La Recherche sur l'Internet Cours

Michel Cabaré Janvier 2002 ver 2.2

TABLE DES MATIERES

TABLE DES MATIERES	2
INTERNET EXPLORER 5.X.	6
LANCEMENT DE INTERNET EXPLORER :	
NAVIGATION DE BASE :	
IMPRIMER UNE PAGE :	
GESTION D'HISTORIQUE :	
PAGE DE DEMARRAGE :	
GESTION DU CACHE :	
GERER LES ADRESSES VIA "LES FAVORIS"	
DES ADRESSES PEU "PARLANTES" :	
AJOUTER UN FAVORIS:	
Utiliser un Favoris :	
ORGANISER LES FAVORIS:	
Créer un dossier	
Déplacer un dossier	
Sauvegarde des Favoris :	
ANNUAIRES & MOTEURS	16
LES ANNUAIRES:	16
LES MOTEUR DE RECHERCHE	17
REFERENCEMENT D'UN SITE	19
SE REFERENCER SUR DES ANNUAIRES:	19
Nomade (www.nomade.tiscali.fr):	
Yahoo (www.yahoo.fr):	22
SE REFERENCER SUR DES MOTEURS:	
Voilà (www.voila.fr):	
Altavista (www.altavista.fr):	24
Google (www.google.fr):	
TAGS META REFERENCEMENT.	27
Revisit After	
Title	
Keywords	
Description	
PRINCIPE GENERAL DE FONCTIONNEMENT D'UN MOTEUR	
Le titre de la page que vous avez saisi dans la balise d'entête	
La description que vous avez saisi dans la balise d'entête	
Les mots clés que vous avez saisi dans la balise d'entête.	
LES PROBLEMES DE FONCTIONNEMENT D'UN MOTEUR	30

Les images	30
Les formulaires	36
Les bases de données	31
Les images mappées	
Les frames	
CRITERE DE CLASSEMENT DES SITES	
POUR ETRE BIEN PLACE	
Créez des liens explicites	
Privilégiez le texte	
Déposez votre nom de domaine	
Divisez un site en plusieurs services	
Renouvelez régulièrement les pages	
Référencez vous régulièrement	
LES OUTILS DE SOUMISSION AUTOMATIQUE	33
Presentation	33
REFERENCEURS ET MOTEURS	33
REFERENCEURS ET ANNUAIRES	
Conclusions	
RECHERCHE PAR MOTEUR OU ANNUAIRE	34
ON NE PEUT TROUVER QUE DES SITES REFERENCES :	
UTILISER DES ANNUAIRES	
UTILISER DES MOTEURS DE RECHERCHE	
VERIFICATION DU REFERENCEMENT	
Faites des échanges de liens	
Ne supprimez pas (trop) de pages	
Vérifiez régulièrement votre référencement	
YAHOO (1° ANNUAIRE)	
PRESENTATION	
RECHERCHE PAR CATEGORIE	
RECHERCHE PAR MOTS CLES	
Nombre de mots cles a employer :	
SENSIBILITE A LA CASSE :	
SENSIBILITE AUX ACCENTS :	
GESTION DES PLURIELS :	
RECHERCHE SUR UNE EXPRESSION PRECISE	
Mots exiges et mots interdits	
RESTRICTIONS A UN CORPS DE DOCUMENT	
CORRESPONDANCE ELARGIE (*)	
COMBINER LA SYNTAXE	
PERTINENCE DU CLASSEMENT	43
RECHERCHE INFRUCTUEUSE	44
LES OPTIONS DE RECHERCHE	45
YAHOO! DANS LE MONDE	45
ALTAVISTA (1° MOTEUR)	46
Presentation	46
RECHERCHE SIMPLE	
La casse	47
	47 47
La casse	47 47 47

Troncation	48
L'opérateur OU	48
L'opérateur ET	
L'opérateur SAUF	
L'opérateur PROCHE DE	
RECHERCHE AVANCEE	
Les operateurs	
Logique Booleenne Intervalle de dates	
Pertinence du classement	
GOOGLE (2° MOTEUR)	
Presentation	
RECHERCHE AVANCEE	
Une vision annuaire ?	
VOILA (MOTEUR+ANNUAIRE)	
Presentation	
La casse	
Les accents	
Ordre des mots	
Recherche de phrases	
L'opérateur OU	
L'opérateur ET	58
L'opérateur SAUF	58
L'opérateur PROCHE DE	
Recherche sur le type de fichiers	
LA RECHERCHE AVANCEE	
PERTINENCE DU CLASSEMENT.	
PAGES JAUNES/BLANCHES/EMAIL	
PLANS ET ITINERAIRES	
RECAPITULATIF - MOTEURS DE RECHERCHE	
CAPACITE DES MOTEURS DE RECHERCHE A ACCEPTER DES CRITERES DE RECHERCHE ETENDU	
CRITERES DE CLASSEMENT DES RESULTATS OBTENUS.	66
METAMOTEURS DE RECHERCHE	68
Un metamoteur, ce n'est pas mieux qu'un moteur ?	68
LES METAMOTEURS DE RECHERCHE.	
LES METACHERCHEURS, TABLEAU COMPARE	
LES AGENTS INTELLIGENTS ET ASPIRATEURS	
AGENTS INTELLIGENTS	71
L'EXEMPLE DE COPERNIC	
par exemple Copernic :	
LES ASPIRATEURS	
ASPIRATEURS DE SITES	
L'EXEMPLE DE MEMO WEB	
RECHERCHE MULTIMEDIA	
PRINCIPAUX MOTEURS DE RECHERCHE MULTIMEDIA	
Photos – images - cliparts	
Cartes – plans - itineraires	73

RECHERCHE ADRESSES MAIL	74
Adresses electroniques & N° de telephone	74
LISTE DE DIFFUSION - FORUMS	75
LISTE DE DIFFUSION	75
Principe de base	
Cas particulier	
Trouver une liste	
EXISTE-T-IL DES LISTES DE DIFFUSION PARLANT DE RECHERCHE?	
FORUM	
EXISTE-T-IL DES FORUMS DE DISCUSSION PARLANT DE RECHERCHE?	
RECHERCHE DE LOGICIELS	78
Telechargement	78
Exemple sur zdnet.fr du logiciel gnutella	
LES FICHIERS SUR SERVEURS FTP	80
FTP FIND.	80
File Mine	
Filez	
Shareware.com	
TuCows	
Exemple de recherche : recherche de winzip	

INTERNET EXPLORER 5.X...

Les pages qui suivent ne se veulent pas un manuel d'utilisation de **Internet Explorer** mais plutôt un guide pour se repérer dans **la navigation** sur le Web à l'aide d'un navigateur...

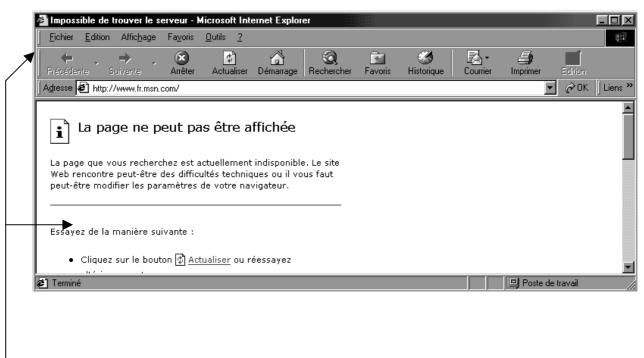
Lancement de Internet Explorer :

L'icône de la barre des tâches ou sur le bureau est facilement reconnaissable...

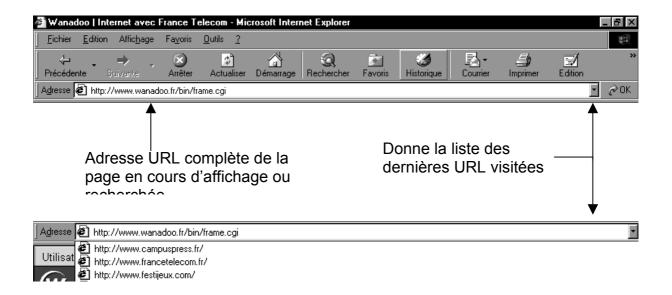




Composition de l'écran:



N.B: Ne confondez pas le navigateur, avec la page qui s'affiche dedans (ici volontairement pas de page, donc un message d'erreur...)



Les outils les plus importants sont

← Précédente	Retourne sur la page affichée précédemment
⇒ Suivante	Retourne sur la page affichée (lorsque l'on est revenu en arrière)
Arrêter	Arrête le chargement de la page
Actualiser	Force le re-chargement de la page
Démarrage	Retourne sur l'adresse configurée comme page par défaut
Rechercher	Lance une recherche via le site prédéfini MSN
* Favoris	Gestion des adresses que l'on à stocké dans les "favoris"
Historique	Gestion des adresse visitées lors des dernières navigations
Courrier	Permet de lancer le client de messagerie (Outlook Express)
Imprimer	Imprime la page actuelle à l'écran

Navigation de Base:

Il faut forcément taper la première fois une URL dans la barre d'adresse sous la forme :

http://www.nom.domaine

par exemple

http://www.sncf.fr

http://www.fnac.com

N.B: Une URL est généralement composée (voir chapitre les adresses internet)

- de caractères tous en minuscule
- aucun caractère spécial (accent, ponctuation)
- pas d'espace

En cas d'erreur sur l'URL on n'aura jamais une réponse catégorique de la part du navigateur (du style "cette adresse n'existe pas") mais plutôt une réponse du style "je n'arrive pas à trouver..."



La page ne peut pas être affichée

La page que vous recherchez est actuellement indisponible. Le site Web rencontre peut-être des difficultés techniques ou il vous faut peut-être modifier les paramètres de votre navigateur.

AVANT DE CHERCHER UN MALFONCTIONNEMENT TECHNIQUE IL FAUT VERIFIER I'URL!

puis naviguer a l'aide des 4 boutons

↓ Précédente	Retourne sur la page affichée précédemment
⇒ Suivante	Retourne sur la page affichée (lorsque l'on est revenu en arrrière)
Arrêter	Arrête le chargement de la page (si trop long ou si on à l'impression que cela ne marche plus)
Actualiser	Force le re-chargement de la page (évite de reprendre la même page éventuellement mise en cache sur la machine lors de la dernière visite)

Imprimer une page:

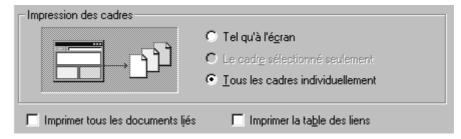
Le principe est simple, il suffit de demander :



Néanmoins cela peut très vite se compliquer par l'existence de cadres constituant le page affichée

Il peut être nécessaire lorsque l'impression ne donne qu'une partie de ce que l'on a à l'écran de demander le menu

Fichier / Imprimer



dans lequel notamment on peut définir ce que l'on souhaite imprimer...



Gestion d'Historique:

Automatiquement Internet Explorer mémorise les URL différentes dans lesquelles on se promène dans un historique :

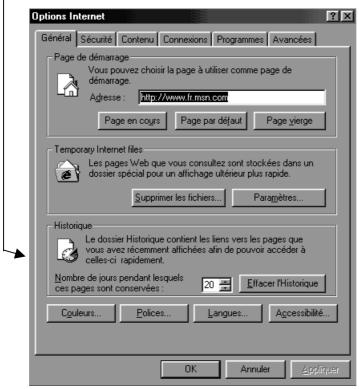


Ce qui permet de retrouver facilement des sites sur lesquelles on vient de passer...

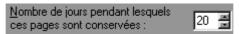
Cela marche relativement bien, mais bien sûr cet historique automatique peut devenir très très imposant ... (génant ?)

On peut "gérer" cet historique en allant dans le menu

Outils / Options Internet onglet Général



et modifier le nombre de jour d'activation de l'historique



ou vider l'historique

Effacer l'Historique

Page de démarrage :

Il s'agit de choisir la page sur laquelle vous souhaitez que votre navigateur s'ouvre lors de son lancement. C'est aussi la page que vous pouvez afficher à tout moment via :

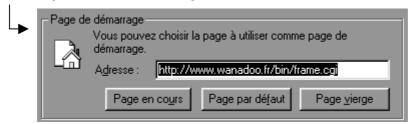


Par défaut cette page est celle de votre fournisseur d'accès ou distributeur de navigateur...

