

LA RECHERCHE DOCUMENTAIRE SUR INTERNET

Dossier Elèves

PREAMBULE

EXEMPLE 1 : Recherches sur l'euro pour l'illustration d'un jeu

EXEMPLE 2 : Recherches pour la réalisation d'une revue de presse

EXEMPLE 3 : Recherches de plans pour construire un télescope

CONDUIRE PAR ETAPES UN PROJET DE RECHERCHE

COMMENT TROUVER LE BON SITE ?

- Mots-clés
- Moteurs de recherche
- Méta-moteurs
- Annuaires
- Portails

PREAMBULE

Votre investissement dans les P.P.C.P. vous procure de multiples atouts pour vos apprentissages et votre vie future. Parmi eux figure en bonne place l'utilisation des nouvelles technologies et notamment l'accès à Internet. C'est l'occasion pour vous de développer de précieuses habiletés et de reculer les limites de l'environnement habituel pour partager les ressources mutualisées aux quatre coins de la planète.

Ce nouvel espace de liberté auquel vous pouvez accéder, faisant de vous un citoyen du monde, est assorti de deux qualités essentielles :

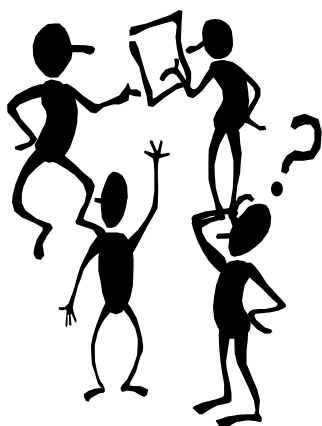
- la confiance que tout un chacun est prêt à vous accorder ;
- le respect, librement consenti par vous même, des règles établies au sein de votre établissement et de votre classe pour une utilisation des ressources conforme à l'éthique et à la morale.

Puisse ce guide, écrit pour vous, vous aider efficacement à apprivoiser cet univers d'informations qu'est Internet !

EXEMPLE 1

La classe de terminale B.E.P. Métiers de la Comptabilité a, en concertation avec leurs professeurs, choisi, dans le cadre des PPCP, un thème d'actualité : "L'EURO". Il s'agit d'élaborer un jeu de type "trivial poursuite" dont les questions se rapporteraient à l'euro.

Un groupe d'élèves de la classe a en charge la confection des cartes à jouer. Outre les questions qui doivent figurer sur celles-ci, ils souhaitent les illustrer à l'aide d'images représentatives de la monnaie. Ils décident donc de procéder à une recherche sur le Web.



Dans un premier temps, ils réfléchissent ensemble (remue méninges) à la problématique afin de bien **cerner le sujet** et d'éviter les fausses routes. Ils ont besoin des images des billets et des pièces de monnaie afin de pouvoir les reproduire sur les cartes à jouer. Le mot clé associé à la recherche est évidemment "euro" mais, utilisé seul, il fournira des réponses liées au "passage à l'euro", à "l'importance juridique et sociale de l'euro", aux "incidences économiques de l'euro"..... Les mots-clés pièces, billets, monnaie sont donc également à considérer.

Le **recensement des sources d'information** à même de fournir les réponses attendues est rapide car ils ne disposent ni d'adresses de sites connus, ni de revues ou de bases de données spécialisées. Ils optent donc pour l'utilisation du moteur de recherche Alta Vista

Compte tenu des mots-clés précédemment identifiés, ils décident de **chercher l'information** à l'aide de la requête : "euro+pieces+billets".

Après avoir **consulté** les résultats obtenus au nombre de 7, ils éliminent les pages trouvées en langue anglaise (sans aucun doute, il aurait fallu limiter la recherche au Web francophone) et

sélectionnent 3 sites dont le descriptif est en rapport avec leur préoccupation :



[Site du Premier ministre - portail du Gouvernement français](http://www.premier-ministre.gouv.fr/fr/p.cfm?ref=21688)

Premier ministre et gouvernement - Actualité politique et internationale, informations sur le gouvernement, le Premier ministre et les différentes...

