

# Préparer ses cours à l'aide d'Internet°: Initiation à la recherche sur la Toile \*

Piet Desmet

Internet, le réseau des réseaux, constitue incontestablement un réservoir de ressources d'une très grande richesse. On y trouve toutes sortes de documents : des textes, des catalogues, des livres ou des revues électroniques, des annuaires, des images, des logiciels, des sons, des vidéos, etc.

Le revers de la médaille est que l'internaute peu expérimenté a parfois le sentiment de « chercher une aiguille dans une botte de foin ». Le but de cette contribution est de présenter quelques stratégies de recherche qui devraient vous aider à naviguer plus efficacement à travers la Toile mondiale. La stratégie la plus évidente pour retrouver un site Web consiste à deviner l'adresse. Voilà pourquoi nous commentons d'abord la structure d'une adresse Web (1.). Si cette stratégie échoue, on peut avoir recours aux listes synthétiques ou « bibliographies » (2.), qui sont des outils de recherche réalisés par des spécialistes du domaine sur lequel vous effectuez une recherche et qui constituent dès lors un excellent point de départ. Viennent ensuite les annuaires ou répertoires – toujours réalisés « à la main », mais qui ne se limitent plus à un seul domaine et qui sont de ce fait déjà plus génériques – ainsi que les moteurs de recherche et les métacheurs (3.), qui sont entièrement automatisés. Récemment, on crée autour des annuaires et/ou des moteurs de recherche des sites qui offrent un éventail plus large de services et qu'on appelle des portails (4.). Nous terminerons cette présentation par quelques conseils pratiques qui devraient vous permettre d'effectuer des recherches plus efficaces (5.)

Reste à préciser que si l'on trouve de tout sur Internet, on n'y trouve pas nécessairement tout ! Dans certains cas, le recours à des ressources documentaires traditionnelles reste donc indispensable.

## 0. Les adresses

### 0.1 Préambule

Une « adresse Web » (ou « adresse URL » ou « URL » [f]) contient un ensemble de données permettant d'avoir accès à certaines pages d'Internet quand on utilise un navigateur Web.

---

\* La version électronique de cette contribution est disponible à l'adresse suivante <http://millennium.arts.kuleuven.ac.be/weboscope/webtheque/recherche2.htm>. Nous tenons à remercier Katrien Luybaert, Yannick Sturbeaux et Henk Verdru de leurs remarques précieuses auprès d'une version antérieure de ce texte.

URL signifie « Uniform (ou Universal) Resource Locator », et définit une sorte d'itinéraire (Locator), menant aux données recherchées (Resource), identique (Uniform) dans toutes les parties du monde et exploitable par tous les logiciels Web ou navigateurs, les plus connus étant Internet Explorer et Netscape Communicator, auxquels sont venus s'ajouter récemment Mozilla (<http://frenchmozilla.sourceforge.net>) et Opera (<http://www.opera.com>).

## **0.2 La structure d'une adresse**

La structure d'une adresse URL peut paraître complexe à première vue. Voici un certain nombre d'adresses :

<http://www.clf.gouv.gc.ca> (Conseil de la langue française)  
<http://www.education.gouv.fr> (Ministère de l'Éducation nationale)  
<http://www.bonjourdefrance.com> (Bonjour de France, journal pour apprendre le français)  
<http://www.arts.kuleuven.ac.be/weboscope>  
<http://www.olf.gouv.qc.ca/ressources/termino/ressling.html>  
<http://www.kulak.ac.be/kulak/kulaklogo.gif>

En fait, on peut comparer la composition d'une URL à la façon dont les fichiers sont organisés sur votre ordinateur personnel. En donnant à votre ordinateur des commandes du type : `c:\my documents\recherche.doc` ou `c:\program files\eudora\eudora.exe`, il peut identifier directement l'endroit où se trouve le fichier en question. En introduisant une URL dans votre navigateur, vous donnez une commande analogue : vous demandez à votre navigateur de charger une page qui se trouve dans tel ou tel répertoire de tel ou tel ordinateur, rattaché au réseau.

*(Advertisie – Encart publicitaire)*

Ces adresses se composent toujours des parties suivantes :

0. Le protocole (= liste de conventions qui précisent la méthode d'accès au document recherché) :

Le protocole HTTP (HyperText Transfer Protocol) est utilisé pour transférer des documents hypertextes ou hypermédias entre un serveur Web (l'ordinateur contenant les pages) et un client Web (l'ordinateur sur lequel vous travaillez). Il n'est d'ailleurs pas obligatoire de reprendre l'indication `http://` dans le champ de saisie du navigateur (LOCATION (Netscape) – ADDRESS (Explorer)).

Si ce protocole est de loin le plus courant, il en existe bien d'autres, notamment le protocole FTP (File Transfer Protocol), qui permet de télécharger des données (des fichiers) d'un ordinateur à un autre.

1. Le nom du serveur (= l'ordinateur hôte (angl. host) qui abrite en principe plusieurs sites). Ce nom commence souvent – mais pas nécessairement – par `www` (World Wide Web) et se termine par l'indication du domaine, qui dévoile la nationalité du serveur ou la nature des informations présentées sur ce serveur :

<code>.com</code> = site commercial	<code>.fr</code> = site français
<code>.edu</code> = site éducatif américain	<code>.be</code> = site belge
<code>.gov</code> = gouvernement américain	<code>.ca</code> = site canadien
<code>.ac.be</code> = université belge	<code>.ch</code> = site suisse

Le nom du serveur se termine par une barre oblique.

2. Le chemin d'accès au document :

Cette partie est facultative. Lorsque le chemin d'accès n'est pas précisé, la barre oblique qui termine le nom du serveur est facultative elle aussi.