C'est évidemment pas toujours génial

On peut "gérer" cette page de démarrage en allant dans le menu

Outils / Options Internet onglet Général



Pour donner une nouvelle page de démarrage, il suffit d'afficher la page dans le navigateur, puis de demander

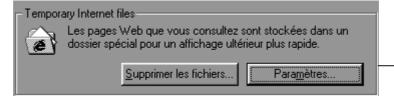
Page en co<u>u</u>rs

Gestion du Cache:

La navigation stocke une grande quantité de fichiers entraînant parfois l'encombrement et la fragmentation du disque

On peut à ce titre gérer le cache en allant dans le menu

Outils / Options Internet onglet Général



et demander de vider le dossier des fichiers temporaires et de gérer le cache internet sur le disque dur



GERER LES ADRESSES VIA "LES FAVORIS"

Des adresses peu "parlantes" :

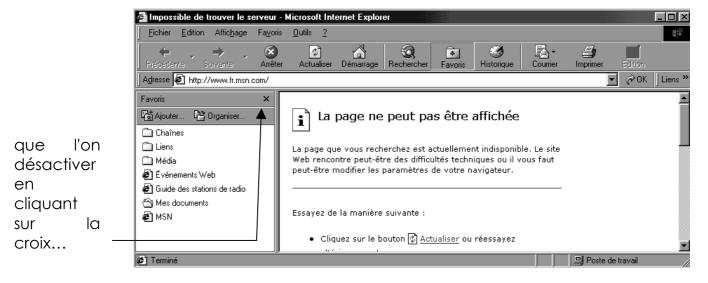
Un des gros problèmes pour un utilisateur d'internet c'est de gérer convenablement la liste d'adresses qu'il utilise, en sachant que ces adresses sont peu mémorisables et se ressemblent toutes

Le principe va consister à leur affecter un nom et à les stoker dans des structures similaires au dossiers-répertoires sur votre ordinateur

cliquer sur



amène:



Le menu Favoris permet de gérer également les favoris

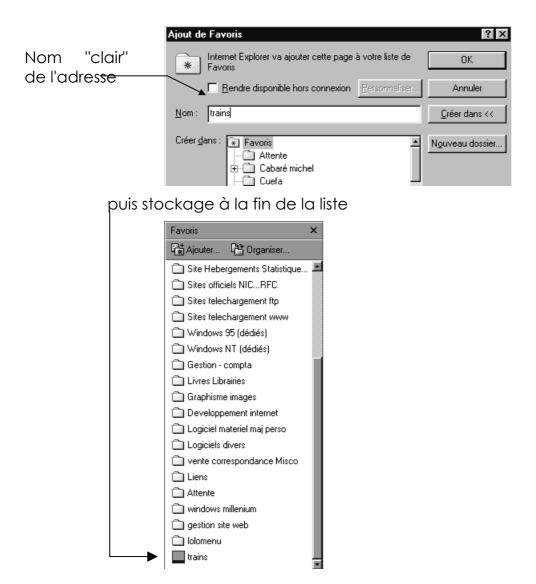


Ajouter un Favoris:

Imaginons que nous souhaitions enregistrer cette adresse http://www.sncf.fr sous un libellé plus parlant comme "trains"

après s'y être rendu, il suffit de demander le Menu

Favoris/ Ajouter aux favoris...



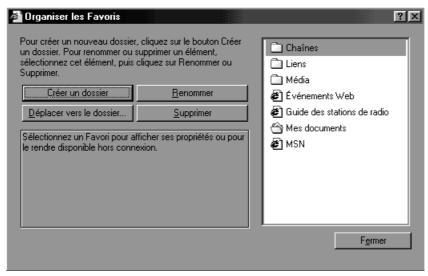
Utiliser un Favoris:

- Si les favoris apparaissent sur la gauche du navigateur il suffit de cliquer sur l'entrée que l'on vient d'ajouter
- Si les favoris ne sont pas à l'écran le menu Favoris doit faire apparaître l'entrée

Organiser les Favoris:

La gestion des adresses existantes dans une arborescence plus complète peut se faire via le menu

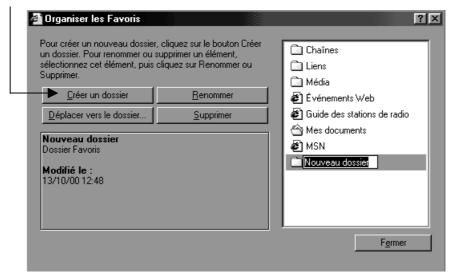
Favoris/ Organiser les favoris...



Créer un dossier

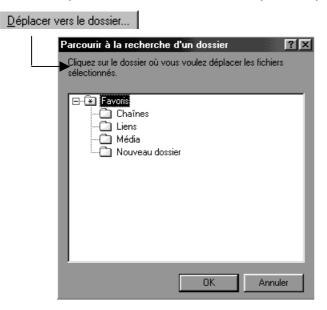
De manière similaire à la création d'un dossier dans l'explorateur windows...

On demande Créer un dossier



Déplacer un dossier

Il faut se placer sur le dossier à déplacer puis demander

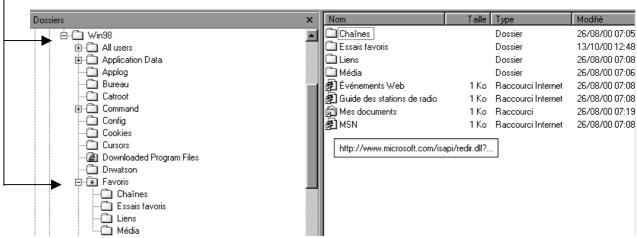


ou il suffit de se placer dessus puis de le faire glisser comme dans l'explorateur windows

Sauvegarde des Favoris :

Notez que toutes les manipulations ici dépendent de l'environnement système, et ne sont valables que pour des machines fonctionnant sous Windows 9x...

Les favoris sont stockés dans un sous dossier de Windows nommé Favoris



ANNUAIRES & MOTEURS

D'apparence similaire, ce sont tous les deux des sites Web dotés d'une fenêtre de saisie destinée à recueillir l'objet de votre requête. Les deux outils se distinguent cependant par leur façon de recenser et d'organiser l'imposante somme d'informations disponible sur le Web.

Les Annuaires:

Un annuaire est un outil de recherche qui recense un certain nombre de sites au travers de fiches descriptives comprenant, en règle générale :

- le titre,
- l'adresse (l'URL)
- et un bref descriptif d'une longueur allant le plus souvent de 15 à 25 mots au maximum.

Chaque site est inscrit dans une ou plusieurs **catégorie(s)** - on parle également de **rubrique(s)** -. Ces catégories thématiques (finances, sports, tourisme, etc.) vont du plus général au plus précis.

Ces outils peuvent ainsi être considérés comme les pages jaunes du Web.

Par exemple, la rubrique principale "Santé" de Yahoo donne naissance à une trentaine de sous-catégories ("Assurances", "Hôpitaux", "Magazines", etc.) qui elles-mêmes se divisent en plusieurs autres parties et ainsi de suite.

Lorsqu'un mot-clé est saisi dans le formulaire proposé, l'annuaire effectue une recherche sur les occurrences de ce terme dans ses fiches descriptives de site, et non pas dans le contenu des pages du site en question. Il s'agit là de la différence la plus notable avec les moteurs de recherche.

"Référencer un site" correspond à la démarche que doit réaliser un concepteur de site pour inscrire un site dans un annuaire (nous verrons plus tard dans un moteur). Le référencement est donc un acte volontaire.

Comprendre le référencement d'un site dans un annuaire permet de comprendre comment rechercher dans cet annuaire.

Les annuaires référencent des sites et plus particulièrement la page d'accueil (Home Page qui souvent s'appelle **index.htm**). Le classement est généralement organisé en catégories et sous catégories. C'est le webmaster qui référence son site en choisissant la catégorie dans laquelle il désire le voir apparaître. Sur le site de l'annuaire, une intervention humaine permettra de vérifier ce choix ainsi que (très rapidement) le contenu.

Les Moteur de Recherche

Le moteur de recherche fonctionne sur un système radicalement différent de celui de l'annuaire. Des **robots** logiciels (appelés crawlers ou spiders) scrutent le Web, vont de page en page (en fait de lien en lien) et sauvegardent au fur et à mesure de leurs pérégrinations le contenu texte des pages rencontrées, constituant ainsi un "index", c'est-à-dire une collection plus ou moins grande de pages Web. Pour exemple, AltaVista en stocke 275 millions. HotBot 150 millions. WebCrawler 2 millions

Le **robot** logiciel **repasse** selon des **délais plus ou moins fréquents** sur les pages qu'il a indexées au préalable, pour en sauvegarder une version plus récente. On dit alors qu'il "rafraîchit sa base (ou son index)".

Lorsque l'internaute saisit un mot clé dans le formulaire proposé, le moteur va en rechercher les occurrences dans son index, c'est-à-dire dans le contenu texte des pages Web sauvegardées au préalable. Une fois le "lot" de pages contenant le terme demandé identifié, le moteur classe les pages par ordre de pertinence, selon un ordre et un algorithme (basé sur certains critères de tri) qui lui est spécifique.

Les robots indexeurs agissent donc exactement comme des surfeurs du Web : ils parcourent les pages en notant précieusement chacun des mots contenus et en visitant tour à tour chacune des URLS présentes

Lorsque les robots indexent un site, ils utilisent en priorité les balises META (voir ce chapitre) pour construire le résumé qui apparaît lors d'une recherche sur ce site. Les META tags sont des balises HTML que l'on insère dans les documents et qui indiquent aux robots comment la page doit être indexée (prise en compte du titre, des mots-clés, de la description...).

L'indexation est réalisée en plein texte, ce qui signifie que les robots considèrent tous les mots d'une page web. Autrement dit, lors d'une recherche sur un mot-clé, les robots répondent en fonction de la note qu'ils accordent à chacun des mots du site : cette note tient compte non seulement des balises META, mais aussi de la position des mots dans la page, de leur fréquence, de leur mise en forme (taille, gras, italique...). Mais les titres des pages, les texte alternatifs des images, sont aussi pris en compte par les robots

Attention, parfois les robots n'aiment pas les "ruses" telles que les écritures de même couleur que le fond (de manière à ce qu'une page contienne un nombre élevé d'occurrence d'un mot clé.... sans que cela nuise à la visibilité du site) et considère ceci comme du spamming ! (sorte de publicité abusive) Du coup la page n'est plus du tout référencée ...

Le moteur de recherche effectue donc ses **recherches sur des pages Web**, alors que l'annuaire vous proposera des **sites Web**. Là est toute la différence qui explique qu'il est absolument impossible de comparer les résultats fournis par les deux outils.

La Recherche Sur Internet

REFERENCEMENT D'UN SITE

Se référencer sur des Annuaires:

2 annuaires principaux ont étés choisis pour illustration

Nomade (www.nomade.tiscali.fr):

L'objectif de Nomade-Tiscali est de répertorier les sites Internet en Français indépendamment de la nationalité de leurs éditeurs ou de la localisation géographique des serveurs d'hébergement.

Nomade est un guide par soumission (annuaire). Ce sont les éditeurs de services qui fournissent une description de leur site. Cette description est vérifiée par l'équipe de Nomade avant d'être mise en ligne

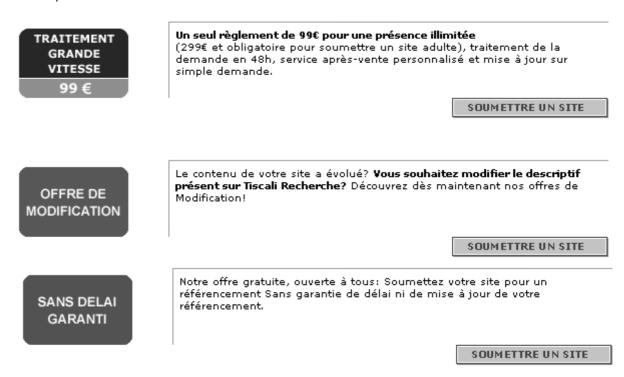


Dans un premier temps le créateur du site va parcourir les catégories et sous catégories afin de trouver celle dans laquelle il désirera faire figurer son site.

Le webmaster trouvera dans la page affichée un lien lui permettant d'ajouter son site à la catégorie dans laquelle il se trouve.



Tiscali innove avec différents style de référencement, mais il s'agit la d'une exception....



A partir de la l'enregistrement n'est qu'une formalité

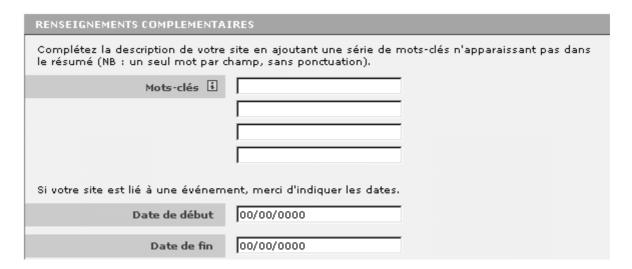


En cas de problème technique, merci de contacter inforef-nomade@fr.tiscali.com

* = champ obligatoire







Yahoo (www.yahoo.fr):

Yahoo est un guide par soumission (annuaire). Ce sont les éditeurs de services qui fournissent une description de leur site. Cette description est vérifiée par l'équipe de Yahoo avant d'être mise en ligne



Dans un premier temps le créateur du site se déplacera dans l'arborescence des catégories afin de trouver la catégorie correspondant au site qu'il désire référencer.

