

ZONE ABONNÉS : Archives de la lettre "R&R"

Avril 2000

■ Yahoo! modifie légèrement la présentation de ses résultats. Si vous allez à l'adresse : http://dir.yahoo.com/Computers_and_Internet/News_and_Media/ Vous vous apercevrez rapidement que plusieurs grandes zones sont de plus en plus souvent présentes dans les pages des rubriques de Yahoo! :

- "Inside Yahoo!" propose les sites et les rubriques propres à Yahoo!, parmi la galaxie de sites qui lui appartiennent.
- "Categories" propose, comme avant, les sous-rubriques disponibles à l'intérieur de la rubrique consultée.
- "Most popular sites" est nouveau et intéressant : il s'agit des sites les plus populaires, mais sans que les critères de choix de ces rubriques ne soient explicités nulle part. Il semblerait que ce soit là le choix des net surfers de Yahoo! (comme la paire de lunettes qui caractérisent les sites à fort contenu pertinent). Il s'agit certainement d'une réplique à l'annuaire de Google qui classe ses sites par indice de popularité. Mais la question qui reste posée est : "quelle est la différence entre la paire de lunettes et la catégorie "Most popular sites" ?" On cherche, on cherche, et on vous tient au courant !

■ D'autre part, Yahoo.com a également modifié, dans ses pages de résultats lors de la saisie d'un mot clé, l'intitulé des rubriques : le lien n'est plus proposé que sur le dernier terme, alors qu'auparavant on pouvait cliquer sur toute la rubrique. La modification intervient certainement pour ne pas induire les internautes en erreur. En effet, auparavant, quelque soit l'endroit où l'on cliquait dans l'intitulé de la rubrique, on allait obligatoirement à la rubrique finale proposée. Or, certains internautes cliquaient à l'intérieur de l'adresse pour aller ailleurs dans l'arborescence de l'annuaire. Une aspirine ? Je m'explique : si Yahoo.com vous proposait, par exemple comme résultat à la requête "search engine" :

Business and Economy > Business to Business > Marketing and Advertising > Internet > Promotion > Search Engine Placement Improvement

Vous ne pouviez aller, par ce lien, qu'à la rubrique "Search Engine Placement Improvement" dans la rubrique Promotion, etc. Or, certaines personnes cliquaient sur les mots "Marketing and Advertising" pour aller directement à cette rubrique, ce qui n'était pas possible.

La nouvelle présentation des adresses n'induit plus en erreur.

■ Le Wap est la nouvelle coqueluche des outils de recherche. La plupart des outils proposent des solutions de recherche dans ce domaine :

FAST :

<http://wap.fast.no/>

Yahoo.co.uk :

<http://wap.yahoo.co.uk/>

Voila :

<http://awap.voila.fr/>

Excite :

<http://www.mobile.excite.com>

Yahoo! France :

<http://wap.yahoo.fr/>

Google :

<http://www.google.com/palm/>

■ Altavista clone Google

Altavista a lancé un nouveau moteur de recherche, nommé Raging Search, qui ressemble comme un frère - jumeau de surcroît - à Google. Si l'index utilisé est bien celui d'Altavista (les résultats sont strictement identiques sur les deux moteurs pour une même requête), la présentation est bien plus sobre et se rattache aux principes mis en avant par les nouveaux outils de recherche, notamment Google et All The Web : pas de publicité, interface la plus sobre possible, quasiment pas de graphique, les résultats, rien que les résultats (exit l'actualité, la météo et tous les atours des portails) et temps de réponse optimisé (Raging serait 20% plus rapide qu'Altavista). Altavista a également annoncé qu'il avait fait un nettoyage de son index pour éliminer le plus possible de pages obsolètes et générant des erreurs 404. Les revenus sur le site Raging devraient venir de liens présentés en bas de pages et vendus à des annonceurs. Une option de configuration permet de mettre en place un filtre familial, la possibilité d'afficher plus ou moins d'information dans les résultats, de prendre en compte une recherche linguistique ainsi que des possibilités de traduction en ligne déjà proposée par Altavista.

Disons qu'un peu plus d'originalité par rapport à Google aurait été le bienvenu (on frise quand même parfois le plagiat), mais cela montre bien que l'avenir des moteurs de recherche passe par la rapidité de l'obtention des résultats et l'arrêt de l'overdose d'information proposée sur les portails. Cela fait quelques années que Yahoo! l'a compris...

<http://www.raging.com/>

<http://www.ragingsearch.com/>

Configuration :

<http://ragingsearch.altavista.com/cgi-bin/query?pg=pref&v=m>

■ Voila a mis en ligne une version Beta de son nouveau moteur de recherche. Ce nouveau moteur (nom de code interne chez France Télécom : "KE") est prévu pour contenir jusqu'à 1,5 milliards d'URLs, voire plus. Il devrait "démarrer" d'ici l'été avec environ 300 millions d'adresses dans son index et 11 millions de pages en langue française. A la fin de l'année 2000, ce chiffre pourrait atteindre 400 à 500 millions. Le but n'est pas obligatoirement de recenser TOUTES les pages du Web mais de proposer un index solide souvent mis à jour. L'objectif de Voila en ce sens est ambitieux : indexation sous 24 heures des sites soumis au moteur, mise en place d'un algorithme de réindexation en fonction de la fréquence de mise à jour des pages (en d'autres termes, une page qui est rarement mise à jour n'est pas indexée souvent, une page d'actualité quotidienne sera "crawlée" tous les jours).

