
Annexe F-5

Des outils de recherche dans Internet

Afin d'identifier les sites Web correspondant à vos besoins ou à ceux de vos élèves, vous devrez utiliser des outils de recherche. Vous éprouverez rapidement le besoin d'en connaître le fonctionnement.

Devant la masse d'information disponible dans Internet et son organisation chaotique, on a senti le besoin de développer des **outils de recherche**.

« On compte certainement plusieurs centaines d'outils de recherche généraux dans Internet. Une même recherche lancée sur ces divers outils ne donnera pas les mêmes résultats. D'un jour à l'autre aussi les résultats seront différents. »¹ Ces outils de recherche sont donc en quelque sorte complémentaires.

« En sciences de l'information, on juge une recherche **performante** lorsqu'elle nous procure **tous** les documents recherchés et **que** les documents recherchés. Selon ces critères, la recherche dans Internet comporte un tel taux de **silence** (on ne trouve pas tous les documents pertinents) et de **bruit** (on récolte souvent une multitude de documents non pertinents) qu'on ne peut la qualifier de performante. »²

Il ne faudrait pas que la frustration engendrée par la recherche dans Internet nous prive de la richesse des informations qu'on peut y trouver. Par conséquent, il vaut mieux se familiariser avec quelques-uns de ces outils et surtout, avoir clairement identifié le propos de notre recherche ainsi que les mots-clés qui s'y rattachent.

Les **outils de recherche** se classent en grandes catégories. En voici trois :

- les **robots proprement dits** ou **banques constituées automatiquement**,
- les **catalogues-sujets pourvus d'un moteur de recherche** et
- les **méta-index**.

¹ GIRI — *Guide d'initiation à la recherche dans Internet*

² *Idem*

Les robots proprement dits ou banques constituées automatiquement :

Les moteurs de recherche se comparent aux index des livres. « Derrière chacun se cachent des robots informatiques qui fouillent continuellement les multiples ordinateurs qui composent le Web. »³ On consulte ces moteurs de recherche à l'aide de mots-clés.

« **AltaVista** (<http://www.altavista.com>) est probablement le plus efficace des moteurs de recherche... bien qu'il soit imparfait, comme tous les autres. Il offre notamment la possibilité d'effectuer des recherches en fonction de la langue de chacun des sites. Il existe aussi certaines versions spécialisées d'AltaVista, dont l'Index des sites éducatifs de la francophonie (<http://isef.ntic.org>) qui a été conçu spécifiquement pour les enseignants francophones ». ⁴ [Il offre aussi deux niveaux de procédure de recherche : **Simple query** et **Advanced query**].

Google (<http://www.google.com/intl/fr/>) est un nouveau moteur de recherche dans Internet. Il gagne de plus en plus d'adeptes en raison de ses performances. En effet, **Google** réussit à donner des réponses qui ont un taux de pertinence beaucoup plus intéressant que tous les autres moteurs.

Voici quelques stratégies pour exploiter AltaVista et Google :

Ces moteurs de recherche effectuent une recherche pas mots-clés provenant du *texte* de la page Web (d'autres outils de recherche indexent seulement les titres des documents).

- « **Recherche par mot de même famille** : l'astérisque peut remplacer jusqu'à cinq caractères à la suite d'un préfixe d'au moins trois lettres. Par exemple, la requête **pays*** permettra d'être dirigée vers des pages Web qui contiennent à la fois le mot **pays** et **paysagistes** ». ⁵
- « **Recherche d'une expression ou d'une phrase** : l'insérer entre guillemets. Par exemple, si la requête **Château Frontenac** permet d'obtenir aussi bien les pages qui contiennent le mot **Château** que celles qui contiennent le mot **Frontenac** la requête « **Château Frontenac** » ne permettra que la sélection des pages qui contiennent les deux mots, l'un à la suite de l'autre ». ⁶

³ *Les chroniques de l'Infobourg.*

⁴ *Idem.*

⁵ *Idem.*

⁶ *Idem.*

-
- « **Majuscules et accents** : un mot écrit avec des minuscules appellera des données inscrites dans la banque en minuscules ou en majuscules. Cependant, si un mot est recherché avec une majuscule, le moteur appellera uniquement les mots correspondants avec majuscule. Par exemple, la requête **clément** permettra d'obtenir à la fois les pages qui contiennent le mot **clément** et celles qui contiennent le nom **Clément**, alors que la requête **Clément** n'indiquera que les pages qui contiennent le prénom, avec la première lettre en majuscule ». ⁷

Encore d'autres moteurs de recherche : **HotBot**, **Lycos**, etc.