Le webmaster trouvera dans la page affichée un lien lui permettant d'ajouter son site à la catégorie dans laquelle il se trouve.

Copyright © 2002 Yahoo! Inc. Tous droits réservés.

<u>Tout savoir sur Yahoo! - Yahoo! Données personnelles - Conditions d'utilisation - Proposer un site - Aide</u>

Se référencer sur des Moteurs:

3 moteurs principaux ont étés choisis pour illustration

Voilà (www.voila.fr):

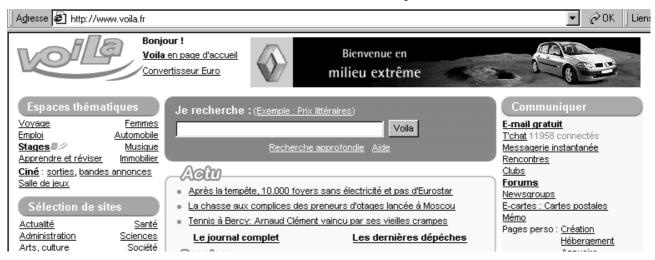
Lancé en juillet 98, Voila est la suite du moteur Echo racheté par France Telecom. Sous forme de portail **Voilà** regroupe un moteur de recherche sur le web francophone, un annuaire intégré (Guide), les actualités, la météo et des informations boursières. Les annuaires <u>Pages Jaunes</u>, Pages Blanches, Les marques, Les rues commerçantes et Les pages pro complètent l'offre pour la France.

Le service s'internationalise avec l'ouverture de <u>Voila.com</u> qui propose une recherche mondiale et pour 5 pays en plus de la France :

The Whole Web - Denmark - Italy - Netherlands - Portugal - Spain

Voila recherche des documents dans une base de données de plus de 11 000 000 de pages web en langue française mise à jour chaque semaine.

Les sites inscrits manuellement sont indexés sous 15 jours.



Il existe ici aussi des offre payantes...

Nos offres de Soumission Express

Soumission Express Basic

Conditions générales de vente.

- Soumission d'un site dans le Guide du Web Voila & Wanadoo
- Garantie d'une évaluation de votre site sous 7 jours ouvrés **
- Service privilégié : Nos équipes s'engagent à vous informer personnellement à chaque étape du traitement de votre demande

199 Euros HT soit 238 Euros TTC Paiement en ligne sécurisé par carte bancaire



mais les standards sont "maintenus:

Notre offre de Soumission Standard

L'offre de Soumission Standard permet de soumettre son site gratuitement dans le Guide du Web Voila et Wanadoo, sans garantie de délais ni de réponse.

Pour accéder à cette offre, choisissez une catégorie dans le <u>Guide</u>. Une fois cette catégorie choisie, vous trouverez un lien intitulé "Soumission standard" vous donnant accès au formulaire de soumission."

Sur le bas de la page d'accueil du site, on demande Ajouter votre site

⊚ Wanadoo

Ajoutez votre site | Aide | Ecrivez-nous | Régie publicitaire | Charte vie privée | Sites à découvrir

wanadoo

qui amène la boite de dialogue suivante:

Référencement	
Référencer vot	re site
	dans la base de données du moteur de recherche Voila , il vous suffit de nous le formulaire ci-dessous.
URL:	http://
Votre E-Mail	:
	Soumettre

Altavista (www.altavista.fr):



Altavista aussi pratique différentes politiques de référencement, mais de manière générale, on peut constater que les offres gratuites sont maintenues...

Trusted Feed ***

- Réservé aux partenaires qui soumettent plus de 500 pages à la base de données d'Altavista.
- Permet la soumission de titres, de mots clés et de signes abstraits personnalisés, et fournit un trafic hautement qualifié grâce à des mises à jour hebdomadaires.
- · Les pages Web sont soumises au format XML.

Inclusion Express **

- Ajoutez rapidement jusqu'à 500 pages à la base de données globale qui alimente les résultats des moteurs de recherche d'AltaVista.
- Générez rapidement et simplement un trafic hautement qualifié.
- Maintenez le contenu des pages ajourne' avec le refrechement quotidien automatique (jour de la semaine seulement) de votre pages.
- Utilisez l'outil de compte rendu en ligne pour gérer la soumission de vos URL.
- Nouveauté! Augmentez la visibilité de vos sites avec des icônes, des logos et plus encore! Pour en savoir plus...

Basic Submit 🤺



- Service GRATUIT.
- Ajoutez ou supprimez de la base de données globale d'Altavista jusqu'à cing pages à la fois.
- Vos soumissions sont évaluées dans un délai de guatre à six semaines.

Pour en savoir plus...

Les sites inscrits manuellement sont indexés sous 4 à 6 semaines

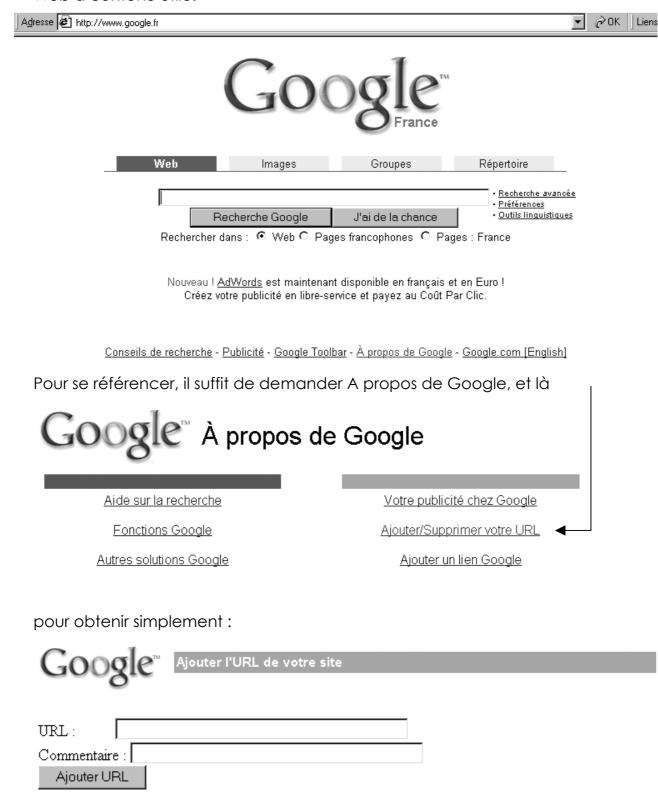
Fonctions de soumission de base

Vous pouvez suggérer l'ajout ou la suppression d'un site parmi les résultats de la recherche d'AltaVista, et ceci gratuitement.

- Vous pouvez soumettre jusqu'à 5 URLs par session.
- Les demandes d'inclusion de nouveaux sites dans les résultats de recherche d'AltaVista sont traitées dans les 4 à 6 semaines.
- Ce service ne garantit pas l'acceptation de votre site.
- D'autre part, il ne garantit pas l'actualisation régulière de votre site.

Google (www.google.fr):

L'index de Google, qui porte sur près de deux milliards d'adresses URL, est le premier du genre et il constitue la collection la plus complète de pages Web à contenu utile.



TAGS META référencement.

Revisit After

Syntaxe: <META NAME="Revisit-after" CONTENT="nombre">

Compatibilité: Certains moteurs

Version: HTML 2.0

Fonctions:

- Permet de spécifier au spider (robot du moteur) de réindexer votre site suivant l'intervalle en jours donnée.

Title

Syntaxe: <TITLE>Titre de votre site</TITLE>

Nombre maxi de caractères: 100

Compatibilité: Tous moteurs

Version: HTML 2.0

Fonctions:

- Permet d'afficher le titre votre site dans la barre supérieure de votre navigateur.
- Lorsqu'un visiteur "bookmark" votre site c'est ce titre qui sera pris en compte comme intitulé de vos favoris ou signets.
- Intitulé pris en compte pas le moteur de recherche comme titre de votre site (première information vue par l'internaute recherchant un site sur un moteur ou annuaire).

Keywords

Syntaxe: <META NAME="Keywords" CONTENT="Vos mots clés ici">

Nombre maxi de caractères : 1000 Maxi

Compatibilité: Tous moteurs

Version: HTML 2.0

Fonction:

- Indique aux moteurs de recherche les mots clés en rapport avec votre site.

L'Erreur à ne pas commettre : Ne répétez pas plusieurs fois de suite le même mot clé.

Note: Cette balise possède l'attribut LANG qui peut définir la langue utilisé pour votre site (fr pour française, en pour Anglaise, en-us pour Américaine, de pour Allemande, it pour Italienne...)

Description

Syntaxe: <META NAME="Description" CONTENT="Description de site...">

Nombre maxi de caractères : 200 Maxi

Compatibilité: Tous moteurs

Version: HTML 2.0

Fonctions:

- Phrase prise en compte pas le moteur de recherche comme description de votre site.
- Les mots que constituent cette phrase sont aussi considérés comme des mots clés par la plupart des moteurs.

Note : Cette balise possède l'attribut LANG qui peut définir la langue utilisé pour votre site.

(fr pour française, en pour Anglaise, en-us pour Américaine, de pour Allemande, it pour Italienne...).

La Recherche Sur Internet

Principe général de fonctionnement d'un moteur

Un moteur de recherche ira (souvent pas toujours) chercher sur les pages de votre site:

Le titre de la page que vous avez saisi dans la balise d'entête.

<Title> Titre de la page</Title>

Ce titre devra être :

- Court
- Significatif
- Un titre par page

Ce titre apparaîtra dans la liste des résultats de la recherche mais aussi dans la fenêtre du navigateur (si vous affichez la page).

La description que vous avez saisi dans la balise d'entête.

<Meta name="Description" contents="texte de la description">

- ▶ Ce texte apparaît dans la liste des sites trouvés à la suite d'une recherche. Il permet de présenter en quelques mots le contenu de la page alors soyez précis et incitateur.
- Le texte devra se composer d'une cinquantaine de mots (si plus, risque d'être tronqué).
- ▶ Pensez à le faire pour chaque page.

Les mots clés que vous avez saisi dans la balise d'entête.

<meta name="KeyWords" contents="liste des mots clés">

- > Ces mots clés devront être Généraux pour quelques un afin de cerner le sujet traité : exemple votre site traite du langage informatique java il vaut mieux faire apparaître quelques termes généraux tels que informatique, web, programmation, langage afin de se dissocier de la danse.
 - Bien sûr ces mots ne seront pas suffisant, une recherche sur informatique va générer une liste de réponses de plusieurs milliers de sites (pourquoi seriez vous dans les premiers ???).
- Ces mots clés devront être Significatifs et précis pour la plupart : la personne qui fait une recherche doit sûrement avoir un besoin précis et peut être sera-t-elle capable de le formaliser dans sa recherche
- N'oubliez pas de donner une localisation précise de votre lieu d'activité : une personne qui recherche un stage informatique le recherchera plutôt dans sa région, dans sa ville.
- Pensez à écrire sans accent
- Pensez à écrire en minuscules
- Pensez à écrire le pluriel et singulier lorsqu'ils sont différents (ex; cheval, chevaux)
- Vous pouvez concentrer essentiellement ces mots sur la page d'accueil : conséquence, vous orientez vos visiteurs vers le point

d'entrée de votre site à eux d'être capable de retrouver la page qui correspond à leur recherche.

Avantage: vous faites connaître à vos visiteurs les autres ressources que vous désirez mettre à leur disposition.

Inconvénient : vous risquez de perdre un certain nombre d'entre eux, des internautes débutants auront peut être un peu de mal à trouver la page qui contient les informations désirées.

Vous pouvez concentrer ces mots sur les pages concernées

Avantage : les visiteurs tombent sur la page qui les intéressent

Inconvénient : s'ils ne sont pas curieux, ils n'iront pas consulter le reste du site.

- On peut inscrire une cinquantaine de mots clés séparés par des virgules
- Le contenu des pages. En général, un moteur parcourt toutes les pages d'un site et en extrait les mots significatifs du contenu.
 La plupart des moteurs de recherche ne fondent pas leur recherche sur les seuls mots clés.

Pour certains moteurs les premières lignes des pages sont celles qui sont prises pour être indexées.

Les problèmes de fonctionnement d'un moteur

Deux types de problèmes peuvent se poser :

- Impossibilité d'indexer des données (cas des images, formulaire, bdd)
- Impossibilité de suivre les liens (image mappée, et frames)

Les images

La plupart des moteurs de recherche savent indexer le texte mais pas les images. Il faut donc les aider un peu en complétant dans la balise HTML d'insertion d'image le paramètre Alt



Là aussi, apportez un soin tout particulier à ce texte

Les formulaires

Les moteurs de recherche n'indexent pas le contenu des formulaires.