<http://www.premier-ministre.gouv.fr/fr/p.cfm?ref=21688>

[euro monnaie unique](http://www.chez.com/eurodevises/)

Découvrez l'union européenne la mise en place de l'euro pièces billets doc audios et videos liens importants

<http://www.chez.com/eurodevises/>

[europibi](http://netia59.ac-lille.fr/was/0594480Z/europibi.htm)

L'Euro : PIÈCES & BILLETS. 1 cent. 5 euro. 2 cent. 10 euro. 5 cent. 20 euro. 10 cent. 50 euro. 20 cent. 100 euro. 50 cent. 200 euro. 1 euro. 500...

<http://netia59.ac-lille.fr/was/0594480Z/europibi.htm>

L'*extraction* des données permet de recueillir des documents intéressants, mais l'analyse de ceux obtenus sur les deux premiers sites sont soit incomplets, soit impossibles à reproduire séparément, conformément aux critères de départ.



Par contre, ils décident d'*exploiter* ceux du troisième site (copié collé partiel ci-dessous) :

	1 cent		5 euro
	2 cent		10 euro
	5 cent		20 euro

L'évaluation des recherches est positive, car les résultats trouvés sont suffisants et répondent aux objectifs fixés. Qui plus est, ils ont été trouvés sur le site de l'académie de LILLE et peuvent donc être **considérés comme fiables**.

Afin de **communiquer** aux autres groupes de la classe le fruit de leurs travaux, le tableau contenant les représentations des billets et pièces est copié-collé, la référence du site consulté est notée, puis le fichier est sauvegardé.

EXEMPLE 2

4 PPCP différents ont été retenus par la classe de première baccalauréat professionnel Secrétariat de MACON. Les 4 groupes d'élèves qui se sont constitués devront, à tour de rôle, mener à terme chaque PPCP.

Un projet a pour objet la réalisation d'une revue de presse économique et juridique (presse locale, nationale, internationale, médias et multimédias) fondée sur l'actualité des 15 derniers jours. Le groupe d'élèves qui a en charge la recherche des informations devra ensuite les présenter aux autres groupes de la classe afin de se familiariser à communiquer oralement.

Dans un premier temps, ce groupe d'élèves prend le temps de la réflexion afin de bien **définir le sujet à traiter**. Que cherche-t-on ? De quelles informations a-t-on besoin ? Quels sont les organismes de presse, les médias susceptibles de détenir ce type d'informations ? Quels mots-clés utiliser ?...



Compte tenu du caractère particulier des informations recherchées (informations liées à l'actualité), ils décident d'**utiliser comme source d'information** l'annuaire Yahoo France.



Actualités et médias

[Sujets d'actualité](#), [Télévision](#), [Journaux](#)

Institutions et politique

[Ministères](#), [Droit](#), [Services publics](#)

Art et culture

[Littérature](#), [Cinéma](#), [Musique](#), [Musées](#)

Références et annuaires

[Dictionnaires](#), [Annuaire](#), [Bibliothèques](#)

Commerce et économie

[B2B](#), [Shopping](#), [Emploi](#), [Immobilier](#)

Santé

[Diététique](#), [Médecine](#), [Organismes](#)

Divertissement

[À voir](#), [Loteries](#), [Humour](#), [Sorties](#)

Sciences et technologies

[Animaux](#), [Astronomie](#), [Physique](#)

Pour **chercher l'information** à partir de la page d'accueil de l'annuaire, ils **consultent** la liste des thèmes proposés, **choisissent** la rubrique Actualités et médias, puis ils cliquent sur la rubrique Revue de presse et parmi les sites proposés ils écartent les pages sans lien avec le domaine économique et juridique et ils **sélectionnent** [dernieres.com](#) et [e.revue](#) qui sont des revues de presse quotidiennes possédant une rubrique "Economie". Après avoir identifié les points d'accès à l'information, le groupe **consulte** et **trie** les données. Celles, les plus significatives, sont **extraites** et **copiées** dans un fichier du traitement de texte.

Après analyse, elles font l'objet d'un **traitement** en vue d'être exploitées. Celles trop anciennes, celles dont les auteurs ne sont pas identifiés et celles à la marge sont écartées.

L'attachement à leur région (la Bourgogne) et plus particulièrement à leur cadre de vie incitent les jeunes à compléter leur travail à partir de nouvelles locales. Ils explorent alors l'abondance Journaux → Actualité locale et régionale → Bourgogne → Journal de Saône-et-Loire et procèdent à l'extraction des données qui figurent à la rubrique "Economie".

