

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

Mars 2000

■ Google propose maintenant la fonction site: qui permet de restreindre une recherche au nom de machine. Par exemple : `site:www.abondance.com` ou `site:www.whitehouse.gov` ou `site:www.abondance.net`.

Exemple :

+google + `site:www.abondance.com`, soit la requête :

[http://www.google.com/search?](http://www.google.com/search?q=%2Bgoogle+%2Bsite%3Awww.abondance.com&num=10)

[q=%2Bgoogle+%2Bsite%3Awww.abondance.com&num=10](http://www.google.com/search?q=%2Bgoogle+%2Bsite%3Awww.abondance.com&num=10) Attention, il semble que la requête "`site:abondance.com`" ne fonctionne pas. Il faut bien saisir "`site:www.abondance.com`".

■ Yahoo propose ses "related searches", à l'instar d'Altavista, de Hotbot et d'Excite. Lors de la saisie d'un mot, il propose des couples de termes contenant celui proposé, en bas d'écran, au dessus du formulaire de saisie. Testez par exemple sur "microsoft" :

<http://search.yahoo.com/bin/search?p=microsoft>

■ 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.

■ L'annuaire de Google semble être sorti de sa version Beta et est désormais pleinement opérationnel. Quelques infos supplémentaires sur ce nouvel outil de recherche :

- Les informations de type "annuaire" sont fournies par l'Open Directory (1,5 millions de sites recensés).

- Lors de la navigation au travers des rubriques en cliquant sur les liens proposés, l'annuaire propose un certain nombre de sites. A gauche de ces sites, deux barres horizontales, l'une verte et l'autre grise, sont présentées. Ces barres sont déterminées par la technologie "PageRank" de Google, pour déterminer la pertinence du site proposé sur le sujet. En fait, il s'agit de la mesure de l'indice de popularité, comme pour la partie "moteur de recherche" de l'outil. Les sites ne sont donc pas classés à l'intérieur d'une rubrique par ordre alphabétique, comme sur l'Open Directory ou tout annuaire "classique". Le système est assez génial, il faut bien l'avouer, est constitue une brèche dans le trivial classement par ordre alphabétique retenu par tous les annuaires jusqu'à maintenant. Enfin un classement par pertinence ou popularité à l'intérieur des rubriques d'un annuaire ! BRAVO !!

- Lors d'une recherche par mots clés, Google propose en premier les rubriques qui semble correspondre à la demande (mais ne contenant pas obligatoirement le(s) mot(s) demandé(s)), puis les sites, et il mixe à ce moment-là les résultats de l'annuaire et du moteur.

- En fait, il semble que les résultats lors de la saisie d'un mot clé dans le formulaire proposé donne le même résultat, que l'on fasse la requête sur Google "version annuaire" et sur Google "version moteur". Une preuve de plus que la grande tendance des outils de recherche est de mixer les résultats des deux types d'outils de recherche, avec une préférence marquée aux annuaires.

Google "version moteur" :

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

Google "version annuaire" :

<http://directory.google.com/>

L'Open Directory :

<http://www.dmoz.org/>

■ Le site Caloweb vient de faire son apparition dans le paysage Internet Français. Ce nouveau moteur de recherche utilise la technologie de Google. On attend cependant encore une confirmation officielle de leur part. Dommage, dans une interview donnée au journal "01 Informatique", le responsable de Caloweb précise que "près d'un milliard de pages Web sont désormais disponibles sur Internet; notre moteur de recherche Caloweb permet de toutes les indexer grâce à la technologie Google", ce qui est évidemment faux (le plus grand index du Web est celui d'Alltheweb avec 300 millions de documents) et ternit un peu le lancement de cet outil de recherche, qui reste à suivre malgré tout, même s'il n'apporte rien de véritablement neuf par rapport à sa version américaine. On aurait surtout apprécié que Caloweb donne une préférence aux pages en français et pallie ainsi un des inconvénients majeurs de Google, qui est l'absence de filtre linguistique. Il n'en n'est rien, les résultats proposés par Caloweb et Google pour une même requête sont en fait parfaitement identiques. Caloweb n'apporte donc qu'une simple interface en français

par rapport au moteur Américain. Un, peu décevant, il faut bien le dire...
<http://www.caloweb.com/>

■ Ecila est actuellement à la recherche de nouveaux développeurs pour l'accompagner dans de nouveaux projets, après une période de léthargie qui a duré de nombreux mois. Que s'est-il passé ? En fait, Ecila, moteur de recherche francophone fondé par Loïc Dachary en 1995, a commencé à se rendre compte l'été dernier que les technologies qu'il utilisait commençaient à montrer leurs limites. Aussi, les responsables techniques ont-ils décidé de concevoir une nouvelle mouture de leur solution de crawl-indexation de pages, notamment en suivant des principes d'"Open Source" sur Linux. Ce travail arrive en ce moment dans sa phase finale. Une nouvelle version du moteur devrait donc voir le jour dans les semaines qui viennent, certainement avant l'été.

■ La version américaine d'Altavista propose depuis peu des "guides de recherche" sur des sujets très divers : comment préparer ses vacances, généalogie, santé, achat de voiture, jardinage, rénovation de maison, etc. Très utile et très bien fait !
http://doc.altavista.com/search_guide/index.html