Le chemin d'accès précise le nom du dossier (ou répertoire) dans lequel la page (ou fichier) se trouve et le nom exact de la page même.

Le nom d'une page standard se termine par l'extension `.htm` ou `.html`.

Lorsque l'adresse se termine par `.gif` ou `.jpg`, il s'agit d'une image.

Les fichiers sonores ou audiovisuels se terminent par différentes extensions, notamment `.wav`, `.mpg` ou `.avi`.

Le tilde (`~` = alt-0-126) sert souvent à indiquer qu'il s'agit d'un site personnel sur un serveur. De tels sites ont tendance à disparaître plus vite et la qualité des informations fournies doit être vérifiée à fond.

### **0.3 Comment identifier l'adresse d'un site ?**

L'adresse du site où vous vous trouvez est indiquée de plusieurs façons :

0. Elle figure tout d'abord dans le champ de saisie LOCATION (Netscape) ou ADDRESS (Explorer).
1. La barre d'état en bas de l'écran vous indique progressivement les différentes opérations que l'ordinateur exécute. Si vous mettez le pointeur sur un lien hypertexte, le pointeur se transforme en une petite main et l'adresse de la page est affichée dans la barre d'état.
2. En cliquant sur le bouton droit de la souris, on peut activer un petit menu. Il suffit de sélectionner la commande VIEW INFO (Netscape) ou PROPRIÉTÉS (Explorer) pour obtenir l'adresse. Vous aurez en même temps l'adresse des images éventuelles.

#### 0.4 Les erreurs dans une adresse

On ne peut se permettre aucune erreur en introduisant l'adresse dans le champ de saisie. C'est pourquoi il est fortement conseillé de copier et de coller les adresses à partir d'une liste de signets (Netscape) ou favoris (Explorer) plutôt que de les retaper.

Il vous arrivera assez régulièrement de tomber sur une adresse qui ne fonctionne plus. Parfois, il s'agit d'une panne temporaire du serveur. Il se peut aussi que la page recherchée ait changé d'adresse. À ce moment-là, on peut toujours raccourcir l'adresse d'une section, c'est-à-dire jusqu'à la barre oblique précédente. On peut continuer cette opération jusqu'à ce que l'adresse ne contienne plus que le protocole et le nom du serveur. Sinon, on doit avoir recours à un moteur de recherche (cf. ci-dessous) pour retrouver la nouvelle adresse du site en question. Finalement, il n'est pas exclu que l'adresse contienne l'une ou l'autre erreur.

Voici quelques suggestions :

- Cherchez des mots mal épelés (p.ex. *frnech* au lieu de *french*) ;
- Les adresses ne contiennent en principe pas de blancs ;
- Soyez attentifs au choix de la casse. Respectez les majuscules et les minuscules (du moins dans la troisième partie de l'adresse).
- Il est assez facile de confondre :
  - .htm et .html
  - \_ et -
  - . et ,
  - 0 et O
  - 1 et l
  - 2 et Z
  - 5 et S
- Si l'adresse ne se termine pas par .htm ou .html, ajoutez une barre oblique à la fin de l'adresse.

#### 0.5 Testez vos connaissances

(basé sur : <http://www.siu.edu/~aatf/self/test.html>)

##### 0.5.1 Étudiez les adresses suivantes et puis répondez aux questions

0. <http://www.cia.gov/>
1. <http://abu.cnam.fr/BIB/auteurs/nervalg.html>
2. [./cola/dfl/mwall.html](http://cola.dfl/mwall.html)
3. <http://www.culture.gouv.fr/culture/actualites>
4. <http://www.cortland.edu/www/flteach/civ/>
5. <http://www.culture.fr/lumiere/documents/files/oeuvres/David.gif>
6. <http://www.fgov.be/>
7. <http://www.richmond.edu/~jpaulsen/gvfrench.html>

Questions

0. Lesquels de ces sites se trouvent en France ?
1. Lequel de ces sites se trouve en Belgique ?
2. Lesquels de ces sites appartiennent à une université américaine ?

3. Lequel de ces sites est un site officiel du gouvernement américain ?
4. Lequel de ces sites est un site personnel ?
5. Quelle adresse renvoie à une image ?
6. Quelle adresse n'est pas un site Web ?

### 0.5.2 Devinez l'adresse des sites suivants

Pour certains sites (relativement bien connus), il devrait être possible de deviner soi-même l'adresse à base des informations données ci-dessus.

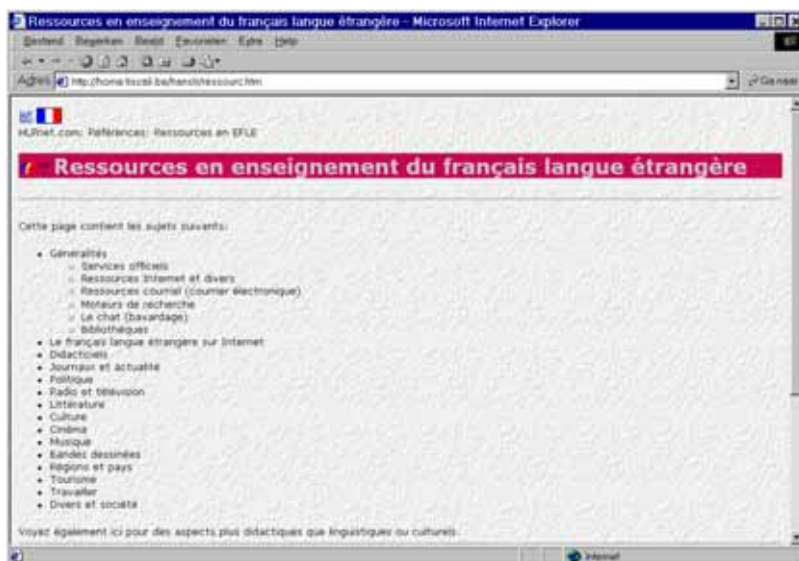
0. Le site de Perrier
1. Le site de La Poste
2. Le site de l'UCL
3. Le site de futuroscope
4. Le site de UCLA (University of California Los Angeles)