Les bases de données

Beaucoup de sites professionnels ont des pages générées à partir de bases de données interrogées par le biais de formulaires? Le contenu des bases n'est pas indexé puisque les formulaires ne sont pas traités.

Les images mappées

La plupart des moteurs de recherche ne savent pas suivre les liens sur les images mappées. Pensez donc à insérer en parallèle un système de navigation sur du texte lorsque vous utilisez une image mappée

Les frames

Certains moteurs de recherche ne savent pas traiter les frames et par conséquent sont dans l'impossibilité de suivre les liens. Pensez donc à réaliser des liens (sur texte ou image) entre les différentes pages qui composent vos frames.

Autre problème concernant les cadres, certains moteurs réagissent comme des navigateurs ne prenant pas en compte les frames

Critère de classement des sites

Aucun moteur de recherche n'affiche ouvertement les critères qu'il utilise pour classer la liste des réponses (pertinence du classement) car ces critères évolus régulièrement.

2 axes malgré tout se distinguent

- La popularité : plus un site à de liens pointant vers lui, plus on considère qu'il est populaire et pas conséquent de bonne qualité.
- La mise à jour régulière: le renouvellement fréquent des pages est un critère de bon classement. On considère qu'un site régulièrement mise à jour est un site qui doit avoir des choses à dire...

Pour être bien placé

Créez des liens explicites

Lorsque vous insérez des liens dans votre page, pensez que le texte de ces liens sera indexé alors utilisez des mots clés pour faire vos liens.

Privilégiez le texte

Rappelez vous que bon nombre de moteurs indexe le contenu des pages et ont du mal avec les images alors soyez texte...

Déposez votre nom de domaine

En général, les moteurs de recherche classe en meilleures positions les sites qui disposent de leur propre nom de domaine.

www.machin.com sera à priori mieux placé que www.monhebergeur.com/machin

Divisez un site en plusieurs services

Votre site possède plusieurs rubriques bien distinctes, référencez chaque rubrique à part.

Renouvelez régulièrement les pages

C'est un critère souvent pris en compte dans le classement des résultats d'une recherche.

Référencez vous régulièrement

Vous changez régulièrement vos pages, forcez (sans abus) les moteurs à revisiter votre site.

La Recherche Sur Internet

LES OUTILS DE SOUMISSION AUTOMATIQUE

Présentation

Référencer son site auprès de la multitude de moteurs et d'annuaires est un travail fastidieux et coûteux (en temps de connexion), aussi peut on faire appel à des logiciels (que l'on appelle des logiciels de référencement ou référenceurs ou encore logiciels de soumission automatique) qui feront le travail à votre place. Ces logiciels existe en général en 2 versions, une payante et une gratuite. On peut supposer que le service payant sera meilleur que l'autre.

Référenceurs et moteurs

La plupart des référencements sur moteurs se limitant à la saisie de l'URL du site et de l'adresse email de son Webmaster, le référencement automatique fonctionnera à merveille !!! Le logiciel de soumission automatique est donc un outils idéal pour se référencer sur plusieurs centaines de moteurs.

Référenceurs et annuaires

Là le problème se pose différemment, en effet les annuaires ont pour la plupart une méthode de référencement différente par exemple les informations fournies pour le référencement sur yahoo ne sont pas les mêmes que celles fournies pour nomade....alors quand on a l'ambition de se référencer sur plusieurs centaines d'annuaires ...

Conclusions

Si vous avez un peu de temps référencez vous sur les sites qui vous paraissent les plus importants. Le référencement sera lent mais de qualité.

RECHERCHE PAR MOTEUR OU ANNUAIRE

On ne peut trouver que des sites référencés :

Il faut comprendre ensuite que à partir du moment ou l'on ne pourra trouver quelque chose que lorsqu'il est référencé, le problème se pose de savoir comment l'auteur d'un site à référencé ses pages...

en effet la structure des pages ayant une forte influence sur la visibilité et le classement, il faut penser aussi à "comment ils ont pu se déclarer"...

D'un côté, un contenu limité mais sagement rangé dans des tiroirs (l'annuaire) de l'autre, une montagne de documents à trier (le moteur). Que choisir?

Pour donner un exemple simple (voire simpliste), on cherchera le site Web d'une entreprise sur un annuaire, mais des renseignements sur un de ses produits (recherche donc à l'intérieur des sites) sur les moteurs de recherche.

Utiliser des annuaires

Si votre recherche est générale (des sites sur le droit constitutionnel ou sur la cuisine créole), posez votre question dans un annuaire. Comme les sites y sont classés en grandes catégories qui s'affinent progressivement, vous serez ainsi orienté et les sites sélectionnés correspondront toujours à votre recherche initiale.

Principaux annuaires francophones (ordre alphabétique)

Nom	Adresse
C'est trouvé	www.ctrouve.com
Francité	www.francite.com
Hachette Net	www.hachette.net
Nomade	www.nomade.fr
Yahoo France	www.yahoo.fr

Principaux annuaires mondiaux (ordre alphabétique)

Nom	Adresse
Encyclopedia Britannica	www.britannica.com
Excite	www.excite.com
LookSmart	www.looksmart.com
Lycos	www-english.lycos.com
Yahoo	www.yahoo.com

Utiliser des moteurs de recherche

Si votre recherche est beaucoup plus pointue ou Si vous ne trouvez pas votre bonheur dans un annuaire, alors dirigez-vous vers un moteur de recherche. Mais attention, pour ne pas être submergé par les "déchets", vous devrez formuler précisément votre requête.

Vous disposez pour cela d'opérateurs logiques et d'éléments de syntaxe. Les plus connus sont les "opérateurs booléens". Il s'agit de mots ou de signes (et, ou, +, -, etc.) destinés à filtrer logiquement la recherche. Ainsi, Si vous entrez les mots jazz + piano dans une fenêtre de recherche, le moteur vous sortira toutes les pages où ces deux mots figurent. Chaque moteur possède sa propre façon d'utiliser les opérateurs. Lisez attentivement leur rubrique d'aide (en effet chaque moteur n'utilise pas forcément la même syntaxe).

Plusieurs serveurs de recherche (Yahoo, AltaVista, Excite, Webcrawler) peuvent lancer une recherche sur une phrase ou un groupe de mot. Si vous êtes en quête de sites sur les vins de la Loire, par exemple, n'entrez pas les mots vins de la loire (la majuscule n'a pas d'importance). Le moteur recherchera en effet séparément chaque mot et vous obtiendrez des milliers de réponses. Tapez plutôt "vins de la Loire" entre guillemets, le serveur n'affichera que les pages contenant la phrase entière.

Principaux moteurs de recherche francophone

k moteurs de recherene maneophone		
Nom	Adresse	
Alta Vista	www.altavista.fr	
Ecila	www.ecila.fr	
Excite France	<u>fr.excite.com</u>	
Infoseek France	www.infoseek.com	
Google	www.google.fr	
Lokace	www.lokace.com	
Lycos France	www.lycos.fr	
Perso-search!	www.perso-search.com	

Voilà	www.voila.fr

Principaux moteurs de recherche mondiaux

Nom	Adresse
Alta Vista	www.altavista.com
Excite	www.excite.com
Hotbot	www.hotbot.com
InfoSeek	www.infoseek.com
Voila	www.voila.com

Vérification du référencement

Altavista: Saisissez la requête: host:monsite.com qui vous indiquera la liste des pages de votre site présentes dans l'index d'AltaVista.

Hotbot : Saisissez la requête : domain:monsite.com qui vous indiquera la liste des pages de votre site présentes dans l'index de HotBot.

Infoseek : Saisissez la requête : site:monsite.com qui vous indiquera la liste des pages de votre site présentes dans l'index d'Infoseek.

Excite: Il n'existe pas de solution simple et "packagée" proposée par Excite. Tapez le nom du serveur (www.monsite.com) dans le formulaire de recherche puis prenez l'option "List by Web site" sur la page de résultats. Les pages sont ainsi regroupées par nom de domaine. Vous devriez y retrouver les vôtres..

Lycos: Allez sur la recherche avancée du moteur et, dans la zone "PAGE FIELD" puis le champ "URL", tapez le nom de domaine de votre site (www.votresite.com).

Northern Light: Saisissez la requête url:monsite.com

Voila : http://voila.fr/Informations/Moteur_guide/recherche_inscription.html

Ecila: Allez à l'adresse http://www.ecila.fr/french/status.html. Saisissez l'adresse de votre site.

Yahoo : Saisissez la requête : u:monsite.com qui vous indiquera immédiatement si votre site est présent sur Yahoo

Faites des échanges de liens

Comme précédemment vu, la popularité d'un site est un facteur déterminant de classement alors faites des échanges avec d'autres webmasters.

Ne supprimez pas (trop) de pages

Cela évite à vos visiteurs l'erreur 404 not found toujours fort désagréable. Redirigez vers une autre page automatiquement ou changez le contenu.

Vérifiez régulièrement votre référencement

Voir chapitre consacré à la vérification.



YAHOO (1° ANNUAIRE)

Adresse: www.yahoo.fr

Présentation

L'objectif de yahoo est de répertorier les **sites Internet en Français indépendamment de la nationalité** de leurs éditeurs ou de la localisation géographique des serveurs d'hébergement.



L'arborescence de Yahoo! est constituée de catégories thématiques imbriquées les unes dans les autres selon les règles suivantes :

- Une distinction nette est faite dans Yahoo! entre les sites commerciaux et les sites non commerciaux. Tous les sites de sociétés sont classés dans les innombrables sous-catégories de Commerce et économie: Sociétés. Tous les sites qui vendent quelque chose sont quelque part dans cette partie de l'arborescence. Cette distinction vous permettra de déterminer par où partir pour votre recherche
- le classement des sites s'effectue selon le critère géographique. La catégorie principale <u>Exploration géographique</u> regroupe en son sein tous les pays, tous les départements, toutes les villes... où les surfeurs de Yahoo! ont pu dénicher des sites en français

Une catégorie est suivie d'un @ lorsqu'elle est placée à plusieurs endroits de l'arborescence.

Recherche par catégorie

cependant les catégories ne sont pas universelles, et ne se retrouvent pas à l'identique d'un annuaire à l'autre ...

Guide des sites Internet - Classement thématique des sites

Actualités et médias

Sports et loisirs

Sujets d'actualité, Télévision, Journaux... Foot, Tourisme, Auto/Moto, Jeux...

Commerce et économie

B2B, Shopping, Emploi, Immobilier...

Art et culture Littérature, Cinéma, Musique, Musées...

Informatique et Internet

Internet, Logiciels, Matériel...

Divertissement

À voir, Loteries, Humour, Sorties...

<u>Santé</u>

Diététique, Médecine, Organismes...

Exploration géographique

Zones régionales, Pays, Europe, France...

Enseignement et formation

Primaire, Secondaire, Supérieur...

Références et annuaires

Dictionnaires, Annuaires, Bibliothèques...

Institutions et politique

Ministères, Droit, Services publics...

Société

Enfants, Gastronomie, Religion...

Sciences et technologies

Animaux, Astronomie, Physique...

Sciences humaines

Archéologie, Histoire, Économie...

Sites de : <u>Algérie - Belgique - Canada - Maroc - Suisse - Tunisie</u> <u>Nouveaux sites - Sites de la semaine - Suggérer un site</u>

Ainsi ici la descente de la catégorie la plus générale à la plus particulière sera:

Exploration géographique

Pays

France

Région

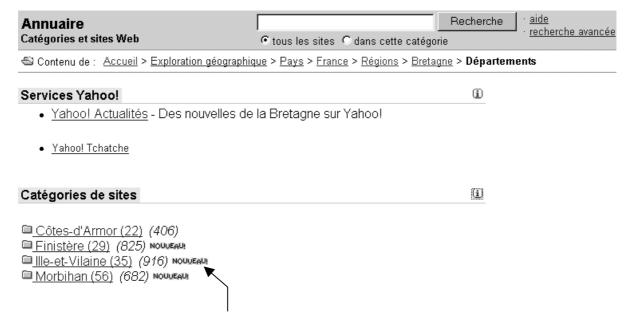
Bretagne

Départements

L' **URL** obtenue est du genre

Agresse ft http://fr.dir.yahoo.com/exploration_geographique/pays/france/regions/bretagne/sports_et_loisirs/tourisme/guides/

avec à l'écran



ill arrive très souvent qu'un site ou une catégorie ait sa place dans plusieurs endroits de l'arborescence. Les catégories qui sont liées à plusieurs endroits sont faciles à reconnaître : une <u>arrobas</u>: **@**) suit leur nom. Si vous croisez une catégorie dont le nom se termine par une @, cela veut dire qu'elle habite à un autre endroit de l'arborescence, et qu'en cliquant sur ce lien vous serez envoyé à cet autre endroit!

Yahoo! France classe les résultats de la façon suivante :

Recherche par mots clés

lors de la recherche



Si la requête est composée de plusieurs mots, yahoo recherche d'abord les sites qui comportent tous les mots puis ceux comportant un des mots.

