

## ■ Google Groups : Deja à la sauce Google

Cet article a été écrit par Béatrice Foenix-Riou, qui publie également l'excellente revue papier "Netsources" (<http://www.fla-consultants.fr/>). Merci à elle. Rappelons qu'en tant qu'abonné au site Abondance, vous avez droit à une remise de 15% sur l'abonnement à cette revue. Béatrice Foenix-Riou propose également l'envoi d'un numéro gratuit si vous le désirez. Pour plus d'infos : <http://boutique.abondance.com/>

### Présentation de Google Groups

-----

Les adeptes des forums de discussion l'auront déjà remarqué : Deja.com, qui donnait accès aux archives des messages échangés sur les forums, n'est plus. Les archives ont été rachetées par Google en février 2001 et un nouvel outil de recherche est désormais disponible à l'adresse : <http://groups.google.com/>

Ce rachat a d'ailleurs suscité de vigoureuses protestations parmi les internautes : contrairement aux critères de recherche sophistiqués disponibles sur Deja, la première version de Google ne disposait que de possibilités basiques et, surtout, ne donnait accès qu'aux messages publiés depuis moins de six mois. Impossible également de poster un message depuis Google (cette fonctionnalité est redevenue possible depuis quelques jours seulement) ; quant aux profils que pouvait enregistrer l'internaute sur Deja, ils avaient tout bonnement disparu.

Des articles véhéments ont alors fleuri sur le Web, certains n'hésitant pas à prôner la révolte contre la reprise des archives de Usenet par une entreprise commerciale, alors qu'il eût été plus logique qu'une telle collection soit détenue par un organisme public comme la Bibliothèque du Congrès des Etats-Unis. Voir en particulier "Deja `Revolt' against Google", Michelle Delio, Feb. 21, 2001, (<http://www.wired.com/news/print/0,1294,41926,00.html>).

C'est sans doute ce mouvement qui a incité Google à accélérer l'enrichissement de sa version beta. Les possibilités de recherche ont d'abord été améliorées pour se rapprocher de celles de la version d'origine, même s'il manque encore des outils comme la troncature ou l'opérateur de proximité. Mais surtout, Google Groups donne accès depuis peu à l'intégralité des archives de Usenet depuis le 29 mars 1995. D'autre part, Google a complété la base de données achetée à Deja - 500 millions de messages - par sa propre compilation de messages, créée depuis août 2000, ce qui représente rien moins que ... 650 millions de messages, échangés sur plusieurs dizaines de milliers de forums !

Deux possibilités de recherche sont proposées :

- La consultation par rubriques successives, à partir d'un classement arborescent des forums. L'écran d'accueil liste les grandes hiérarchies (biz., comp., humanities., news...) et donne pour chacune des précisions sur le contenu ; en cliquant sur une hiérarchie, on obtient la liste des subdivisions (biz.ecommerce, biz.marketplace...), puis celle des forums, avec les titres des derniers sujets de discussion ; en cliquant sur le titre d'un "thread", on affiche l'ensemble des messages (questions-réponses) liés entre eux.

Bien que cela ne soit pas précisé, il semble cependant que le classement thématique ne concerne aujourd'hui qu'une sélection de forums.

L'option "Browse complete list of groups" pour sa part donne accès à la liste alphabétique des forums de discussion ;

- Une recherche par mots sur le contenu des messages est également offerte, depuis l'écran d'accueil ou à partir d'une grille de recherche avancée (choix "Advanced Groups Search"). Cette dernière permet de limiter la requête à différents champs du message : nom d'un forum ou d'une hiérarchie de forums, auteur, sujet, numéro d'identification, mais aussi langue du message (dans une liste de 26) et/ou date ou période de publication. L'algorithme de pertinence utilisé pour classer les messages est dérivé de celui pris en compte pour le moteur de recherche Google (pages web).

Quel que soit le mode de recherche, on aboutit en fin de course au texte des messages publiés, avec pour chacun le nom et l'e-mail de l'émetteur, le sujet, le forum, la date et le texte du message. D'autre part, en cliquant sur le nom de l'émetteur, on obtient la liste de tous les messages qu'il a postés depuis 1995 ! Cette option peut s'avérer fort utile lorsqu'il s'agit de suivre ce qui se dit sur une entreprise, ou encore d'identifier un expert ; mais il ne faut pas l'oublier lorsque l'on participe à des discussions : attention, tout ce que vous dites est "enregistré"...

Au total, certaines des fonctionnalités de Deja font encore défaut (troncature, proximité). Mais en l'état, Google Groups représente la mémoire des discussions sur le Net ; il reste à ce titre incontournable. Avant de se lancer sur le Web pour trouver la réponse à une question, il est toujours utile de vérifier que celle-ci n'a pas été posée dans les forums, et que leurs archives ne contiennent pas déjà les réponses pertinentes de professionnels !

### Mode d'emploi

-----

- Domaine d'exploration : Archives des messages postés dans les forums depuis 1995, soit 650 millions de messages.

- Opérateurs booléens : Oui. En mode recherche simple (écran d'accueil), l'opérateur AND est utilisé par défaut ; il est donc inutile de le saisir. Sinon, le moteur comprend l'opérateur OR (en majuscules), et le symbole - pour l'exclusion (AND NOT). Les parenthèses ne sont pas comprises (contrairement à Deja). En mode recherche avancée, la grille permet de préciser que la requête doit se faire avec tous les mots (AND), un des mots (OR), ou à l'exclusion d'un mot (AND NOT).

- Opérateurs de proximité : Non, contrairement à Deja.