## 1. Les listes synthétiques ou « webliographies »

Une liste synthétique ou « webliographie » offre un choix de sites autour d'une thématique particulière. La sélection est faite par des spécialistes du domaine et les sites sont présentés à travers un classement thématique. Souvent les liens hypertextes sont accompagnés d'un bref commentaire informatif et/ou évaluatif.

Pour le FLE, il existe entre-temps plusieurs dizaines de listes synthétiques, qui émanent de facultés universitaires, d'associations de profs ou d'enseignants individuels. Il est évidemment impossible de les présenter toutes dans le cadre de cette présentation. Nous nous permettons de renvoyer à la rubrique « recherche » de notre propre liste synthétique qu'est le weboscope (cf. ci-dessous) qui présente un classement géographique des principales listes pour le FLE.

- Belgique : liste de Hans Leroy  
(<http://home.tiscalinet.be/hanslr/ressourc.htm>)



- Belgique : Virtuele Landen - La douce France (CLT Louvain)  
(<http://www.clt.be/vila/frvi.htm>)



- Belgique : Bopothek  
(<http://bop.vgc.be/bopothek/frans>)
- Belgique : le weboscope

Le weboscope (<http://www.arts.kuleuven.ac.be/weboscope>) est un site qui fournit un inventaire critique des ressources disponibles sur Internet pour la didactique du FLE et s'adresse donc à tout enseignant qui se propose de faire un emploi judicieux d'Internet dans ses cours de français. Récemment, l'interface du site a été complètement revue et la sélection des sites a été actualisée assez systématiquement.



Les premières rubriques du site couvrent les quatre compétences, à savoir

lire (journaux, magazines, littérature, bandes dessinées, etc.),  
écrire (correspondance entre classes, exercices de composition, IRC, etc.),  
écouter (journal parlé, télévision, chansons, etc.) et  
parler (exercices de conversation, programmes d'échange, etc.).

Vient ensuite le domaine des connaissances, regroupant

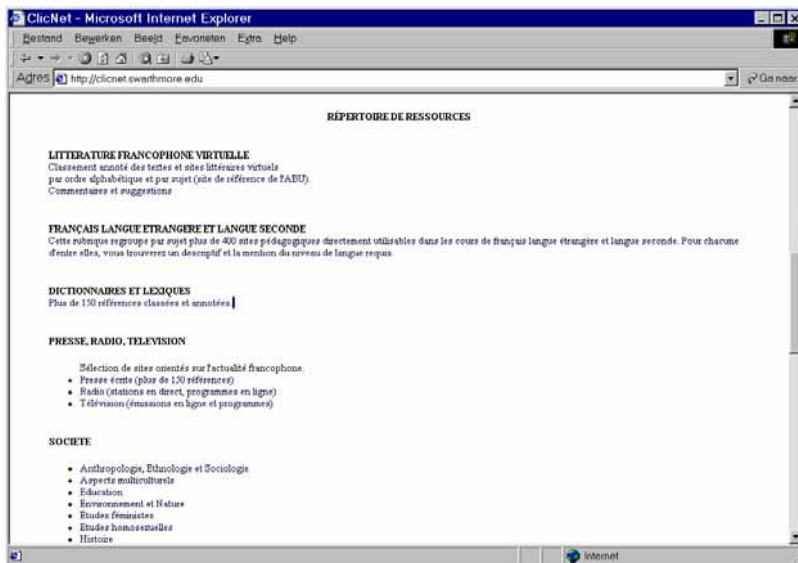
la grammaire (descriptions grammaticales, conjugaison automatique, exercices),  
le vocabulaire (dictionnaires, listes de vocabulaire, exercices) et  
l'interculturel (francophonie, histoire, tourisme, etc.).

Le dernier domaine du weboscope contient des matériaux supplémentaires (préparations de cours, exercices, didacticiels, etc.), des pistes pour la recherche d'informations (moteurs de recherche, listes synthétiques, etc.), et quelques rubriques pour les utilisateurs qui aimeraient s'informer davantage sur les nouvelles technologies.

- France : Noé – un site pour les profs  
(<http://noe-education.org/D14.php3>)
- France : FrancoFil  
([http://www.francofil.net/fr/fle\\_fr.html](http://www.francofil.net/fr/fle_fr.html))
- États-Unis : Le Quartier français du Village planétaire  
(<http://www.richmond.edu/~jpaullen/gvfrench.html>)