Section du document : 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.

Hiérarchie des catégories: Parmi les catégories correspondantes, celles qui sont situées en haut de l'arborescence Yahoo! France (donc plus générales) sont mieux classées que les catégories inférieures (plus précises).

Catégories Yahoo! (1 - 18 sur 18)

Exploration géographique > Pays > France > Régions

Bretagne

🜓 Exploration géographique > Pays > France > Régions > **Bretagne** > Commerce et économie > Sociétés

• Crédit mutuel de Bretagne (CMD)

Exploration géographique > Pays > France > Régions > Bretagne > Enseignement et formation

• Institut universitaire de formation des maîtres (IUFM) - Bretagne

Nombre de mots clés à employer :

Pour affiner une recherche qui a donné de trop nombreux résultats, tapez un terme plus spécifique ou ajoutez un mot clé (cf Nomade)

Sensibilité à la casse :

Non: ECOLE=ecole

Sensibilité aux accents :

Oui : ecole différent de école (plus limité)

Faites un essai avec "péché", puis "peche" puis "pêche"

Gestion des pluriels :

Non: cheval est différent de chevaux

Recherche sur une expression précise

Vous pouvez également faire une recherche sur une expression stricte en tapant celle ci **entre guillemets**. Mettre un groupe de mots entre guillemets permet en effet de restreindre les résultats de la recherche aux documents contenant les mêmes mots, dans le même ordre.

comparez: maladie de la vache folle à "maladie de la vache folle"

Mots exigés et mots interdits

L'insertion de l'un des opérateurs ci-dessous permet d'exiger ou d'interdire l'apparition d'un ou plusieurs mots dans les résultats de la recherche.

+ : Attacher un + devant un mot exige que le mot en question apparaisse dans chacun des résultats de la recherche (réglage par défaut)

comparez: france à france +île

- : Attacher un - devant un mot exige que le mot en question n'apparaisse dans aucun des résultats de la recherche.

comparez: **guerre** à **guerre – etoile**

Restrictions à un corps de document

L'insertion de l'un des opérateurs ci-dessous devant un mot recherché permet de restreindre la recherche à un corps de document.

t: - restreint la recherche au seul titre des documents comparez: <u>jules verne</u> à <u>t:jules verne</u>

u: - restreint la recherche à la seule adresse Web (URL)

comparez: <u>culture</u> à <u>u:culture</u>

Correspondance élargie (*)

Attacher un * à la fin d'une chaîne de caractère élargira la recherche à tous les mots commençant par la chaîne de caractère précisée.

comparez: <u>lune</u> à <u>lune*</u>

Combiner la syntaxe

Vous pouvez combiner n'importe lesquelles de ces syntaxes tant que vous le faites dans le bon ordre, qui est celui utilisé dans cette page, à savoir : +, -, t:, u:, '"' et enfin *.

comparez: (correct) <u>+t:football - américain</u> versus (incorrect) <u>t:+football - américain</u>

Pertinence du classement

Qu'est-ce qui fait que un site apparaît en tête de liste ou non lors de votre recherche?

plusieurs facteurs dont les suivants :

- le nombre de mots clés correspondants. Plus les mots ou chaînes de mots correspondent à une entrée de la base Yahoo!, plus cette dernière apparaîtra en haut des résultats.
- l'exactitude des correspondances. Elles sont classées plus haut que les correspondances approximatives.
- l'endroit où la requête a été trouvée. Une correspondance dans le titre d'un site a priorité sur la même correspondance dans le commentaire d'un site.

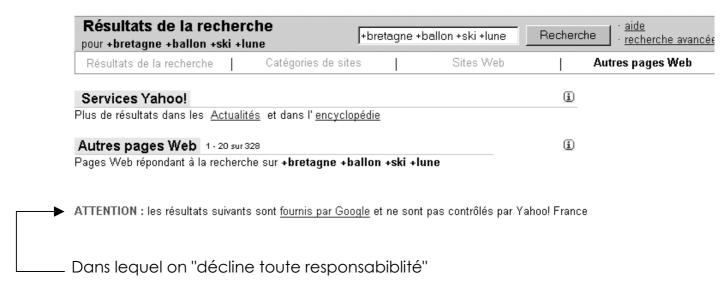
encore une fois une grosse partie du travail relève du référencement...

Recherche infructueuse

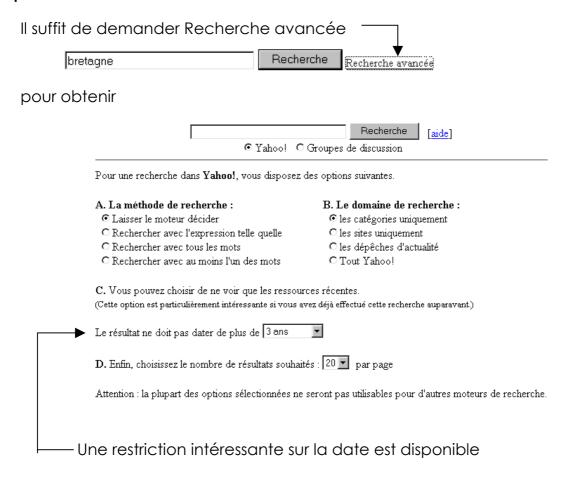
Lorsque **Yahoo!** ne trouve pas de résultat dans sa base de données, il peut passer la main au moteur de recherche **Google** qui effectue une recherche en texte intégral. Cela sedemande via Autres pages Web



on obtient alors



Les options de recherche



Yahoo! dans le monde

Yahoo! dans le monde

Europe : Allemagne · Danemark · Espagne · Italie · Norvège · R.U. & Irlande · Suède

Asie: Australie & N.Z. · Chine · Corée · HK · Inde · Japon · Singapour · Taiwan · en anglais · en chinois

Amériques : Yahoo! [US] : Argentine : Brésil : Canada : Mexique : en espagnol

Autres services Yahoo!

 $\underline{\textit{Culture/Divertissement}}: \underline{\textit{Cinéma}} \cdot \underline{\textit{Musique}} \cdot \underline{\textit{People}} \cdot \underline{\textit{Horoscope}} \cdot \underline{\textit{Logithèque}} \cdot \underline{\textit{Programmes Télé}}$

Finance : Bourse · Banque en ligne · Conversion de devises · Impôts · Vidéo

Actualités : À la Une : Monde : France : Local : Économie : Multimédia : Sciences : Santé : Dossiers

Sport: A la Une Football Formule 1 Tennis Rugby Basket Cyclisme et aussi...

Àvoir : Nouveautés · Sélection de la semaine · Le choix des surfeurs

ALTAVISTA (1° MOTEUR)



Adresse: www.altavista.fr

Présentation

L'objectif de altavista est de répertorier les **sites Internet en Français indépendamment de la nationalité** de leurs éditeurs ou de la localisation géographique des serveurs d'hébergement.

AltaVista a été lancé par la société Digital Equipment en décembre 1995, société rachetée par Compaq en 1998. En 1999 il est devenu la propriété de la société américaine CMGI.

L'interface depuis février 2001 est beaucoup plus dépouillée, ce qui sera apprécié des strictes utilisateurs du moteur de recherche.

<u>AltaVista France</u>, lancé en février 2000, compte 11 millions de pages indexées en français.

Ce moteur se veut le plus complet des outils de recherche et effectue une indexation du texte intégral de 1,2 milliards de pages Web pour n'en garder que 350 millions (mai 2000). Il propose aussi 30 millions de fichiers multimédia (sons, images, vidéos). L'index est réactualisé en fonction de la stabilité des sites, déterminée par le robot Scooter à partir des visites précédentes.

D'autre part AltaVista fait appel à Deja pour les groupes de News et offre un service appellé **Translate**: Ce service est fourni par la société **Systran**. Il suffit de donner du texte (texte ASCII) ou un URL. La traduction opère sur 2000 octets au maximum. Dix couples de langues sont disponibles.

Alta Vista offre deux modes de recherche pour interroger les serveurs web:

- o Une recherche simple
- o Une recherche avancée

Recherche simple

Pas d'opérateurs booléens, la question est une suite de mots simples ou composés (phrase). L'opérateur implicite est le OU.



La casse

Requête saisie en minuscules : toutes les occurrences sont recherchées.	mac trouve mac, MAC ou Mac.
Requête saisie en majuscules : l'occurrence exacte est recherchée.	Mac trouve Mac mais pas mac ou MAC.

Les accents

Requête saisie sans accents : toutes les occurrences sont recherchées.	erable trouve érable, erable.
Requête saisie avec accents : l'occurrence exacte est recherchée.	érable trouve érable mais pas erable.

Ordre des mots

L'ordre	dans	lequel	sont	Important : paris dakar donne u résultat différent de dakar paris
saisis	les	mots	est	
importa	nt.			

Recherche de phrases

Recherche les pages	"les memoires d'un ane" est
contenant les mots ou la	différent de les memoires d'un
phrase exacte.	ane.

Troncation

xxxx*	chocolat*	trouve	chocolat,
* remplace une suite de caractères (jusqu'à 5). Peut	chocolatine ou	chocolaté.	
être au milieu d'un mot.			

L'opérateur OU

Opérateur par défaut	moteur	rec	herche	trouve	les
	pages	qui	contie	nnent	soit
	moteur,				
	deux.				

L'opérateur ET

+	+moteur +recherche recherche
	les pages qui contiennent à la
	fois moteur ET recherche.

L'opérateur SAUF

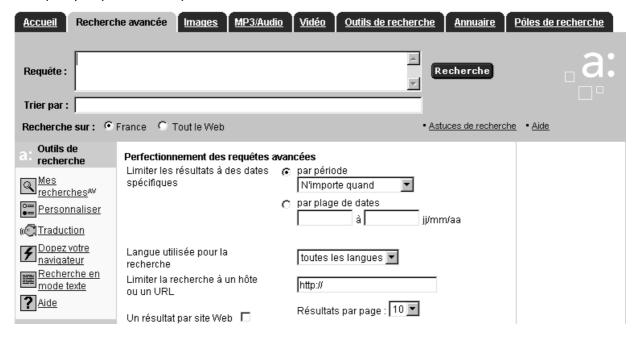
-	moteur -recherch* recherche
	les pages qui contiennent
	moteur SANS recherch

L'opérateur PROCHE DE

~	moteur ~auto* recherche les
	pages qui contiennent moteur
	et auto proches l'un de l'autre
	(contenus dans une fenêtre de
	10 mots au maximum)

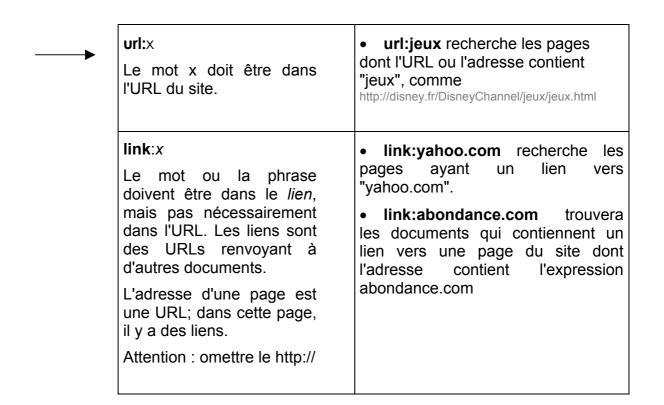
Recherche avancée

il faut obligatoirement utiliser les opérateurs AND, OR, NOT ou NEAR en majuscule pour combiner plusieurs termes. Les parenthèses sont utilisées lorsqu'il y a plusieurs opérateurs.



Les opérateurs

Il y a beaucoup d'opérateurs disponibles, les plus utiles sont marqués d'une flèche



	text:x Le mot ou la phrase peuvent être n'importe où dans le texte de la page. Note : cela exclut images, liens et URL	text:"ce plat pays qui est le mien" recherche les pages dont le texte contient cette phrase.
	image:x Recherche les images dont le nom contient x.	image:coquelicot recherche les images dont le nom contient "coquelicot" comme coquelicot.jpg ou groscoquelicots.bmp.
	title:x Le mot ou la phrase doit être dans le champ titre ou dans le meta-titre. Note : pas d'espace entre title, les deux points et l'expression.	title:chocolat recherche les pages dont le titre contient chocolat. title:"chocolat au lait" title:((chocolat OR praline) AND (serbe OR irakienne))
	host:x Recherche les pages sur un hôte déterminé. Un hôte est par exemple "multimania.com" ou "www.yahoo.com". A ne pas confondre avec le domaine qui est "fr" ou "com".	host:multimania recherchera les documents qui sont disponibles sur une machine dont le nom contient le mot multimania
	like:x Recherche de sites similaires ne fonctionne pas sur www.altavista.fr	like:http://www.tf1.fr/ recherchera les sites similaires à tf1.
	domain:x Recherche les pages appartenant à un domaine déterminé.	domain:org recherche les pages appartenant au domaine "org".
	anchor :x Recherche sur les intitulés des liens hypertexte	anchor:abondance trouvera les documents qui contiennent un lien dont l'intitulé (donc le texte en couleur et souligné) contient le mot abondance.