Compte tenu qu'ils ne pourront pas procéder à ces recherches tous les jours, ils souhaitent accéder aux informations économiques et juridiques des jours antérieurs. Ils s'engagent donc une nouvelle recherche en sélectionnant cette fois Sujets d'actualité qui permet d'accéder à diverses rubriques dont une, relative à l'économie, elle-même subdivisée en secteurs.

Économie	
 Difficultés du secteur aéronautique	19 septembre, 18h16
 Emploi	19 septembre, 20h05
 FMI et Banque mondiale	18 septembre, 7h32
 La préparation à l'euro sur Yahoo! Finance	
 Les marchés pétroliers	18 septembre, 22h44
 Mondialisation et OMC	19 septembre, 16h09



La consultation, par exemple, de celui relatif à l'emploi permet d'accéder aux articles se rapportant à ce thème. Les auteurs des articles sont mentionnés (exemple Agence France Presse), ils sont datés et il est possible d'accéder à ceux précédemment édités.

A l'occasion de l'**évaluation** collective qui suit, le groupe **estime** avoir répondu au questionnaire initial car les données recueillies sont suffisantes.

Il ne reste plus qu'à adopter un plan, une présentation, à citer les références et à **communiquer** le fruit des recherches aux autres groupes de la classe.

EXEMPLE 3

Les classes de BEP Maintenance des Systèmes Mécaniques Automatisés (MSMA) et BEP Productique Mécanique option Usinage (PMU) sensibilisées par une journée d'observation astronomique en planétarium décident, pour leur PPCP, de construire un télescope. Les conseils techniques et les contraintes de fabrication sont prodiguées par un membre d'une société astronomique. Plusieurs problèmes restent à résoudre, notamment ceux liés à l'équilibre général du système, au support du miroir et à la transportabilité de l'ensemble.



Des recherches sur Internet sont donc envisagées pour enrichir et compléter la documentation déjà en leur possession, laquelle se réfère au télescope de type "DOBSON". Après **réflexion**, le groupe **recense les informations** dont il a besoin, à savoir le schéma d'un dispositif stable (équilibre indifférent), supportant un miroir (masse d'environ 15 kg) et orientable dans toutes les directions.

Le *métamoteur COPERNIC 2000* est *retenu* pour *recenser les informations* sur le Web.

Pour *rechercher l'information* sur le Web en français, les jeunes écrivent la requête suivante : "construction d'un télescope". L'expression étant entre guillemets ("..."), cela signifie que seuls seront identifiés les sites dans le descriptif desquels la requête figure exactement comme elle est écrite..



La *consultation* des réponses obtenues au nombre de 4 révèle que la recherche doit être *reconsidérée* car l'une est un carnet d'adresses et les autres sont inintéressantes car la construction du télescope n'est pas traitée.

Estimant ne pas avoir répondu au questionnaire initial, ils *modifient la stratégie de recherche*, utilisent les 2 mots-clés construction et télescope et lancent une nouvelle recherche.

Après un *examen* rapide des descriptifs des 29 sites trouvés ils *retiennent* ceux mentionnant les télescopes de type DOBSON, puisque l'une des consignes de départ faisait référence à ce modèle.

Le premier site visité est intéressant. De nombreuses photos figurent, des liens sont proposés et une rubrique construction apparaît. Toutefois, la consultation renvoie à un site anglophone avec des dimensions en pouces, ce qui incite les élèves à consulter l'autre site précédemment identifié.