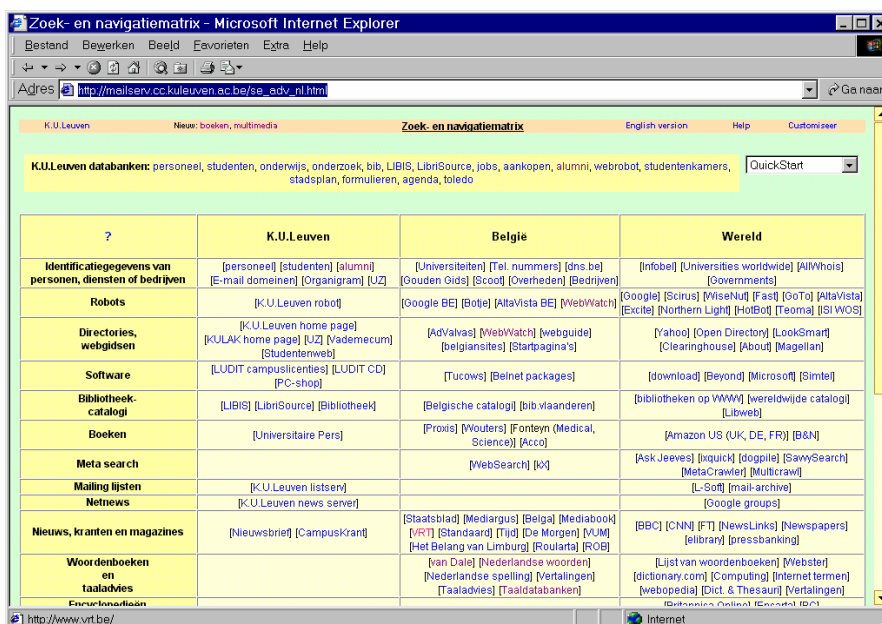


- États-Unis : ClicNet  
(<http://clicnet.swarthmore.edu>)



## 2. Annuaire, moteurs de recherche et métachercheurs

On désigne souvent les outils de recherche sous le terme générique de « moteurs de recherche ». Toutefois, il y a lieu en principe de distinguer entre les annuaires thématiques, les moteurs de recherche et les métachercheurs. Une recherche efficace passe d'ailleurs quasi inévitablement par la consultation de ces différents types d'outils de recherche. Certains sites combinent d'ailleurs un annuaire thématique avec un moteur de recherche. Pour une liste des principaux outils, nous renvoyons à la rubrique « recherche » du weboscope ou à la matrice de recherche de Herman Van Uytven de la K.U.Leuven ([http://mailserv.cc.kuleuven.ac.be/se\\_adv\\_nl.html](http://mailserv.cc.kuleuven.ac.be/se_adv_nl.html)).





## 2.1 Les annuaires ou répertoires

Les annuaires (ou répertoires/guides) thématiques répartissent les sites Web suivant une classification thématique. Ils sont réalisés « à la main » par des équipes qui cataloguent l'information. Contrairement aux listes synthétiques, ces annuaires ne se limitent pas à un seul domaine. Des rédacteurs répertorient et parfois commentent des sites et les présentent par grandes catégories thématiques organisées hiérarchiquement. L'un des annuaires les plus connus est sans aucun doute Yahoo! (Yet Another Hierarchical Officious Oracle) (<http://www.yahoo.com>), dont il existe également une variante régionale Yahoo! France (<http://www.yahoo.fr>). Un autre annuaire utile, consacré exclusivement au Web francophone, est Voila (<http://www.voila.fr>). Parmi les annuaires consacrés au Web francophone, on peut citer encore Excite (<http://www.excite.fr>), Fr@ncité (<http://www.francite.com>), Le Fureteur (<http://www.lefureteur.com>), Lycos (<http://www.lycos.fr>) ou Tiscali (<http://www.tiscali.fr>)

Les annuaires thématiques peuvent être consultés en gros de deux façons :

- La méthode descendante (*top-down*) consiste à partir de la catégorie la plus générale (p. ex. *Sciences humaines*) et à descendre progressivement à des rubriques plus spécifiques (p. ex. *Langues et linguistique*). Cette méthode est particulièrement indiquée pour prendre connaissance des sites tout entiers consacrés à un sujet donné et pour en arriver rapidement à une vue d'ensemble sur ce sujet.
- La méthode ascendante (*bottom-up*) consiste à introduire un mot-clé dans le champ de saisie, ce qui vous permet de passer tout de suite à une rubrique spécifique de l'annuaire. Cette méthode est plus efficace à condition que vous soyez à même de synthétiser votre recherche en un mot-clé spécifique (p. ex. *universités*). Une fois le mot-clé introduit, Yahoo! cherche successivement dans les trois domaines de sa base de données, à savoir les catégories, les sites mêmes et les dépêches récentes d'agences de presse comme UPI ou Reuters<sup>1</sup>.

Yahoo! classe les résultats de la façon suivante :

- Pour la méthode descendante : parmi les catégories correspondantes, celles qui sont situées en haut de l'arborescence (donc plus générales) sont mieux classées que les catégories inférieures (plus précises)
- Pour la méthode ascendante :
  - même principe que pour la méthode descendante
  - les documents dont les mots du titre correspondent aux mots-clés sont mieux classés que ceux dont les mots-clés figurent dans le corps ou dans l'adresse
  - les documents correspondant à la plupart des mots-clés sont mieux classés que les autres.

Il convient finalement de faire remarquer que les annuaires ont également des limites :

- N'y sont présents que les sites qui veulent y figurer. Un site dont l'administrateur

---

<sup>1</sup> Il convient de faire remarquer que Yahoo! cherche en fait encore dans une quatrième base de données, qui est celle du moteur de recherche Google (cf. 3.2.). Suite à cette alliance entre Yahoo! et Google, nous disposons d'un outil particulièrement puissant qui combine un des meilleurs annuaires avec un des moteurs de recherche les plus performants. Il faut bien évidemment se rendre compte que les résultats de Google ne sont pas contrôlés par la rédaction de Yahoo!.

ne désire pas qu'il figure dans Yahoo! n'y sera pas indexé, sauf si quelqu'un d'autre le propose à la place du responsable (éventualité peu probable).