Dans le mode de recherche par défaut du nouveau moteur, les réponses qui correspondent exactement à l'expression demandée apparaissent maintenant en premier. Si vous tapez plusieurs mots, ce sont les documents qui comportent tous ces mots qui vont également apparaître en premier. Viennent ensuite ceux qui contiennent une partie seulement des mots. Une option "classer par popularité" permet de tenir compte de l'indice de popularité (le nombre de liens présent sur le Web vers les sites présentés dans les résultats de la requête, voir en fin d'article), comme sur la plupart des moteurs majeurs actuels. Le système analyse donc les liens croisés des réponses de la liste entre elles : plus un site est référencé par les autres, meilleure sera sa position dans la liste-réponse. Le nouveau système affiche également, en gras lorsqu'ils sont affichés, les mots que vous avez recherchés pour chacune des urls retrouvées.

Un nouveau procédé permet également d'accélérer l'affichage des 10.000 recherches les plus fréquentes. L'assistant de recherche avancée bénéficie d'une ergonomie simplifiée, et permet désormais d'effectuer des recherches dans des parties spécifiques de pages web : vous pouvez, par exemple, faire porter votre recherche sur le titre, ou les mots-clés des pages html . Une nouvelle aide vous indique enfin la syntaxe à employer, et décrit tous les paramètres à utiliser ainsi que ceux qui sont à venir prochainement. Les nouveaux opérateurs de recherche (outre les "classiques +, - et guillemets) sont les suivants :

text: Le ou les mots spécifiés doivent être contenus dans n'importe quel endroit de la page hormis les tags TITLE, A, IMAGE, META KEYWORD et META DESCRIPTION. Exemple : **text:bourse**.

title: Le ou les mots spécifiés doivent être contenus dans le titre (balise TITLE). Exemple : **title:publicité**

anchor: Le ou les mots spécifiés doivent être contenus dans le texte d'un lien html.
Exemple : **anchor:voila**

alt: Le ou les mots spécifiés doivent être contenus dans le texte associé à la propriété ALT d'une image. Exemple : **alt:eclipse**

keywords: Le ou les mots spécifiés doivent faire partie des mot-clés (META KEYWORDS).
Exemple : **keywords:commerce**

description: Le ou les mots spécifiés doivent être contenus dans la description du document (META DESCRIPTION). Exemple : **description:mode +paris**

url: Le ou les mots spécifiés doivent faire partie de l'URL de la page. Exemple : **url:voila**

domain: La page doit faire partie du domaine spécifié. Exemple : **domain:voila.fr**

link: La page doit contenir un lien vers l'URL spécifiée. Exemple : **link:www.voila.fr**

feature: Permet de rechercher des pages contenant un objet. Les valeurs sont les suivantes : image, sound, video et interactif (pour les fichiers javascript, java, flash). On doit combiner ce filtre avec une recherche de mots-clés. Exemple : **+feature:video +simpsons**

On peut également appliquer les opérateurs de recherche sur plusieurs mots à l'aide des parenthèses :

+title:(guide new york)

+(adaptations cinématographiques oeuvres littéraires) -(droits auteurs)

+"orange mécanique" -(film cinema)

Version beta de Voila :

<http://www.voila.fr/beta>

Pour tout savoir sur l'indice de popularité :

<http://www.abondance.com/docs/popularite.html>

■ Excite France cherche dans les vidéos et les images

La nouvelle recherche Audio-Vidéo d'Excite, fonction complémentaire du moteur, apporte à l'internaute des résultats qui ne contiennent que des fichiers audio et vidéo : images, photos, extraits de films, morceaux de musique, clips ou discours peuvent être trouvés grâce à ce moteur spécifique. Excite Audio-Vidéo offre également une fonctionnalité avancée qui permet à chaque internaute de chercher parmi tous les formats audio et vidéo existants ou bien de ne chercher que certains formats (AVI, MIDI, MPEG, MP3, Real, QuickTime ou WAV) en fonction des logiciels disponibles sur son ordinateur.

■ Altavista propose son moteur à ses clients

Le moteur Altavista va proposer sa technologie de moteur de recherche aux clients qui désire mettre en place un moteur de recherche sur leur site de commerce électronique, sous le nom de "AltaVista Search Engine 3.0 for e-business".

<http://solutions.altavista.com>

■ Voila licencie son moteur

Voila propose de "vendre" sa technologie de moteur et de "TSE" (Thematic Search Engines), sur le site "Voila Search Solutions". Plusieurs index sont disponibles, dans de nombreux langages.