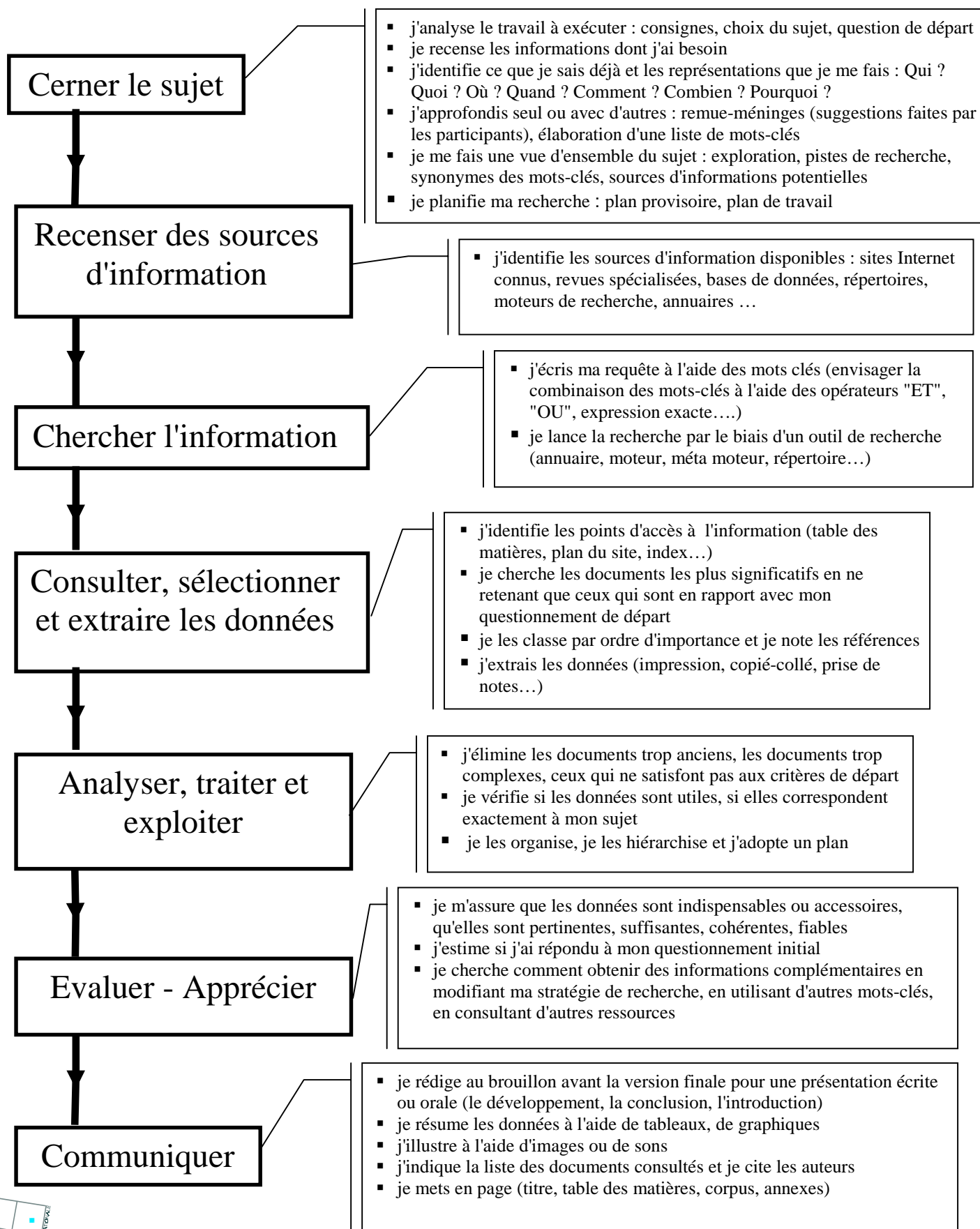
Exemple d'un barillet 18 points en mode mixte, triangles de flottaison et supports astatiques.

Après avoir examiné le plan du site, ils cliquent sur la rubrique *barillet astatique*, obtiennent le plan recherché et *estiment* avoir terminé la recherche.

Pour *présenter* leurs travaux, ils notent les références du site consulté et sélectionnent quelques images qui viendront illustrer leur compte rendu.



CONDUIRE PAR ETAPES UN PROJET DE RECHERCHE



COMMENT TROUVER LE BON SITE ?

Mots-clés

Dans un travail de recherche documentaire, un mot-clé est un mot porteur de sens qui permet d'interroger une base de données. Les mots-clés permettent de synthétiser un thème, un projet, une problématique, une recherche.... Exemple : train-avion-bateau-métro-moto ont pour mot clé "moyen de transport".