- N'y sont présents que les sites satisfaisant aux critères de classement : contenu du site, affichage de logos, de bannières, etc. Il y a donc un tri, un choix, voire une censure.

### **Exercice pratique**

Servez-vous de Yahoo! France pour trouver :

- des sites sur l'enseignement du français langue étrangère (méthode descendante)
- des sites de la presse francophone (journaux/magazines) (méthode descendante)
- des sites sur la francophonie (méthode ascendante)

Effectuez éventuellement les mêmes recherches à l'aide de l'annuaire Voila.

Comparez les résultats des deux annuaires.

### **2.2 Les moteurs de recherche**

Les moteurs de recherche sont entièrement automatisés : des robots parcourent régulièrement les pages Web, et réalisent l'indexation partielle ou totale de leur contenu, constituant ainsi de gigantesques bases de données interrogeables par mots-clés. Voilà pourquoi les moteurs de recherche ont de loin les bases de données les plus importantes et les mieux mises à jour. Parmi les moteurs de recherche les plus importants, on peut citer AltaVista (<http://www.altavista.fr>) et surtout Google (<http://www.google.fr>), qui est sans doute le moteur de recherche le plus performant.

Ce type d'outil est particulièrement approprié à la recherche portant sur des termes spécifiques et pointus, des noms propres ou des expressions. Il convient d'ailleurs de formuler sa requête de la façon la plus précise possible (cf. ci-dessous).

Si quasiment tous les moteurs de recherche fonctionnent à partir de mots-clés, il en existe quelques-uns qui permettent de formuler une recherche sous la forme de phrases intégrales (en langage naturel). Ainsi, dans Ask Jeeves (<http://www.ask.com>) vous pouvez introduire n'importe quelle question... en anglais. Le moteur vous propose différentes reformulations de votre question, à partir desquelles vous pouvez effectuer une recherche.

### **2.3 Les métachercheurs ou multi-moteurs de recherche**

Les métachercheurs fournissent un instantané des résultats fournis par les plus gros moteurs de recherche (une sorte d'écrémage), tout en intégrant les résultats (pour éviter de voir deux fois le même site). Ixquick (<http://ixquick.com/fra>) est un métachercheur avec une interface en français. Parmi les métachercheurs américains les plus connus, on peut citer Search.com (<http://www.search.com>), Mamma (<http://www.mamma.com>) et KartOO ([www.kartoo.com](http://www.kartoo.com)).

Les métachercheurs donnent une idée des sites que l'on va obtenir, de leur façon de traiter le sujet et un certain nombre d'éléments importants :

- Le nombre de résultats par moteur de recherche montre ceux sur lesquels on trouvera le plus grand nombre d'informations.

- Les termes employés dans les titres et les descriptions des sites potentiellement intéressants.

En fait, mieux vaut commencer une recherche par la consultation d'un outil de ce type. Il ne faut pas hésiter à effectuer plusieurs recherches avec des termes différents et sur plusieurs métachercheurs. Une fois que l'on a ces points de départ, on peut passer à la recherche sur les moteurs ou sur les annuaires.

### 3. Les portails

Récemment, on essaie de créer autour des annuaires et/ou des moteurs de recherche existants des sites qui offrent un éventail de services plus large. Ces portails ou portes d'accès au réseau Internet peuvent proposer, par exemple, un service de courrier électronique gratuit, une sélection de moteurs de recherche, des bavardoirs ou salons (de discussion) (*chatrooms*), les actualités, la météo, les cotes de la Bourse, un classement de sites par thèmes (ex. : cinéma, voyage, sport), des raccourcis pour les achats en ligne, etc. L'objectif des portails est de devenir le site de démarrage du plus grand nombre d'internautes.

Parmi les portails les plus importants dans le monde francophone, on peut citer non seulement Microsoft Network ou MSN (<http://www.msn.fr>) et le Netscape Netcenter (<http://www.netscape.fr>), mais aussi Club-Internet (<http://www.club-internet.fr>). De plus, les sites des annuaires comme Yahoo! France ou Voila se profilent de plus en plus comme de véritables portails. Parmi les portails belges, on peut citer notamment AdValvas (<http://www.advalvas.be/indexfr.html>) ou Skynet (<http://www.skynet.be>).

De plus, certaines listes synthétiques ou « webliographies » essaient également d'offrir la plupart des fonctionnalités prévues dans un portail. Un des exemples les plus réussis est sans doute La porte des lettres, un portail sur les lettres et les sites éducatifs en français (<http://www.portail.lettres.net>). Un site aux objectifs quelque peu plus larges est Educlic, le portail des professionnels de l'éducation (<http://educlic.education.fr>). Par ailleurs, le site Momes.Net, le portail de la communauté internationale des jeunes francophones, vaut certainement le détour (<http://www.momes.net>).



## 4. Comment effectuer une recherche efficace ?

### 4.1 Conseils généraux

- Utiliser plusieurs moteurs de recherche simultanément (pour autant que l'on n'ait pas encore reçu de réponse satisfaisante après avoir consulté un seul moteur de recherche). Chaque moteur couvre un nombre de pages différent des autres et sa méthode d'indexation lui est propre.
- Utiliser toutes les fonctionnalités d'un moteur de recherche : beaucoup de moteurs proposent des recherches avancées et la possibilité de filtrer les résultats obtenus à l'aide d'opérateurs ou d'expressions booléennes. Cela permet de limiter considérablement le nombre de résultats en laissant de côté les pages inutiles.
- Être le plus spécifique possible : il faut éviter si possible les termes vagues ou trop génériques, comme par exemple, sport, voiture, informatique, Internet, histoire, etc. Surtout sur les moteurs de recherche importants, il faut veiller à employer des mots-clés qui cernent le plus précisément possible l'objet de sa recherche.