Logique booléenne

x AND y x + y	Recherche les pages contenant x ET y On peut utiliser le + à la place du AND	pages contenant à la fois
x OR y x y	Recherche les pages contenant x OU y. On peut utiliser le (symbole obtenu par Altgr 6) à la place du OR	pages contenant "dopage" OU
x AND NOT y x – y		festina AND NOT dopage trouve les pages contenant "festina", mais pas "dopage". (rares)
x NEAR y x ~y	Recherche les pages dont x est éloigné de y d'au plus 10 mots. On peut utiliser le ~ (symbole obtenu par Alt 126) à la place du NEAR Pratique pour trouver des termes associés dans le même contexte.	recherche les pages qui contiennent moteur et automobile proches l'un de l'autre (contenus dans une
()	Les parenthèses groupent des actions.	dopage AND (football OR cyclisme) trouve les résultats contenant dopage ET football, OU dopage ET cyclisme.

Intervalle de dates



La date prise en compte est la date de dernière **modification** des documents

Pertinence du classement

la requête est trouvée dans les premiers mots du document, spécialement dans le titre.

les termes sont trouvés les uns à côté des autres dans le document.

le document contient plus de mots de la requête que les autres documents.

Importance relative des différents critères		
Titre	***	
Balise META keywords	*	
Indice de popularité de la page	**	
Corps du texte	**	

Google Google

GOOGLE (2° MOTEUR)

Adresse: www.google.fr

Présentation

C'est le dernier né dont les resultats se font de plus en plus époustouflants...



Lorsque vous soumettez une requête de base, Google renvoie uniquement les pages Web qui contiennent **tous les mots** spécifiés (qu'ils soient adjacents ou non); par conséquent, pour mieux cibler votre recherche et obtenir une liste de résultats à la fois plus courte et plus pertinente, il suffit d'ajouter d'autres termes de recherche à votre requête

Pour exclure un mot de votre recherche, placez le signe moins (-) immédiatement devant ce mot. (Le signe moins doit être précédé d'un espace.)

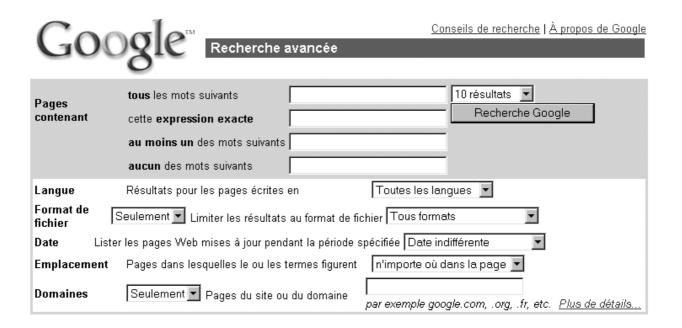
Pour rechercher une expression (par exemple, «imprimante laser couleur»), utilisez des guillemets. Lorsque vous placez plusieurs mots entre guillemets (par exemple, "mots entre guillemets"), Google les considère comme une seule chaîne de caractères et recherche uniquement les documents qui contiennent ces mots dans la même séquence.

Google ne supporte pas les recherches à base de caractères joker. Autrement dit, Google utilise les mots exactement tels que vous les entrez dans le champ de recherche. Exemple: Si vous entrez le terme « cheval » ou « cheva* », Google ne fait pas porter la recherche sur « chevaux », « chevaline » ou « chevalerie », mais uniquement sur le terme « cheval » ou « cheva* » (soit la chaîne de caractères « cheva » suivie d'un astérisque)

Les recherches Google ne tiennent pas compte de la « casse » des lettres (MAJUSCULES/minuscules)

Par défaut, les recherches Google ne tiennent pas compte des accents ou autres signes diacritiques (cédille, tilde espagnol, umlaut allemand, etc.). Ainsi, les termes [FRANÇAIS] et [FRANCAIS] retrouvent les mêmes pages. Pour indiquer que ces deux termes ont une signification différente, utilisez un signe plus (+), soit les termes de recherche [+FRANÇAIS] et [+FRANCAIS]

Recherche avancée



Une vision... annuaire?

Le répertoire Google (http://directory.google.com) permet de rechercher les sites Web les plus pertinents pour un sujet donné et dans un secteur donné. Exemple -- Si vous recherchez le terme « bardot » dans la catégorie Arts > Audio-visuel > Cinéma > Acteurs et actrices du répertoire Google, les résultats concerneront uniquement l'actrice, et non les animaux qu'elle protège ni un cabinet immobilier de Châteaubriant...

En « descendant » ainsi dans différentes sous-catégories du répertoire, vous parvenez très rapidement à réduire le nombre des pages qui vous intéressent et à en augmenter la pertinence.



The web organized by topic into categories.

Arts	Home	Regional
Movies, Music, Television,	Consumers, Homeowners, Family,	Asia, Europe, North America,
Business Companies, Finance, Jobs,	Kids and Teens Computers, Entertainment, School,	Science Biology, Psychology, Physics,
Computers	News	Shopping
Internet, Hardware, Software,	Media, Newspapers, Current Events,	Autos, Clothing, Gifts,



VOILA (MOTEUR+ANNUAIRE)

Présentation

Lancé en juillet 98, Voila est la suite du moteur Echo racheté par France Telecom. Sous forme de portail **Voilà** regroupe un moteur de recherche sur le web francophone, un annuaire intégré (Guide), les actualités, la météo et des informations boursières. Les annuaires <u>Pages Jaunes</u>, Pages Blanches, Les marques, Les rues commerçantes et Les pages pro complètent l'offre pour la France.

Voila recherche des documents dans une base de données de plus de 11 000 000 de pages web en langue française mise à jour chaque semaine. Les sites inscrits manuellement sont indexés sous 15 jours.

Voila offre un formulaire de recherche simple et un formulaire de recherche avancée.

La recherche simple



La casse

Sans	importance :	toutes	IBM trouve ibm, IBM ou Ibm.
les	occurrences	sont	
rech	erchées.		

Les accents

Sans importance : toutes	bébé donne le même résultat
les lettres accentuées ou	que bebé ou bebe
non sont cherchées	

Ordre des mots

saisis les mots est sans	paris dakar donne le même résultat que dakar paris
importance.	

Recherche de phrases

Recherche les pages	"les memoires d'un ane" est
contenant les mots ou la	différent de les memoires d'un
phrase exacte.	ane.

L'opérateur OU

Opérateur par défaut	moteur recherche trouve les pages qui contiennent soit moteur, soit recherche, soit les deux.
Opérateur par défaut	moteur recherche trouve les pages qui contiennent soit moteur, soit recherche, soit les deux.

L'opérateur ET

+	+moteur +recherche recherche
	les pages qui contiennent à la
	fois moteur ET recherche.

L'opérateur SAUF



L'opérateur PROCHE DE

~	moteur ~auto* recherche les
	pages qui contiennent moteur
	et auto proches l'un de l'autre
	(contenus dans une fenêtre de
	10 mots au maximum)

Recherche sur le type de fichiers

Menu déroulant : Son, image, vidéos



La recherche avancée

On y accède par Recherche approfondie



on peut voir que les mêmes opérateurs que Yahoo sont disponibles, (on peut les taper ...)

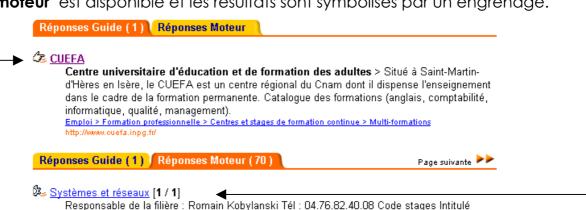
Rechercher par type de so	игсе						
Baisir une requête, sélectionner l	a source ci-dessous puis	valider:					
e recherche :	voila	a!					
C Dépêches de l'actualité	C Sélection de sites	Tout I	e web francophone	C Guide	mondial	C Encyclopé	die
Plus d'options							
Attention : ces options de rec Les champs marqués d'une :		'à la source	"tout le web francoph	one".			
Je recherche :							
Tous les mots suivants : *				Placés :	n'importe o	où	¥
Aucun des mots suivants :							
Dans le domaine :		(ex. : w	ww.voila.fr ; .fr ; guide.v	oila.fr/1100	0/11155)		
Contenant des fichiers de ty	pe: C Sons C Images	C Vidéos					
Mis à jour avant :	n'importe quand	▼					
Dans la thématique : (un	<u>e seule thématique à la fo</u>	is)					
Arts, culture		▼	Emploi, Enseignen	nent			▼
Tourisme, voyage		▼	Informatique, intern	iet			▼
Villes		▼	Business, économ	ie			▼
Sports		V	Administration, poli	tique			▼
Vie pratique		V	Science, médecine				▼
Charme		▼					

Pertinence du classement

Importance relative des différents critères							
Titre	***						
Balise META keywords	*						
Indice de popularité de la page	**						
Corps du texte	*						
Autres champs	L'URL est très importante, ainsi que le positionnement des mots dans des balises spécifiques (<h1> ou par exemple)</h1>						

Les réponses du "Guide de Voila" (partie annuaire du portail) apparaissent dans la partie "**Réponses Guide**" et sont symbolisées par une main

Pour les réponses émanant du moteur (les pages web), l'onglet "**Réponses moteur**" est disponible et les résultats sont symbolisés par un engrenage.



Responsable de la filière : Romain Kobylanski Tél : 04.76.82.40.08 Code stages Intitulé
Public concerné et objectifs Niveau d'entrée Durée Coût S1 Utiliser UNIX
http://www.cuefa.inpg.ft/pages/tcont/fichesprog/systemes.htm | 22 Ko | + infos

Récemment, une nouvelle catégorie, nommée Encyclopédie voila a fait son apparition...

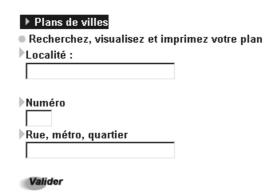
Encyclopedie Voila avec Hachette

Pages jaunes/blanches/email



Plans et Itinéraires





ltinéraire en Europe Durée du trajet, carburant, Départ : péages dépendent des critères suivants. Modifiez - les pour une Arrivée : réponse personnalisée. EXPRESS • votre véhicule : **Etapes** Routière (facultatif) votre carburant : 1 Essence 1.1 EURO/L 2 vos indemnités : 0.0 EURO /km 3 votre devise: **EURO** 4 <u>les distances</u>: kmOrdre des étapes : ordre indiqué <u>col d'altitude</u> : Autoriser

Valider

RECAPITULATIF - MOTEURS DE RECHERCHE

	<u>Altavista</u>	<u>Excite</u>	<u>Hotbot</u>	<u>Infoseek</u>	<u>Webcrawler</u>	<u>Lycos</u>	Northern Light
Millions de pages indexées	150	55	110	45	2	50	140
Millions de pages visitées par jour	10	3	Jusqu'à 10	-	-	6 à 10	+ de 3
Fréquence de rafraîchissement	1 jour à 6 semaines	1 jour à 6 semaines	1 jour à 2 semaines	1 jour à 2 mois	Hebdomadaire	-	2 à 4 semaines
Pages soumises	1 à 2 jours	< 2 semaines	< 2 semaines	< 2 jours	< 2 semaines	2 à 3 semaines	2 à 4 semaines
Pages non soumises	1 jour à 1 mois	Jusqu'à 6 semaines	2 semaines	1 à 2 mois	< 2 semaines	2 à 3 semaines	2 à 4 semaines
<u>Profondeur</u>	Sans limites	Sans limites	Sans limites	Echantillon	Echantillon	Echantillon	Sans limites
Support des cadres	Oui	Non	Non	Non	Non	Non	Oui
Support des images en coordonnées	Oui	Non	Non	Oui	Oui	Non	Oui
Protection par mot de passe	Non	Oui	Non	Oui	Non	Oui	Oui
Popularité du lien	Non	Oui	Oui	Non	Oui	Oui	Non
Contrôle de la fréquence de mise à jour	Oui	Non	Oui	Oui	Non	Non	Non
Robots.txt	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Meta – Robots	Oui	Oui	Oui	Oui	Oui	Oui	Oui

Recherche Sur Internet v: 2.0 Cabaré Michel Page 63/81

<u>Millions de pages indexées</u> : nombres à prendre avec des pincettes, à cause de la difficulté de ce calcul. Les moteurs n'ont pas tous la même structure et les mêmes méthodes d'indexation, certaines pages peuvent être comptées en double. Néanmoins, plus le moteur est gros, plus la chance de trouver la page que vous cherchez est importante.