Dans le cadre de l'utilisation d'un moteur de recherche, les mots-clés utilisés peuvent conduire à des problèmes de **synonymie** (une demande d'information sur les "automobiles" laissera de côté un grand nombre de pages qui évoquent les "voitures"), **d'homonymie** (une requête sur "grève" vous fera entrer dans l'univers de la revendication, mais aussi dans celui, plus paisible, de la mer ou encore à Paris, sur la Place de Grève...) et de **polysémie** (une simple requête sur "jaguar" donnera des milliers de pages concernant, entre autres, l'automobile de ce nom, le félin, mais aussi une équipe de football américain). Si l'on maîtrise bien son sujet, la stratégie de recherche consiste à aller immédiatement au plus précis pour restreindre le nombre de documents. Sinon il est préférable d'aller du plus simple au plus compliqué en s'assurant, à chaque étape, c'est à dire à chaque rajout d'un mot-clé, de la présence de documents sur le sujet donné.

Améliorer la recherche par mots-clés

Il y a peu de chance qu'un seul terme utilisé comme mot-clé fournisse des résultats appropriés et peu nombreux. Les moteurs de recherche tiennent compte de cet aléa en permettant une recherche combinant les chaînes de caractères à chercher comme les opérateurs "et", "ou"... , les guillemets et les parenthèses. Ainsi, l'utilisation de mots-clés multiples donnera souvent de meilleurs résultats. Par défaut, écrire les mots-clés en minuscules et sans accents.

L'espace () : par défaut, l'espace est considéré comme un séparateur de chaîne, mais aussi, avec certains moteurs, comme l'expression de l'addition de chaînes. Il équivaut alors à [+] ou [ET] ou [AND].

Les guillemets (" ") : le moteur de recherche les interprète comme une chaîne unique et non comme des mots séparés. Les résultats correspondent aux pages qui contiennent la chaîne de caractères exacte.

La troncature (*) : le caractère générique est utilisé pour chercher plusieurs formes possibles d'un mot-clé. En ajoutant un astérisque (*) à la fin d'un mot, comme dans program*, le moteur cherchera le terme "program" ainsi que "programme", et "programmation" ; "programmons"

Les opérateurs booléens ET/OU/SAUF : ils doivent être utilisés en majuscules.

Utilisez ET pour vous assurer que tous les mots-clés apparaissent dans le résultat de la recherche.

Utilisez OU pour vous assurer qu'au moins un des mots-clés apparaîtra dans le résultat de la recherche.

Utilisez SAUF pour vous assurer qu'un certain terme n'apparaît pas dans le résultat de la recherche.

L'opérateur PRES DE est utilisé pour s'assurer que les deux mots-clés sont proches l'un de l'autre, à l'intérieur d'un groupe de 10 mots.

Les parenthèses () peuvent être utilisées pour grouper des mots-clés en utilisant les opérateurs booléens afin d'effectuer une recherche encore plus précise.

Moteurs de recherche

Un moteur de recherche est un logiciel appelé robot qui parcourt en permanence le réseau et archive le contenu des pages qu'il rencontre (environ 3 milliards de pages sont disponibles). Il note les termes significatifs qui les composent : **les mots-clés** et leur attribue un index (actuellement un moteur n'indexe qu'une petite partie, entre 10 et 15 % de la toile mondiale, mais les progrès sont constants). Ensuite, la sélection est inversée et à chaque requête formulée à l'aide d'un mot clé est

associée la liste des pages qui le contiennent. Par exemple, sur le moteur de recherche "Google", le terme de "Mp3" est associé à plus de 10 millions de pages. Ainsi, l'internaute qui définit sa requête à l'aide du mot clé "Mp3" pourra choisir parmi les 10 millions de pages qui lui sont proposées. Bon courage !

Le nombre de réponses est presque toujours très élevé et peu de sites répondent de manière satisfaisante. C'est pourquoi il importe de bien définir sa requête en recourant à plusieurs mots clés et en utilisant les options de recherche avancées, comme les opérateurs "et", "ou"..., les guillemets et les parenthèses.

Le tableau qui figure- ci-dessous recense la mise en œuvre des opérateurs pour les principaux moteurs.