Les deux conséquences d'une mauvaise requête sont les suivantes :

- Pas ou peu de résultats : la requête est trop spécifique, il faut élargir le champ de recherche en ouvrant la requête. La base de données du moteur de recherche est peut-être aussi trop limitée pour vos besoins.
- Trop de résultats : il faut restreindre sa requête en employant des mots-clés plus spécifiques.

Un bon réflexe pour faire des requêtes pertinentes :

N'avoir aucun ou trop de résultats revient dans la pratique à la même chose : ça ne sert à rien. Mieux vaut restreindre un peu trop son champ d'investigation et perdre 20 % de sites intéressants, qu'être trop général et en perdre 80 %, faute de pouvoir dépouiller tous les résultats. De toute façon, il est impossible de faire une requête parfaite qui engloberait tous les sites intéressants, ne serait-ce que parce qu'un moteur de recherche ne possède jamais à lui tout seul toutes les références intéressantes.

### 4.2 La définition du sujet de sa recherche

L'absence de résultats intéressants à une recherche sur le Web a souvent pour cause une mauvaise définition du sujet de la recherche. On doit pouvoir exprimer le sujet d'une recherche sous plusieurs angles :

- Sa **catégorie** : sous quel angle veut-on aborder le sujet ? Scientifique, littéraire, testimonial, humoristique, historique, poétique, etc. Par exemple, si on cherche des renseignements sur le tennis, veut-on des statistiques économiques, une histoire de ce sport, des informations pratiques ? Chaque angle d'étude fera l'objet d'une méthode de recherche particulière, utilisant des outils spécifiques.
- Le sujet d'une recherche s'exprime aussi en **mots-clés** : quels sont les termes, expressions, exemples qui définissent le mieux le sujet de sa recherche. Il faut tout simplement définir une liste de mots que l'on retrouvera certainement dans les pages que l'on recherche. Pour reprendre l'exemple précédent sur le tennis, si on

cherche des informations historiques, on pensera au terme « jeu de paume », mais si on veut des résultats sportifs, on emploiera « ATP », entre autres termes.

### 4.3 Cerner sa recherche : les opérateurs

Sur les moteurs de recherche généraux, c'est-à-dire ceux ayant de très importantes bases de données, n'importe quelle requête simple retourne plusieurs dizaines de milliers de résultats. Sauf chance exceptionnelle, les sites potentiellement intéressants ayant trait au sujet recherché sont noyés dans des milliers d'informations inutiles. Lorsque c'est possible, l'utilisation d'opérateurs restreint considérablement le champ de la recherche en diminuant le « bruit » (i.e. tous les résultats inutiles générés par une requête).

Il convient de préciser que les tableaux ci-dessous présentent une utilisation théorique des opérateurs, basée sur les conventions proposées par Altavista. Dans la pratique, chaque moteur de recherche a sa syntaxe propre. Les variations et les spécificités sont indiquées dans les aides de ces moteurs.

#### 4.3.1 Limitation de certains champs

opérateur	fonction	exemples
<b>title:</b> <i>x</i>	Le mot ou la phrase doit être dans le champ <i>titre</i> (ou dans le <i>méta-titre</i> ). Attention : pas d'espace entre <i>title</i> , les deux points et l'expression !	<b>title:</b> grammaire française <b>title:</b> “musée du louvre”
<b>url:</b> <i>x</i>	Le mot doit faire partie de l'adresse du site.	<b>url:</b> <b>fle.fr</b>
<b>link:</b> <i>x</i>	Cette commande permet de chercher quel site a établi un lien vers cette page. Deux pages liées traitent souvent du même sujet, ou sont au moins complémentaires.	<b>link:</b> <b>www.fle.fr</b> <b>link:</b> <b>www.olyfran.com</b>
<b>image:</b> <i>x</i>	Recherche les images dont le nom contient <i>x</i> .	<b>image:</b> “Arc de Triomphe”
<b>anchor:</b> <i>x</i>	Recherche les pages ayant des boutons ou des liens contenant <i>x</i> dans le texte du lien hypertexte.	<b>anchor:</b> <b>dictionnaires</b>
<b>domain:</b> <i>x</i>	Recherche les pages appartenant à un domaine déterminé.	<b>domain:</b> <b>edu</b> <b>AND NOT domain:</b> <b>com</b> (permet d'exclure les sites commerciaux)

### 4.3.2 Recherche de phrases

“x”	Recherche les pages contenant littéralement x.	“Il est cinq heures. Paris s’éveille” “vache folle”
-----	------------------------------------------------	--------------------------------------------------------

Inutile de vous rappeler que cette fonction est particulièrement utile lorsque vous croyez qu’un élève a recopié tel ou tel passage du Web sans citer ses sources !

En utilisant les guillemets, on peut même se servir du Web comme un énorme corpus électronique. On peut y trouver un début de réponse à un certain nombre de questions d’ordre linguistique.

- P. ex. “cliquer un lien” (5 pages) vs “cliquer sur un lien (510 pages)  
 “une URL” (16 054 pages) vs “un URL” (1 240 pages)  
 Que signifie le sigle “AFNOR” ?