Les catalogues-sujets pourvus d'un moteur de recherche :

Les répertoires se comparent aux tables de matières que nous retrouvons dans les livres. Leur contenu est indexé par des personnes et non pas par des robots informatiques. « Les milliers de liens qu'ils contiennent sont organisés par catégories [...]. Par un simple clic sur une catégorie, vous obtenez une liste de sous-catégories, puis une liste de liens vers des sites qui correspondent à un champ d'intérêt. » ⁸

Certains de ces grands catalogues offrent aussi une fonction de recherche par interrogation à l'aide de mots-clés.

Yahoo! (<http://www.yahoo.com>) est l'un des répertoires les plus populaires. « Il se démarque par son volume et son organisation ». Il en existe également une version française (<http://www.yahoo.fr>). ⁹

« Voici quelques stratégies [...] pour utiliser Yahoo! à son meilleur

- **Recherche par mot de même famille** : l'astérisque doit être ajouté à la droite d'un préfixe ou d'un mot.
- **Recherche d'une expression ou d'une phrase** : l'insérer entre guillemets.

Notez que cet outil de recherche ne fait aucune distinction entre les minuscules et les majuscules, pas plus qu'entre les caractères accentués ou non ». ¹⁰

⁷ *Les chroniques de l'Infobourg.*

⁸ *Idem.*

⁹ *Idem.*

¹⁰ *Idem.*

« Rapidement, vous découvrirez que le contenu des répertoires de recherche est moins volumineux que le contenu des moteurs de recherche, mais que l'organisation qu'ils présentent, généralement par sujets, facilite grandement leur utilisation. »¹¹

Il existe d'autres catalogues-sujets : **La Toile du Québec** (<http://www.toile.qc.ca>) spécialisé dans les sites québécois ou francophones, **Infomine** (<http://www.mining.com>) très riche répertoire de niveau universitaire, mais en langue anglaise seulement, ...

Les méta-index

Il existe quelques outils permettant d'effectuer une recherche dans plusieurs index à la fois, cela à partir d'un seul formulaire. Ainsi, une même requête sera envoyée simultanément à plusieurs banques de données et le méta-index fera la collection des résultats. De par leur nature, ces systèmes peuvent s'avérer plutôt lents. De plus, ils ne peuvent pas tirer profit au maximum des possibilités particulières de certains outils de recherche. Voici deux exemples de méta-index : **MetaCrawler** et **Copernic**.

- **MetaCrawler** (<http://www.metacrawler.com/index.html>) ne possède pas de base de données. Il utilise plusieurs outils de recherche dont : AltaVista, Excite, Google, Infoseek, Lycos et Webcrawler. Il combine les résultats de tous ces outils de recherche et les place par ordre décroissant de pertinence. Il affiche à l'écran les résultats de tous ces outils de recherche. Il est plus lent que les autres outils; cependant, il offre plus de diversité dans les résultats.
- **Copernic** est un autre exemple de méta-index mais qui demande l'installation d'un logiciel sur son poste de travail. C'est un logiciel québécois bien coté qui s'obtient sur le site Internet <http://www.copernic.com/>. Ce logiciel gratuit peut être installé sur un Macintosh comme sur un ordinateur PC. Il combine plusieurs outils de recherche dont : AltaVista, Deja.com, Excite, HotBot, Infoseek, Lycos, Magellan, WebCrawler et Yahoo!

¹¹ *Les chroniques de l'Infobourg*