Millions de pages visitées par jour : nombre de pages pouvant être visitées par jour. Plus ce nombre est important, plus la probabilité d'avoir une base de données à jour est forte. Mais ce n'est pas le seul critère du degré de mise à jour.

<u>Fréquence de rafraîchissement</u>: étant donnée la rapidité de croissance du Web, un moteur de recherche peut être rapidement dépassé. L'âge des listes de sites peut varier de quelques jours à quelques mois. En effet, si les robots indexent relativement rapidement les pages soumises, il n'en est pas de même pour les pages non soumises, et certains robots parcourent plus souvent certaines parties plus populaires du Web que d'autres moins populaires.

<u>Pages soumises</u> : délai d'indexation des pages soumises aux moteurs de recherche. Ces délais sont beaucoup plus courts qu'en ce qui concerne les pages non soumises.

<u>Pages non soumises</u>: une fois qu'une page a été soumise, le robot suit les liens à partir de cette page. Ces autres pages ont souvent un délai d'indexation supérieur aux pages directement soumises, car les moteurs remettent à plus tard la visite des pages non soumises tout en indexant presque instantanément les pages soumises.

Profondeur : cela a trait aux pages non soumises. Cela indique le nombre de pages au-delà de la page soumise que le robot va indexer. Existent deux possibilités :

- Sans limites: le robot essaie d'aller partout, le plus loin possible. Dans la pratique, il n'indexe évidemment pas toujours toutes les pages.
- Echantillon: les robots rassemblent un échantillon plus ou moins gros des pages. Généralement, plus le site est populaire, mieux il sera représenté.

Beaucoup d'obstacles peuvent arrêter un robot : cadres, images en coordonnées, mots de passe, etc. (voir plus bas).

Support des cadres : capacité des moteurs de recherche à gérer les cadres dans les sites indexés.

Support des images en coordonnées : capacité des moteurs de recherche à gérer les images en coordonnées.

<u>Protection par mot de passe</u>: capacité des moteurs de recherche à gérer les pages protégées par mot de passe. Certains moteurs de recherche peuvent pénétrer dans un site protégé, si l'administrateur des pages a fourni le mot de passe lors de l'indexation du moteur de recherche. Les utilisateurs devront toujours passer par un formulaire d'entrée, mais la page protégée sera indexée dans le moteur de recherche.

<u>Popularité du lien</u> : les moteurs de recherche peuvent analyser la popularité d'une page en analysant le nombre de liens qui pointent vers cette page. En dessous d'un certain seuil de popularité, les sites ne sont pas indexés.

Contrôle de la fréquence de mise à jour : certains moteurs de recherche ont les moyens de connaître la fréquence de mise à jour des pages. Plus souvent la page est mise à jour, plus elle sera visitée par le robot.

Robots.txt : robot.txt est un fichier mis dans le répertoire d'un site qui permet d'arrêter les robots des moteurs de recherche.

<u>Meta - Robots</u>: c'est un tag spécial permettant aux propriétaires de sites de préciser que la page ne doit pas être indexée: <META NAME= "ROBOTS" CONTENT="NOINDEX">

Capacité des moteurs de recherche à accepter des critères de recherche étendus.

Un moteur de recherche doit pouvoir offrir à ses utilisateurs la possibilité d'effectuer des requêtes plus ou moins complexes : recherche sur certaines parties d'Internet, du Web, dans le titre d'une page, dans son URL, utilisation d'opérateurs booléens, etc.

	<u>Altavista</u>	<u>Excite</u>	<u>Hotbot</u>	Infoseek	<u>Yahoo</u>	Webcrawler	Lycos	Northern Light
Sensibilité à la casse	Oui	Non	Oui	Oui	Non	Non		
Troncation *	* 5 caractères	Non	Oui	Non	Oui à la fin du mot		Non	* et %
Phrase	Oui	Oui	Oui	Possible en recherche avancée	Oui	Oui	Oui	
Groupement ()	Oui	Oui	Oui	Non		Oui	Oui	Oui
Recherche Booléenne	OR AND NEAR AND NOT	AND OR AND NOT	AND OR NOT	Non	Non	AND OR NOT	AND OR NOT ADJ NEAR	AND OR NOT
+ -	Oui	Oui	Oui	Oui	Oui	Oui		Oui
Limiteurs de champs	title, url, link, text,image, anchor, host, domain, applet	Non	Domain, depth, feature(acrobat, applet, activex, audio, embed, flash, form, frame, image, script, shockwave, table, video, vrml), outgoingurltext, newsgroup, scriptlanguage, title, after, before, within	Link, site, url, title, alt	t (titre) u (url)		Non	url, title, pub, company, ticker, text
Age des documents	Oui	Non	Oui	Non	Oui	Non	Non	Oui

Critères de classement des résultats obtenus.

	<u>Altavista</u>	<u>Excite</u>	<u>Hotbot</u>	Infoseek	Webcrawler	<u>Lycos</u>	Northern Light
Popularité du lien	Non	Oui	Non	Oui	Oui	Oui	Non
<u>Les Meta balises</u> <u>favorisent le classement</u>	Non	Non	Oui	Oui	Non	Non	Non
Rafraîchissement Meta	Spam	Accepté	Accepté	Spam	Accepté	Accepté	Accepté
<u>Texte invisible</u>	Spam	Accepté	Spam	Spam	Accepté	Spam	Spam
Texte minuscule	Spam	Accepté	Spam	Accepté	Spam	Spam	Accepté
Indexation des textes ALT	Oui	Non	Non	Oui	Non	Oui	Non
Indexation des commentaires	Non	Non	Oui	Non	Non	Non	Non
Stemming	Non	Non	Non	Oui	Non	Oui	Oui
Emplacement des mots- clefs sur la page	trouvée dans les premiers mots du document, spécialement dans le titre. - les termes sont trouvés les uns à côté des autres dans le document. - le document	présents comme les plus satisfaisant aux critères. Ex : chien chien chien chat présentera d'abord la page où on parle de chiens. Habituellement, si ces termes apparaissent dans le titre, ces			Prône une recherche en langage naturel. En entrant comme requête " le plus grand pays du monde ", Webcrawler supprime les éléments accessoires (le bruit) et opère une recherche avec AND sur les autres. Les documents présentant tous les termes sont proposés en premier, mais les pages présentant un seul des termes sont aussi proposés. Le problème est que la définition du " bruit " selon Webcrawler est plutôt floue.		

Popularité du lien : indique si le fait que la page soit pointée dans d'autres sites accroît son rang dans le classement.

<u>Les meta-balises favorisent le classement</u> : indique si l'existence des rubriques "keywords" (mots-clefs) et "description" dans les meta-balises de la page accroissent son rang dans le classement.

<u>Texte invisible</u>: cette technique consiste à placer du texte de la même couleur que le fond de la page (donc invisible), et dont le but est d'être lu par les robots des moteurs de recherche. Indique si cette méthode est acceptée par les robots, ou considérée comme du Spam, et dans ce cas la page peut ne pas être indexée.

<u>Texte minuscule</u>: cette technique consiste à placer du texte de très petite taille sur la page que l'on désire indexer, en vue d'être lu uniquement par des robots. Certaines pages dont la proportion de très petit texte est importante peuvent être refusées à l'indexation, ou le texte lui-même peut ne pas être pris en compte.

<u>Indexation des textes ALT</u> : ce sont les textes associés à des images, que l'on voit sous certains navigateurs quand le curseur est sur l'image. Indexation des commentaires : voir commentaire "indexation des textes ALT" ci-dessus.

<u>Stemming</u>: le "stemming" est la capacité qu'a un moteur de recherche d'adopter différentes formes pour le mot-clef proposé. Par exemple, "art" peut devenir "arts" ou "artistique" ou "artiste".

Emplacement des mots-clefs sur la page : de façon générale, les moteurs de recherche présentent d'abord les pages où figurent tous les mots-clefs de la requête, ou figurent dans le titre, puis les pages où une partie des mots-clefs figurent, etc. lci figurent les particularités de certains moteurs, qui appliquent cette règle avec des paramètres différents.

Ces critères sont toutefois assez flous et forment un élément de la politique des moteurs qui, grâce à la pondération des résultats, choisissent en fait de privilégier certains types de sites.

Recherche Sur Internet v: 2.0 Cabaré Michel Page 67/81

METAMOTEURS DE RECHERCHE

Un métamoteur, ce n'est pas mieux qu'un moteur?

Vaste débat... Les métamoteurs sont des outils qui, pour une même requête, interrogent plusieurs moteurs de façon simultanée, rapatrient les résultats, les synthétisent et vous proposent un récapitulatif des réponses données. L'idée est séduisante...

Cependant, il n'est pas permis, sur ces services, d'utiliser les fonctionnalités avancées des moteurs de recherche (notamment les fonctions de type host:, url:, title: d'AltaVista, par exemple), tout simplement parce qu'elles varient grandement selon les outils interrogés. Il est évident que la mise en place de ces fonctionnalités dans le cadre d'une recherche simultanée sur plusieurs moteurs est loin d'être aisée, voire tout simplement impossible.

D'autre part, les métamoteurs font la synthèse de résultats fournis par plusieurs moteurs différents, classant chacun leurs résultats de façons différentes, sans utiliser les mêmes critères de pertinence. Une synthèse de documents classés de façons ainsi disparates est-elle si simple que cela à effectuer, et surtout, est-elle plus pertinente ? Peut-on comparer le résultat de Yahoo! avec celui d'AltaVista pour un même mot clé ? La question reste posée...

L'utilisation de ce type de métamoteurs engendre un autre problème de fond : quasiment tous les moteurs de recherche sur lesquels ils s'appuient se financent grâce aux bandeaux publicitaires qu'ils affichent. Or, les promoteurs de cette couche logicielle supplémentaire que sont les métamoteurs ne répercutent pas systématiquement (ou pas du tout) ces bandeaux, préférant même parfois proposer leurs propres annonces. Le recours à ces métamoteurs réduit donc de façon substantielle le nombre d'accès au moteur de recherche traditionnel, ce qui compromet ses recettes publicitaires et risque, à terme, de signer son arrêt de mort.

D'autre part, se pose un problème d'éthique : est-il juste d'utiliser pour son propre compte les technologies et investissements mis en œuvre par d'autres sociétés, sans contrepartie financière ? Chacun méditera sur ce point...

Les métamoteurs de recherche

Si votre demande est très ciblée, vous pouvez essayer un multi-moteurs. Ces services de recherche interrogent simultanément plusieurs moteurs ou annuaires et trient le résultat. Cela évite d'interroger plusieurs serveurs à la suite Si les résultats ne sont pas satisfaisants. En revanche, on perd un peu en rapidité et les requêtes doivent rester simples pour s'adapter à tous les moteurs. Les principaux multi-moteurs

Principaux méta-moteurs de recherche

Nom	Adresse
Metacrawler	www.metacrawler.com
	OU
	www.go2net.com/search.ht
	<u>ml</u>
Find it	www.iTools.com/find-it/find-
	<u>it.html</u>
General moteur	www.acorus.fr/general
Research it	www.iTools.com/research-
	<u>it/research-it.html</u>

Les Métachercheurs, tableau comparé

Dans la colonne **" Moteurs principaux utilisés"** figure la liste des moteurs principaux utilisés par les métachercheurs

A=Altavista, E=Excite, H=Hotbot, I=Infoseek, L=Lycos, W=Webcrawler, Y=Yahoo

Dans la colonne "Manque" figure l'absence des principaux moteurs (A, E, H, I, L, W, Y).

Dans la colonne "Autres" figurent tous les autres moteurs.

La colonne "**Traitement des opérateurs**" indique quels opérateurs sont pris en charge et traduits par les métachercheurs. Quand ils sont pris en charge, les requêtes sont adaptées aux exigences des moteurs de recherche (AND peut devenir +, etc.); dans le cas contraire, les opérateurs sont envoyés en l'état aux moteurs de recherche, qui ne les comprennent pas forcément.

La colonne "Choix des moteurs" indique s'il est possible de choisir les moteurs de recherche au lancement de la requête.

La colonne "Présentation des résultats" précise les critères de classement des résultats.