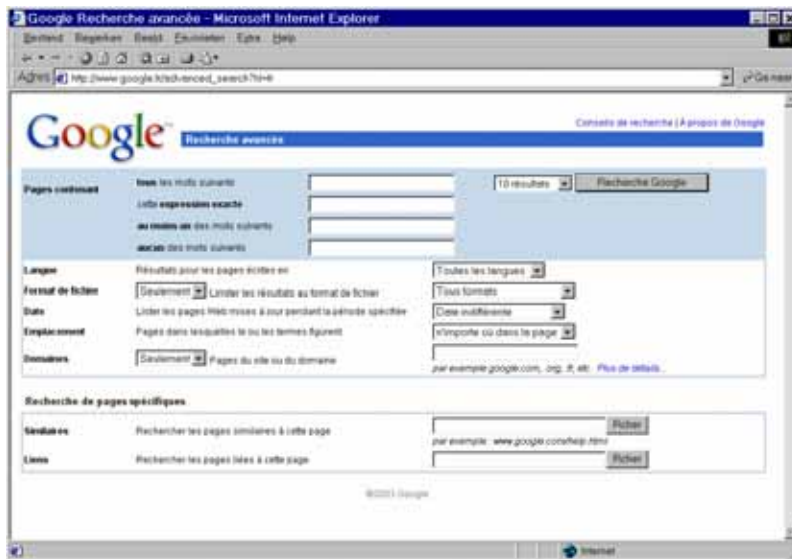
### 4.3.3 Troncation

xxx*	Le joker (*, angl. <i>wildcard</i> ) remplace une suite de caractères (max. 5). Il peut se placer également au milieu du mot. Le joker ne remplace ni les majuscules, ni les chiffres, ni les signes de ponctuation.	<b>grammaire*</b> trouve grammaire et grammaires <b>francopho*</b> trouve francophone, francophones, francophonie, etc. <b>Dosto*evsk*</b> trouve Dostoïevski, Dostoïevsky, Dostoievski, etc.
------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### 4.3.4 Recherche booléenne

x AND y	Recherche les pages contenant x ET y	“auberge de jeunesse” AND Paris
x OR y	Recherche les pages contenant x OU y.	Baudelaire OU “Les Fleurs du Mal”
x AND NOT y	Recherche les pages contenant x MAIS NON PAS y.	festina AND NOT dopage
x NEAR y	Recherche les pages dans lesquelles x est éloigné de y d’au plus 10 occurrences.	Prince NEAR Philippe
( )	Les parenthèses groupent des actions	dopage AND (football OR cyclisme)

Dans la plupart des moteurs de recherche, on peut se passer de ces expressions booléennes en activant le menu « recherche avancée » qui présente ces mêmes opérations en langage naturel.



#### 4.3.5 La casse

<p><b>Majuscules, minuscules et caractères accentués</b></p>	<p>Dans les requêtes, les mots en minuscules équivalent à des mots en minuscules, en majuscules ou en majuscules et minuscules mélangées. Les majuscules équivalent uniquement à des majuscules. Même principe pour les accents.</p>	<p><b>aimé</b> équivaut à Aimé, AIME ou aIMé.</p> <p><b>Aimé</b> équivaut à Aimé, AiMé, mais non pas à aimé.</p> <p><b>AIME</b> n'équivaut qu'à AIME.</p>
--------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------

On ne saurait donc sous-estimer l'importance de la casse dans les moteurs de recherche. En règle générale, n'utilisez les majuscules que pour chercher un nom propre existant aussi en tant que nom commun. Par exemple, si je recherche des renseignements sur la région de la Champagne, je taperai "Champagne" et non "champagne".

### 5. Quelques applications concrètes

Comment procéderiez-vous pour rechercher les informations suivantes ?  
Quelles sont les différentes stratégies de recherche possibles ?

0. Un site sur Jean-Jacques Goldman (données biographiques, discographie, paroles de ses chansons, fragments).
1. Un site sur le métro parisien.
2. Les sites des partis politiques français.
3. Les paroles de la chanson "À tâtons" d'Axelle Red.
4. Un texte sur :
  - la maladie de la vache folle,
  - l'emploi du GSM par les jeunes.
0. Les résultats d'un sondage sur les élections présidentielles de 2002.

1. La une d'un journal français d'hier.
2. Des exercices grammaticaux.
3. Une image de la Grande Arche de la Défense.
4. Des matériaux en vue de la préparation d'un cours consacré au Petit Prince.

## 6. Bibliographie

Cette présentation se base directement sur les sites et les ouvrages suivants ainsi que sur les fonctions d'aide des outils de recherche :

- Abrate, Jayne. 1999. *Utiliser le World Wide Web dans la classe de français : tout pour le débutant et même quelque chose pour l'internaute rusé.*  
<http://www.siu.edu/~aatf/self/begin.html>
- Chirurgien, Estelle. 2001. *Comment trouver ?* Paris: Éditions MultiMondes.
- Corda, Alessandra. 1999. *Internet in het talenonderwijs.* Bussum : Coutinho.
- Crepuq. 1996. GIRI - Guide d'initiation à la recherche dans Internet.  
<http://www.bibl.ulaval.ca/vitrine/giri>.
- Davies, Nancy. 1999. *Activités de français sur Internet.* Paris : CLE International.
- Dubreucq, Martine. 2002. "Trouver l'aiguille dans la botte de foin". *Le Français dans le Monde* 321. 39-41.
- De Grooff, Dirk. 1999. *Snel en efficiënt zoeken op het Internet.* Leuven : Davidsfonds.
- Rousseau, M. 1999. *Le Manuel de survie du chercheur sur Internet.*  
<http://www.multimania.com/mrousseau>.
- Truyen, Fred – Tolleneer, Jan. 1998. *Historische Informatiekunde. Beheer van tekstuele data met de Personal Computer.* Leuven : Acco.