Opérateurs	Moteurs de recherche					
	ALTA VISTA	GOOGLE	HOTBOT	CALOWEB	VOILA	LOKACE
ET	AND ou +	par défaut	par défaut	par défaut	+ ou ET ou AND	par défaut
OU	par défaut	non	OR	OR ou OU	par défaut ou OR ou OU	non
SAUF	AND NOT ou -	-	NOT	-	-	-
PHRASE EXACTE	"..."	"..."	"..."	"..."	"..."	"..."
TRONCATURE A DROITE	*	non	*	non	non	non
ACCENTS	oui	oui	oui	indifférent	non	indifférent
MAJUSCULES	indifférent	indifférent	indifférent	indifférent	indifférent	indifférent
MOTS ORDONNES	oui	oui	indifférent	oui	indifférent	indifférent
Remarque : la plupart des moteurs proposent une recherche avancée qui facilite la formulation de la recherche.						

Remarque : les pages sélectionnées sont listées et numérotées. L'extraction qui figure ci-dessous présente les 10 premières réponses parmi les 1 696 adresses répondant aux mots clés énergies+nouvelles.



APPLICATION :

Dans le cadre d'un PPCP ayant pour thème la qualité de l'eau, il s'avère nécessaire de conduire une recherche sur la production d'eau potable.

Le tableau ci-dessous présente les résultats en français de la recherche réalisée sur tout le Web.

Syntaxe de la requête	Moteur de recherche	Nombre de réponses
eau potable	Caloweb	43 700
eau potable	Google	43 100
eau potable	AltaVista	520 062

Le nombre de réponses affichées aujourd'hui ne sera sans doute pas le même demain car la fréquence de rafraîchissement de l'indexation est variable d'un moteur à l'autre. Néanmoins, on remarque que le nombre de documents trouvés varie d'un moteur à l'autre. En effet, les sites référencés ne sont pas nécessairement les mêmes et le nombre de sites indexés varie. En outre, cela ne préfigure en rien quant à la qualité des réponses proposées.

Syntaxe de la requête	Moteur de recherche	Nombre de réponses
eau potable	AltaVista	= 500 000
production eau potable	AltaVista	= 1 000 000
eau+potable	AltaVista	35 199
production+eau+potable	AltaVista	9

L'importance du nombre de réponses fournies à la première requête invite à reconsidérer celle-ci en augmentant le nombre de mots-clés. Le résultat n'est pas à la hauteur des attentes car le nombre des réponses a sensiblement augmenté, ce qui, au demeurant, n'est pas surprenant car le moteur a identifié le mot production, le mot eau et le mot potable, ces trois mots-clés n'étant pas forcément présents simultanément sur les sites visités.

L'utilisation du connecteur ET ou + contraint cette fois le moteur à ne retenir que les sites présentant simultanément les deux mots clés eau et potable et, de ce fait, le nombre de réponses a fortement diminué. A fortiori, si l'on impose un troisième mot-clé, le nombre de réponses diminue considérablement.

Cet exemple permet d'appréhender l'importance et le soin qu'il faut accorder à la formulation de la requête et à l'utilisation des connecteurs logiques.

La syntaxe à respecter (mettre un espace ou non avant ou après les connecteurs) est variable d'un moteur à l'autre. Certains connecteurs (ET par exemple) sont utilisés par défaut par quelques moteurs. Les fonctionnalités de la recherche avancée (quand elle existe), si elles permettent d'utiliser plus simplement les connecteurs, diffèrent également d'un moteur à l'autre. Il semble donc préférable et plus efficace de choisir un moteur de recherche et de bien le connaître plutôt que de butiner de l'un à l'autre.

Les moteurs de recherche francophones : AltaVista France ([http:// fr.altavista.com/](http://fr.altavista.com/)), Caloweb (<http://www.caloweb.com/>), Ecila (<http://www.ecila.fr/>), Excite , France (<http://fr.excite.com/>), HotBot France (<http://www.hotbot.lycos.fr/>), Google français ([http:// www.google.com/intl/fr/](http://www.google.com/intl/fr/)), Lokace (<http://www.lokace.com/>), Netscape Recherche (<http://home.netscape.com/fr/escapes/search/ntschrnd-1.html>), Voila ([http:// www.voila.fr/](http://www.voila.fr/)), Perso-search ! ([http:// www.perso-search.com/](http://www.perso-search.com/)).

Méta moteurs

Les méta moteurs interrogent en une fois différents outils de recherche. Compte tenu que chaque moteur a sa propre base de données et qu'aucun n'est exhaustif, la réponse obtenue est plus efficace car elle correspond à la compilation des réponses obtenues par chacun des moteurs.