<http://www.voilasearch.com/>

■ Ask Jeeves lance une gamme de produits de recherche

Ask Jeeves vient d'annoncer la disponibilité d'une gamme de trois logiciels, nommée "Webwide Navigation Services" basée sur la technologie DirectHit, rachetée il y a quelques semaines de cela. Les trois produits se nomment Jeeves Popularity Search (moteur de recherche), Jeeves Directory Search (annuaire basé à la fois sur l'Open Directory et sur DirectHit) et Jeeves Compare (comparaison de prix pour les sites de commerce électronique).

<http://www.ask.com/>

■ 500 millions de pages pour Inktomi

Inktomi a annoncé la disponibilité d'un nouvel index, nommé GEN3, de 500 millions de pages, avec la possibilité d'aller jusqu'à un milliard de documents. Rappelons quelques clients d'Inktomi : America Online, Goto.com, LookSmart, HotBot, MSN, Nomade et Yahoo! La solution serait disponible immédiatement.

■ Presenceweb rachète Francopholistes

L'excellent annuaire de listes de discussions francophones Francopholistes est passé dans les mains de la société Presenceweb qui a remporté l'appel d'offres du Comité Réseaux des Universités, créateur originel du site.

<http://www.francopholistes.com/>

■ ON EST TOUS GAGA DE GOOGLE !

Google vient de lancer ses versions internationales dont l'une permet d'effectuer des recherches uniquement en français. Une page de configuration permet de choisir parmi 11 langues celle(s) que l'on désire utiliser (système très comparable au "Language Settings" d'Altavista). Puis une page spécifique à votre demande, c'est-à-dire permettant d'effectuer des recherches uniquement dans les langues demandées, vous est proposée. Simple et efficace.

La page de configuration linguistique : http://www.google.com/lang_options.html

La page d'accueil du "Google français" : <http://www.google.com/intl/fr/>

Google propose également des fonctions de clustering. Cela signifie, que dans ses résultats, il n'indique plus que deux pages au maximum par site web. Pour obtenir les autres pages, il faut cliquer sur un lien intitulé "[More results from www.siteweb.com]". La plupart des moteurs de recherche majeurs proposent déjà cette fonctionnalité.

Il propose, pour chaque résultat, la fonction "Show Matches" qui présente la page indiquée, avec le mot recherché qui est surligné de couleur jaune. Une fonction "Email These Results" permet, quant à elle, d'envoyer la page de résultats à une personne que vous connaissez, par messagerie électronique.

Enfin, "pour le fun", allez jeter un coup d'oeil à la page qui présente de très nombreux logos, projets abandonnés ou réalisés à l'occasion d'un jour particulier (1er avril, Saint Valentin, Halloween, etc.) :

<http://www.google.com/logos.html>

■ Comment cherchent les internautes ?

La société américaine NPD a sorti début avril une étude qu'elle a l'habitude de mener tous les trimestres depuis 1997. Cette étude porte sur les habitudes des internautes lorsqu'ils recherchent de l'information en général, mais surtout sur leur utilisation des outils de recherche. En voici les principaux enseignements :

- 78% des internautes estiment qu'ils trouvent ce qu'ils cherchent sur l'Internet la plupart

du temps ou souvent. Ce chiffre pourrait paraître positif à première vue, pourtant, il est en baisse constante depuis 1997, date à laquelle il était de 86,6%. Dans le même temps, le pourcentage de ceux qui ne trouvent jamais est passé de 0,2% en 1997 à 3,3% en avril 2000 !

- Lorsqu'il ne trouve pas à la première tentative, l'internaute tente un deuxième essai sur le même site à 76,4%, plutôt que d'effectuer la même requête sur un autre outil de recherche (14,8% des personnes interrogées seulement). Cette tendance va plutôt en s'accroissant avec les années.

- Sur un moteur de recherche, l'internaute utilise en majorité plusieurs mots clés (40,5%), plutôt qu'un mot unique (30,5%). Il est à noter que la plus forte progression depuis trois ans revient aux requêtes en langage naturel, qui passent de 0% en 1997 à 11,6% en l'an 2000! Le site Ask Jeeves y est certainement pour quelque chose.

- Google est l'outil de recherche qui obtient la meilleure note avec 95% des internautes qui l'ont jugé pertinent pour leur recherche. Les outils "audités" étaient pour cette étude : Altavista, AOL Search, Ask Jeeves, Excite, Go (Infoseek), Google, GoTo, Hotbot, Looksmart, Lycos, Netscape search, Webcrawler et Yahoo! L'étude s'est faite en ligne à l'aide de fenêtre de type "pop-up" qui apparaissaient sur la page d'accueil des outils de recherche en octobre et novembre dernier. 40 000 personnes ont ainsi répondu au questionnaire proposé.

Notons, pour compléter ces indications, que sur des annuaires comme Yahoo! france ou Nomade, il y a autant d'internautes qui effectuent des recherches directement par mots clés que d'internautes qui suivent les liens proposés pour suivre l'arborescence de l'annuaire en "descendant les rubriques" pour aller à l'information recherchée.

<http://www.npd.com/>