y a pas ou peu de contrôle éditorial. Tout le monde peut créer un site aussi bien une université, qu'un réseau mafieux, ou une secte. Il y a aussi

des sites qui sous des apparences scientifiques diffusent des informations racistes ou qui ont des liens avec des sites qui diffusent des propos tendancieux.

L'invasion publicitaire: De nombreux sites sont envahis par des bandeaux et par des fenêtres publicitaires. Certains renvois aussi automatiquement vers un site commercial.

FIABILITÉ ET PERTINENCE DES SOURCES INTERNET

Comme les sources «classiques», les sources Internet doivent être évaluées selon deux grands critères: leur degré de qualité ou de fiabilité et leur degré de pertinence.

Evaluation de la pertinence

Quelques questions permettent d'évaluer la pertinence des informations: sont-elles appropriées? Sont-elles pertinentes par rapport au lieu, à l'époque, à la thématique du champ de la recherche? Le contenu n'est-il pas trop simpliste ou au contraire trop spécialisé?

Identifier la source

Pour pouvoir évaluer, il faut d'abord identifier les différents paramètres de la source que l'on veut utiliser.

Identifier la provenance

Déterminer la provenance d'un document est essentiel pour évaluer sa fiabilité et son intérêt. Cette démarche indispensable pour tout type de source, l'est encore plus lorsqu'il s'agit d'un document provenant d'Internet, mais elle souvent plus difficile

tiques: html, pdf, jpeg, ppt; des logiciels divers,...

- **une instabilité:** les informations font l'objet de modifications, de nouveaux sites sont créés, d'autres disparaissent; idem pour les pages WEB; les adresses sont parfois modifiées ce qui entraîne des difficultés pour retrouver des pages que l'on avait consultées quelques mois auparavant ou bien les liens d'une page renvoient à une autre page qui n'existe plus ou qui a changé d'adresse.

- **une hypertextualité:** un document WEB contrairement aux documents «classiques» (ex.: documents sur support papier) est un document ouvert grâce aux liens hypertextes qui renvoient vers d'autres documents web. Cette possibilité de naviguer sur la toile est un atout, mais il ne faut pas s'y perdre.

- **une juxtaposition et une indifférenciation:** les réponses obtenues suite à une requête via un moteur de recherche, sont présentées sur le même pied, qu'il s'agisse d'ouvrages scientifiques destinés à des spécialistes, d'articles de la presse quotidienne destinés à un large public, de pages personnelles, d'une bibliographie, d'une présentation de programmes de cours, d'une publicité,...

Par comparaison, dans une bibliothèque, des professionnels ont opéré un tri dans les informations en fonction de leur contenu, de leur intérêt, parfois aussi en fonction de leur fiabilité et ils les ont classés suivant une série de critères préétablis: type de support, type d'informations, thématiques,... Autrement dit le lecteur sait quel type de document il consulte;

avec les documents web c'est à lui d'opérer le tri!

- **une multiplicité des sites:** exemples:
 - sites publics / sites avec accès restreint
 - sites gratuits / sites payants
 - sites donnant accès à des textes intégraux / sites dits d'informations secondaires

LES OUTILS DE RECHERCHE

Il est important de connaître les outils de recherche sur Internet pour pouvoir choisir celui ou ceux qui sont les plus adéquats par rapport aux objectifs de recherche que l'on mène.

Ce sont principalement: les annuaires ou répertoires, les moteurs de recherche, les métamoteurs. Comment sont-ils élaborés? par qui? pour quel type de recherche sont-ils adaptés?

On peut aussi faire une recherche par adresse URL.

Les répertoires ou annuaires

«Un annuaire Internet est un inventaire thématique, généraliste ou spécialisé, de ressources Internet, classées par catégories et sous-catégories, selon une classification donnée et navigable par des liens hypertextes.»¹

Ces catalogues thématiques de sites fournissent des informations que l'on peut qualifier de fiables (voir plus loin L'évaluation des sources Internet), car ils sont réalisés par des spécialistes. Chaque annuaire a des correspondants, comme c'est le cas pour certaines revues ou journaux. Ce sont des experts ou professionnels d'un domaine qui

sont chargés de collecter les sites et de les classer. À noter cependant que certains annuaires recensent des sites qui sont directement proposés par leurs éditeurs.

Ces outils sont surtout utiles quand on débute une recherche ou lorsqu'on aborde une thématique relativement large. Par exemple, lors de la phase exploratoire d'un mémoire de fin d'études: vous avez l'idée d'un sujet que vous voudriez traiter mais votre question de départ n'est pas encore très précise. Dans ce cas, une recherche dans un annuaire permet de repérer des sites sur lesquels vous trouverez des études qui vous permettront de cerner le sujet et, par là, de préciser la question.