APPLICATION :

COPERNIC (<http://www.copernic.com>) propose 27 sites correspondant à la requête "production d'eau potable". Ce méta moteur présente l'avantage de conserver en mémoire non seulement les résultats de la recherche, mais aussi le contenu des pages dont l'utilisateur aura demandé le téléchargement. De plus, il affiche un indice de pertinence pour chaque résultat trouvé.

Annuaire

Un annuaire est un moteur de recherche indexé, non par des robots, mais à la main par des opérateurs qui répertorient les sites catégorie par catégorie. On comprend dès lors que les listes de réponses se réfèrent à un nombre de sites limité (exemple 120 000 sur l'annuaire Yahoo) car le nombre de pages croît au rythme d'un million par jour.

Le classement des informations par rubrique est plus structurant pour les élèves et à chaque adresse est associée une description de quelques lignes ce qui évite les risques d'égarement sur des sites inappropriés.

Les annuaires présentent l'avantage de trouver les pages les plus fraîches dont les mises à jour sont fréquentes, le dernier reportage vidéo ou la dernière promotion. Peu importe la météo d'hier, c'est celle d'aujourd'hui qui compte. Les sites de recherche associent fréquemment moteurs et annuaires.

APPLICATION :

La recherche sur la production d'eau potable est, cette fois, réalisée avec l'annuaire Yahoo.

1^{ère} possibilité : réaliser une recherche thématique.

Après avoir parcouru l'arborescence : Commerce et économie → Produits et services pour les professionnels → Environnement → Qualité et traitement de l'eau → Bureaux d'études et ingénierie, on obtient la liste des sites correspondants :

- **Saur**- Production et distribution d'eau potable, assainissement et épuration, conception et construction d'ouvrages de traitements d'eau.
- **Rhodia Eco Services**- Traitement des déchets (industriels, sols), produits et technologies de traitement des eaux (eau potable, eaux usées, effluents). Production et régénération d'acide sulfurique et dérivés.
- **Stereau**- Filiale du Groupe Saur spécialisée dans la conception et la construction d'usines de production d'eau potable et de stations de traitement des eaux usées.

D'autres arborescences peuvent être explorées :

- Sciences et technologies → Sciences de la Terre → Hydrologie → Instituts
- Commerce et économie → Produits et services pour les professionnels → Environnement → Gestion des déchets → Déchets industriels → Recyclage
- Commerce et économie → Produits et services pour les professionnels → Environnement → Qualité et traitement de l'eau → Équipement et matériel → Fabricants

2^{ème} possibilité : utiliser le moteur de recherche intégré à l'annuaire Yahoo.

Syntaxe de la requête	Annuaire	Nombre de réponses
eau+potable	Yahoo	29
production+eau+potable	Yahoo	6

L'indexation étant manuelle, le nombre de sites identifiés est bien évidemment plus restreint. Les recherches sont moins exhaustives mais elles sont souvent plus rapides et plus fructueuses.

Remarque : il est possible de combiner l'arborescence et le moteur de façon à n'obtenir que les réponses catégorielles liées, par exemple, à l'environnement ou au commerce ou...

Les Annuaire francophones

C'est Trouvè (<http://www.ctrouve.com/>),

Francité (<http://www.francite.com/>),

Club-internet (ex Hachette net) (<http://www.hachette.net/>),

INDEXA (<http://www.indexa.fr/>),

Lycos France (<http://www.lycos.fr/>),

Nomade (<http://www.nomade.fr/>),

Le Guide de Voila (<http://themes-search.voila.fr/>),

Souany (<http://www.souany.com/>),

Sympatico (<http://pre.sympatico.ca/fr/chercher/recherche.html>),

Yahoo! France (<http://www.yahoo.fr/>).

Portails

Ce sont des points d'entrée de l'internaute sur le web, ils sont constitués d'un annuaire et d'un moteur de recherche. Un grand nombre de services sont offerts (messagerie, dépêches d'agence de presse, sélection de sites, météorologie, cours de la bourse shopping...). En classe, ils présentent un intérêt pédagogique si l'on a besoin fréquemment du même type d'information (relevé météo, actualité...).

Sites portail en français : <http://www.eureka-fr.com/> ; <http://www.excite.fr/> ; <http://www.francite.com/>.