## "Corrigés" des applications

### 1.5.1 Étudiez les adresses suivantes et répondez aux questions

0. Sites français : 2, 4, 6
1. Site belge : 7 (site du gouvernement fédéral)
2. Université américaine : 5, 8
3. Gouvernement américain : 1 (= Central Intelligence Agency)
4. Site personnel (cf. présence du tilde) : 8
5. Image : 6 (cf. format gif)
6. Pas un site Web : 3 (protocole http fait défaut)

### 1.5.2 Devinez l'adresse des sites suivants

0. [www.perrier.fr](http://www.perrier.fr) (= site de Perrier France) / [www.perrier.com](http://www.perrier.com) (= site en anglais)
1. [www.laposte.be](http://www.laposte.be)
2. [www.ucl.ac.be](http://www.ucl.ac.be)
3. [www.futuroscope.com](http://www.futuroscope.com) (ou <http://www.futuroscope.fr>) /
4. [www.ucla.edu/](http://www.ucla.edu/)



### 3.1 Les annuaires – exercice pratique

#### 0. Enseignement du FLE :

Yahoo! : enseignement et formation / par secteurs et matières / enseignement des langues / par idiomes / français comme langue étrangère : associations et organismes (23 sites), ressources pédagogiques (8 sites) ; listes synthétiques (4 sites)

(méthode ascendante : introduire “français langue étrangère” dans le champ de saisie : 16 sites dans l’annuaire)

Voilà : enseignement / formation, disciplines / langues / français, français langue étrangère : 49 sites (méthode ascendante : 62 sites)

#### 1. Presse francophone :

Yahoo! : actualités et médias / journaux et magazines: 658 sites

Voilà : actualité / presse : presse nationale d’actualité (12 sites), presse périodique régionale (56 sites), presse spécialisée (457 sites), etc.

#### 2. Francophonie : “francophonie” :

Yahoo! : 2 catégories et 83 sites

Voilà : 52 sites

## 6. Quelques applications concrètes

#### 1. Jean-Jacques Goldman :

- annuaire comme Yahoo! France : “Jean-Jacques Goldman” : rubrique « annuaire de sites » 19 sites, sélectionnés par la rédaction (= sans doute la meilleure piste)
- moteur de recherche comme Altavista ou Google : “Jean-Jacques Goldman” (n’oubliez ni les guillemets ni les majuscules !) -> plusieurs dizaines de sites
- la même recherche + recherche « MP3/Audio » : aucun résultat (pas de fichiers mp3 disponibles)
- liste synthétique comme Le weboscope : rubrique écouter / chansons

#### 2. Le métro parisien :

- deviner soi-même l’adresse : [www.ratp.fr](http://www.ratp.fr)
- annuaire comme Yahoo! France : essayez d’avoir un mot-clé aussi précis que possible -> “RATP” (1 catégorie, 14 sites) mieux que “métro” (11 catégories, 211 sites) mieux que “transports” (269 catégories, 784 sites)
- moteur de recherche comme Altavista : ici aussi plus le mot-clé est précis, plus le résultat sera satisfaisant. (p.ex. métro NEAR Paris / métro AND paris\*)
- liste synthétique comme Le weboscope : rubrique interculturel / transports (France : Air France, SNCF, RATP – Belgique : Thalys, TEC, STIB)

#### 3. Les partis politiques français :

le meilleur outil de recherche est clairement un annuaire comme Yahoo! France : “partis politiques” -> 32 catégories (e.a. Institutions et politique / Politique / Partis Politiques) + 11 sites (e.a. liste des partis politiques établie par le gouvernement français)

4. *À tâtons* d'Axelle Red :

- introduire "A tâtons" (n'oubliez ni les guillemets ni les majuscules) dans un moteur de recherche comme Alta Vista (+ langue de recherche : French) ou Google
- liste synthétique comme le weboscope : écouter / chansons

5. Texte sur la maladie de la vache folle/sur l'emploi du GSM par les jeunes :

la meilleure piste est sans doute de consulter les archives d'un journal ou d'un magazine. Pour ce faire, consultez la rubrique « lire » du weboscope, puis « journaux » ou « magazines ». Prenez p.ex. Le Soir, passez à la rubrique « archives » et tapez "vache folle"/ "GSM" AND "jeunesse". De même, la consultation des archives du Nouvel Observateur vous donne tout de suite quelques articles sur la vache folle.

6. Sondage :

- moteur de recherche comme Google : « sondage » et « élections présidentielles » (Rechercher dans « pages francophones »)
- consultez un annuaire comme Yahoo! France : "sondage"-> 12 catégories, parmi lesquelles on trouve : Exploration géographique / Pays / France / Institutions et politique / Politique / Scrutins / Sondages et enquêtes d'opinion -> permet de passer à la rubrique « Sofres France » et d'accéder ainsi au site : "Présidentielle 2002"

7. La une d'un journal français d'hier :

- annuaire comme Yahoo! France : catégorie actualités et médias / journaux
- moteur de recherche comme Altavista : recherche avec le titre d'un journal comme mot-clé
- liste synthétique comme le weboscope : rubrique lire / journaux

8. Des exercices grammaticaux :

liste synthétique comme le weboscope : rubrique grammaire / exercices (classés par sujet grammatical et par niveau !) ou rubrique matériaux / exercices / grammaire

9. Image de la Grande Arche :

moteur de recherche comme Google, rubrique « Images » : "Grande Arche" (545 images) ou Altavista, rubrique « Image »: "Grande Arche" (366 images)

10. Cours – Petit Prince :

- liste synthétique comme Le Quartier français du village planétaire – rubrique « Activités – Petit Prince »
- annuaire comme Yahoo! France : 1 catégorie et 13 sites dans l'annuaire de sites