Parmi les annuaires thématiques généralistes, on peut citer à titre d'exemple: Yahoo <http://fr.yahoo.com>
Un annuaire à recommander parmi les annuaires spécialisés: [revues.org](http://www.revues.org) auquel participent plusieurs revues de sciences humaines et de sciences sociales.

La plupart des annuaires possèdent un moteur de recherche interne au site.

Les moteurs de recherche

«Un moteur de recherche peut être défini comme un outil automatisé d'indexation et de recherche de ressources du web, interrogeable par mots-clés»²

Contrairement aux annuaires qui sont réalisés par des personnes, les moteurs de recherche sont des logiciels (des robots de recherche) qui balayent la toile. Autrement

dit, il s'agit d'une procédure entièrement automatisée.

Chaque moteur comporte trois modules: le module de collecte des données qui balaie le WEB pour récolter les données; le module d'indexation des données; le module de requête pour effectuer les recherches.

L'indexation se fait automatiquement. Sans rentrer dans les détails du fonctionnement des modules, il faut savoir que tous les moteurs n'indexent pas de la même manière et, par conséquent, les réponses à une même question peuvent être différentes suivant que l'on utilise tel ou tel moteur de recherche. D'autre part, la gestion des résultats diffère aussi: par exemple, GOOGLE les classe par indice de popularité, c'est-à-dire en fonction du nombre de liens faits à la page, alors qu' Altavista a opté pour un indice de pertinence, c'est-à-dire en fonction de la fréquence des termes de recherche sur la page.

En conclusion: même si GOOGLE tient aujourd'hui le haut du pavé, utilisez toujours plusieurs moteurs de recherche! Comme pour les annuaires, il existe des moteurs généralistes, exemple: GOOGLE et des moteurs spécialisés, exemple: In-tenso, pour les sciences humaines.

Les moteurs de recherche proposent différents modes de recherche. Ils permettent, en autres, de faire des recherches par mots-clés simples ou avancées à l'aide des opérateurs booléens: et/et/et; et/ou; et/et/mais pas

Important. Pour faire une bonne recherche, il faut bien décomposer sa question en mots-clés. Si pour une requête, vous n'obtenez aucun résultat ou très peu de résultats, demandez-vous si vous avez utilisé les bons mots-clés et réessayez en utilisant des synonymes.

Si au contraire, vous avez une longue liste de résultats, utilisez le mode recherche avancée qui vous permet de préciser votre demande.

Cette remarque vaut aussi pour une recherche dans des catalogues de bibliothèques.

Autre remarque importante qui découle de ce qui a été dit plus haut: les moteurs de recherche sont des outils efficaces lorsque votre requête est suffisamment précise, c'est-à-dire quand vous savez précisément ce que vous recherchez. Pour reprendre l'exemple du travail de fin d'études, il vaut mieux utiliser les moteurs de recherche lorsque la question de départ est bien définie!

Note: Avec GOOGLE, vous pouvez aussi faire une nouvelle recherche à l'intérieur de la liste.

Les métamoteurs

«Un métamoteur est un outil de recherche qui interroge en parallèle plusieurs outils (moteurs, annuaires)». ³ Ils permettent d'interroger plusieurs moteurs à la fois, sélectionnent les documents les plus pertinents dans chaque moteur et procèdent notamment à l'élimination des doublons.

L'adresse URL

Tout document qui se trouve sur le WEB possède une adresse URL: Uniform Resource Locator) peut aussi être un outil de recherche: quand vous consultez le numéro d'une revue spécialisée ou un article de presse, notez l'adresse du site de cette revue ou de ce journal et consultez le: vous aurez une grande probabilité de trouver d'autres articles concernant votre recherche.

Ceci vaut aussi pour les documents audiovisuels.

L'adresse URL est propre à chaque document, un peu comme la cote d'un livre dans une bibliothèque et indique son emplacement exacte sur le réseau ainsi que le protocole pour y accéder. Attention, donc: quand vous transcrivez une adresse de bien respecter la ponctuation, les espaces, les majuscules, les minuscules.

Enfin, n'oubliez pas de noter l'adresse URL des documents consultés, et ce pour deux raisons: pouvoir facilement y retourner, par exemple si vous avez égaré le document et surtout pour pouvoir citer les références dans la bibliographie de vos travaux.

CRITIQUE DES SOURCES INTERNET

Au même titre qu'un développement industriel non contrôlé entraîne une pollution de l'environnement, le développement d'Internet a engendré une «infopollution», ou plutôt des infopollutions. On distingue quatre types: la surabondance de l'information, la désinformation, la contamination et l'invasion publicitaire.