		Les Mé	tachercheurs			
Les Métachercheurs	Moteurs principaux	Moteui Mangue	Autres	Traitement des	Choix des moteurs	Présentation des résultats
All4one www.all4one.com	utilisés A, E, H, L	Infoseek, Webcrawler, Yahoo	-	opérateurs -	-	Par moteur
Cyber411 www.c4.com	A, E, H, L, W, Y	Infoseek	Dejanews - Galaxy - Goto - Looksmart - Magellan - PlanetSearch - Search.com - Snap - Thunderstone - What- U-Seek	""()OR AND NOT NEAR	Oui	Par moteur
<u>Debriefing</u> www.debriefing.com/france	A, E, H, I, L, W, Y					
Dogpile dogpile.com	A, E, H, I, L, W, Y	-	WEB: Excite Guide - Goto Com - LycosA2Z - Magellan - PlanetSearch - Thunderstone - What-U-Seek USENET: Altavista News - Dejanews - Dejanews' Old Database - Reference.com. FTP: Fast FTP Search. INFOS: Excite News - Infoseek News Wires - Yahoo News Headlines.	""()OR AND NOT NEAR	-	Par moteur
Inference-find www.infind.com	A, E, I, W, Y	Hotbot, Lycos	-	-	-	Par domaine
Méga Francité www.mega.francite.com						
Metacrawler www.metacrawler.com	A, E, I, L, W, Y	Hotbot	Looksmart - The Mining Co Thunderstone.	" " + OR Intégré au formulaire	Oui	Par coefficient d'adéquation à la requête et par site
Metafind www.metafind.com	A, E, I, W	Hotbot, Lycos, Yahoo	PlanetSearch	" " () OR AND NOT NEAR	-	Au choix : par moteur, alphabétiqueme nt, domaine, mot-clef
Savvy Search www.savvysearch.com/lan g/french.html	A, E, I, H, L, W	Yahoo	Direct Hit - Galaxy - Google - NationalDirectory - Thunderstone pour la recherche générale et beaucoup plus pour des recherches spécialisées (Usenet, infos, etc)	""()+-OR AND NOT NEAR		Par moteur puis par site. Format bref, normal, étendu.
Trouvez ! www.trouvez.com/						

LES AGENTS INTELLIGENTS ET ASPIRATEURS

Agents intelligents

Un méta-chercheur qui vous cherche dans la plupart des moteurs de recherche et archive ses résultats sur votre disque dur. Indispensable.

BullsEye

www.intelliseek.com

Copernic

www.copernic.com

Express by infoseek express.infoseek.com

Hurricane Websearch www.aatecomm.com

Inforia Quest www.inforia.com

l'exemple de Copernic

Vous pouvez aussi vous servir de logiciels pour vous assister dans vos recherches. Les agents intelligents sont là pour ça. Ces programmes que vous installez sur votre ordinateur fonctionnent un peu comme les sites de recherche sauf que vous formulez votre requête hors connexion. Quand vous avez tapé vos **mots clés**, le logiciel se connecte à plusieurs annuaires ou moteurs de recherche, trie ,met en forme le résultat et rapatrie le tout sur votre disque dur tout en éliminant les répétitions. Vous avez alors tout le temps de consulter les réponses.

par exemple Copernic:

on peut télécharger ce produit à l'adresse suivante

http://www.copernic.com/fr

LES ASPIRATEURS

Aspirateurs de sites

Permettent de rapatrier le contenu d'un site pour en étudier le contenu localement. Permet une lecture off-line.

Ecatch

www.ecatch.com

Memo web

www.memoweb.com/fr/indexfr.htm

TelePort Pro

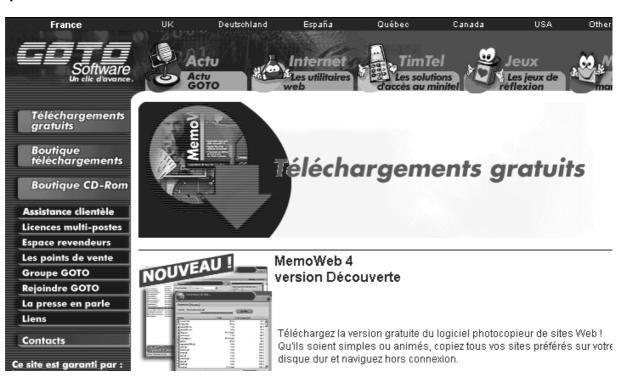
www.tenmax.com/pro.html

Webzip

www.spidersoft.com/webzip/home.htm

N'hésitez pas à consulter la page http://www.idf.net/info/table.html consacrée aux aspirateurs.

l'exemple de Memo web



RECHERCHE MULTIMEDIA

Principaux moteurs de recherche multimédia

Nom	Adresse
Alta vista	www.altavista.com
	www.altavista.fr
Lycos multimédia	http://multimedia.lycos.com/
Voila	www.voila.fr

Photos – images - cliparts

Nom	Adresse
Clipart	www.clipart.com
Go Graph	www.webgraphique.com
Ditto	www.ditto.com
Corbis (payant)	http://search.corbis.com
Photodisc (payant)	www.photodisc.com
Power Pics (payant)	www.powerpics.com

Cartes – plans - itinéraires

Université du Texas
www.lib.utexas.edu/Libs/PCL/Map_collection/Map_collection.html
Inter Carto
www.intercarto.prd.fr/index.htm
Map Quest
www.mapquest.com
Iti à <u>www.iti.fr</u>
PlanFax
www.planfax.com
Ismap
http://www.ismap.com/geo/lsmap.htm

RECHERCHE ADRESSES MAIL

Adresses électroniques & N° de téléphone

Enfin, plusieurs sites sont consacrés à la recherche de personnes ou d'entreprises (e-mails, adresses postales et numéros de téléphone).

Principaux annuaires

<u>x annuaires</u>	
Nom	Adresse
Annu Téléphone fr	www.annu.com
BigFoot Téléphone fr et email	fr.bigfoot.com
Yahoo People Search	people.yahoo.com
Tél. US et Email US	
Tél. FR et Email FR	fr.people.yahoo.com
Info Space	www.infospace.com
Tél. US et Email US	
Voilà	www.annuairemail.voila.fr
Tél. FR et Email FR + autres pays	
Switchboard	www.switchboard.com
Tél. US et Email US	
Lycos Network WhoWhere	whowhere.lycos.com
Tél. US et Email US	french.whowhere.lycos.com
Tél. FR et Email FR	
Four11	foreign.four11.com/
	ου
	people.yahoo.com
EuroInfo	http://www2.eu-
Téléphone et Annuaire inversé fr et Europe	info.com/france/

LISTE DE DIFFUSION - FORUMS

Liste de diffusion

Le CRU	www.cru.fr/listes/thematique .html
Nouvellement vendu → →	http://www.francopholistes.com/

Principe de base.

Les **listes de diffusion** permettent à un groupe de personnes de communiquer sur un thème donné par l'intermédiaire du **courrier électronique**. Chacun écrit à l'adresse électronique de la liste. Un automate se charge de distribuer les messages à tous les participants. Il est nécessaire de s'abonner à la liste pour pouvoir participer aux discussions.

Cas particulier.

Certaines listes de diffusion n'ont pas pour but des discussions mais juste l'émission d'une information (bulletin, magazine, revue de presse,...). Dans ce cas, l'abonné ne peut pas s'exprimer dans la liste.

Trouver une liste

Rechercher une liste par Nouveauté F	<u> Périodique Catégorie Nom</u>		
Entrez un ou plusieurs mots-clés pour trouver une liste.			
	Trouver		
Recherchez une liste			
Catégories			
• Arts et Culture	 Biens de consommation 		
Architecture, art populaire, folklore,	Habitat, bien-être, jouets, sexe,		
cuisine, gastronomie, musique,	véhicules, mode, beauté, jardinage,		
arts, cinéma, photographie,	fleurs, plein-air, voyance, astrologie,		
litérature, spectacles,	communication,		

Exemple de liste de diffusion consacrées aux mairies et collectivités locales



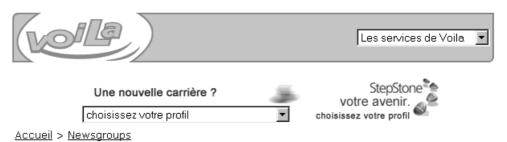
Existe-t-il des listes de diffusion parlant de recherche ?

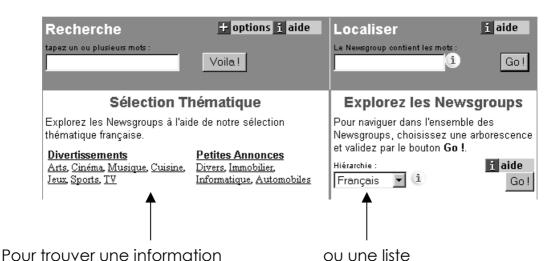
Il existe une excellente liste de diffusion francophone parlant des moteurs de recherche, MotRech. Pour en savoir plus, allez à l'adresse http://www.chez.com/jcharron/motrech/presentation.html

Forum

Quelques sites de recherche spécialisés dans les groupes de discussion

Voilà NewsGroup	news.voila.fr
DéjàNews	www.dejanews.com





Existe-t-il des forums de discussion parlant de recherche ?

En français, fr.network.internet qui est très général.

En anglais, **alt.internet.search** (mais il en existe de nombreux autres) est à votre disposition.

RECHERCHE DE LOGICIELS

Téléchargement

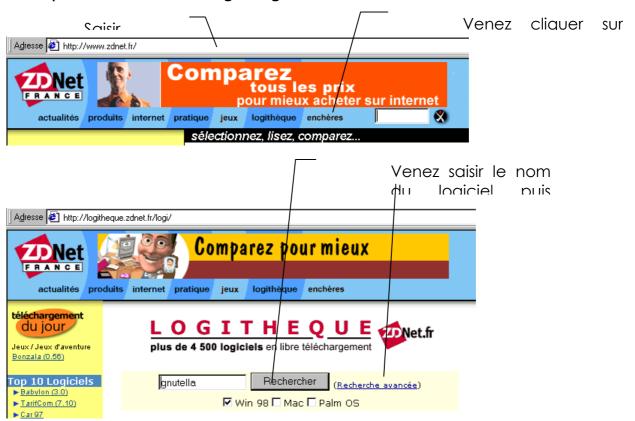
<u>www.tucows.com</u> <u>www.download.com</u>

<u>www.telecharger.fr</u> <u>www.telecharger.com</u>

www.winfiles.com www.anshare.com

www.zdnet.fr

Exemple sur zdnet.fr du logiciel gnutella



logithèque résultat

Bureautique Communication Développement Éducation Entreprise

Internet Jeux

Loisirs et famille Multimédia Utilitaires Palm

Freeware | téléchargé 148 fois | 🗖

Gnotella (0.70)

NEW!Recherche et partage de fichiers sur le Net.

W2K:Windows 95, Windows 98, Windows NT

Freeware | téléchargé 3226 fois | ***

Gnutella (.56)

Rechercher et télécharger des fichiers MP3.

Windows 95, Windows 98

Venez cliquer sur le nom

LES FICHIERS SUR SERVEURS FTP

Les sites suivants permettent de rechercher un logiciel par son nom, sa description ...

Ftp Find

Adresse: www.ftpfind.voila.fr

Recherche FTP de Voilà

File Mine

Adresse: www.filemine.com/

On donne un terme et on précise le système d'exploitation ou on choisit dans une catégorie

Filez

Adresse: www.filez.com

Ce site affirme signaler 75 millions de fichiers répartis sur 5000 serveurs. On peut limiter une recherche en précisant un nom de domaine.

Shareware.com

Adresse: http://shareware.cnet.com/

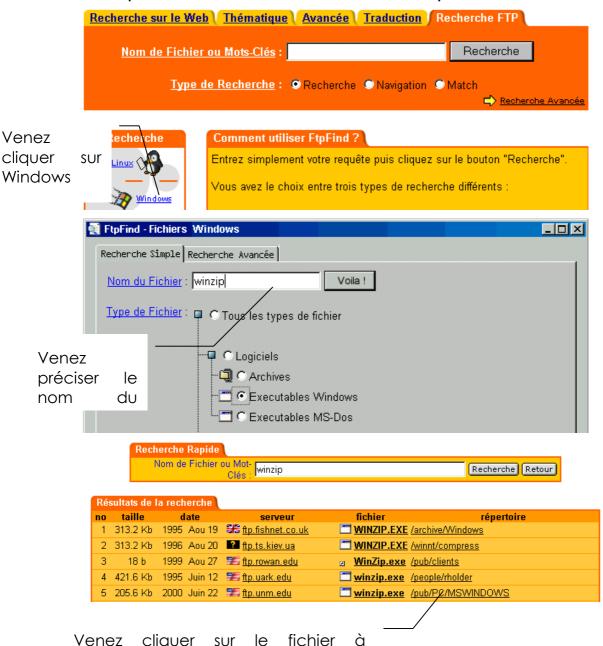
Après avoir indiqué le système d'exploitation concerné, on donne quelques termes combinés par Et et SAUF si nécessaire. On peut limiter par date. La recherche se fait dans le texte de description qui accompagne tout logiciel et dans le nom du produit.

TuCows

Adresse: http://www.tucows.asi.fr/

Ce site est une mine de logiciels permettant d'exploiter au mieux les ressources de l'Internet. De nombreux miroirs permettent un téléchargement sans problème.

Exemple de recherche: recherche de winzip



On vous propose ensuite d'enregistrer le fichier sur votre disque

Vous devez ensuite préciser dans quel dossier vous désirez enregistrer le fichier.