- Mots composés / phrase : Oui, il faut saisir les termes "entre guillemets".

- Recherche sur champs : Oui. En mode recherche avancée, il est possible de faire une recherche sur certains champs, dans une grille spécifique : auteur, sujet, nom (ou partie du nom) du forum, numéro d'identification du message, langue (dans une liste de 26), date ou période d'émission. En mode recherche simple, certains opérateurs peuvent être utilisés directement :

- \* group: recherche dans les messages d'un forum particulier ; on peut élargir à l'ensemble des forums d'une hiérarchie en utilisant la troncature à droite (Ex.: group:fr.reseaux\*) ou au contraire limiter aux forums d'une catégorie particulière en utilisant la troncature à gauche (Ex.: group:\*forsale)

- \* insubject: recherche dans les titres des messages

- \* author: (suivi du nom ou de l'adresse e-mail d'un auteur) recherche des articles d'un auteur (Ex.: author:john\_smith@dejanews.com).

- Troncature : Non (contrairement à Deja). Pour une recherche exhaustive, il faut donc comparer les résultats obtenus en écrivant les mots au singulier, au pluriel...

- Ordre des mots : Une recherche avec deux termes donne les mêmes résultats quel que soit l'ordre des mots.

- Caractères admis : Interprète indifféremment les majuscules et les minuscules. En revanche, l'accentuation est prise en compte par le moteur. Pour une recherche sur un mot accentué, il est donc prudent de comparer les résultats en écrivant le mot avec et sans les accents.

- Critères de classement : A l'issue d'une recherche par mots, les messages sont classés par pertinence, mais la liste des résultats offre le choix d'un classement par date. A l'issue d'une recherche par arborescence, les résultats sont classés par ordre ante-chronologique.

- Format de visualisation : A l'issue d'une recherche par mots, le système affiche la liste des forums répondant le mieux à la question, suivie de la liste des messages classés par pertinence. Pour chaque message, on trouve l'indication du sujet, un extrait comportant les mots de la requête (surlignés), le nom du forum, la date, le nom de l'auteur et un lien vers l'ensemble des questions-réponses (thread). A l'issue d'une consultation par arborescence, on obtient en bout de course la liste des derniers sujets postés dans le forum, avec la date d'envoi, le titre du thread (avec un lien vers les messages), le nombre d'articles et le nom du dernier auteur.

Quel que soit le mode de recherche, on obtient pour chaque message le nom et l'e-mail de l'émetteur, le sujet, le forum, la date et le texte du message. En cliquant sur le nom de l'émetteur, on obtient la liste de tous les messages qu'il a postés, depuis 1995.

- Paramétrage de l'affichage : En mode recherche avancée, il est possible de choisir le nombre de réponses par page : 10 (option par défaut), 20, 30, 50 ou 100.

Pour en savoir plus :

Google Groups :  
<http://groups.google.com>

Style Guide for Posting to Usenet :  
[http://groups.google.com/googlegroups/posting\\_style.html](http://groups.google.com/googlegroups/posting_style.html)

FAQ (including instructions on removing posts) :  
[http://groups.google.com/googlegroups/posting\\_faq.html](http://groups.google.com/googlegroups/posting_faq.html)

## ■ Deux questions aux annuaires francophones

Une rubrique qui revient périodiquement dans la lettre "Recherche et Référencement" : il s'agit de poser une ou deux question(s), mais à tous les outils de recherche majeurs du Web francophone, qu'il s'agisse des annuaires ou des moteurs, selon la question. Si vous avez des interrogations générales qui pourraient s'intégrer dans cette rubrique, n'hésitez pas ;-)

Ce mois-ci :

- **Question 1** : "Lorsqu'on a soumis un site à la date D sur l'annuaire A, combien faut-il attendre de temps avant de vérifier si le site est inscrit sur l'outil de recherche ?"

- **Question 2** : "Si le site n'est pas inscrit à cette date, quelle est la procédure à envisager ? Un mail ? Une nouvelle soumission ?"

Voici donc les réponses des annuaires à ces deux questions :

### **Agnès Bedel (Guide de Voilà) :**

Question 1 : " Le délai d'inscription est normalement de 2 mois. Pour cause de travaux sur la base et sur les flux de traitements des sites, nous ne prenons actuellement les inscriptions qu'au compte-goutte. Les travaux seront terminés à la fin du mois de mai, et les inscriptions reprendront en juin avec les mêmes délais.

A partir de juin, des nouveautés sont prévues pour optimiser le système de référencement, notamment avec l'envoi de mails à l'utilisateur pour l'informer des différents étapes de prise en compte de sa soumission."

Question 2 : " Il n'est pas utile de soumettre 2 fois le même site. Des réclamations peuvent être envoyées par mail, cependant, à l'adresse [webmaster@fr.voilà.net](mailto:webmaster@fr.voilà.net). Dans le prochain processus de soumission, il n'est plus possible de soumettre 2 fois une même url, et ceci pour éviter que des doublons s'accumulent dans la base des soumissions en attente. Si l'utilisateur soumet 2 fois une même url, on l'informerait par mail que cette url est en attente de traitement. Quand son url sera prise en compte, il en sera également informé. "

### **Jérôme Hérard (L'Ile des Médias, Directeur Général Adjoint) :**

Question 1 : "Le temps normal de prise en compte d'un site est 1 mois."

Question 2 : "Attendre 15 jours et soumettre à nouveau le site. A la deuxième tentative, inutile d'insister plus si le site a été refusé..."

### **Michelle Carrupt (Open Directory, éditeur ODP France) :**

Question 1 : "Il n'y a pas vraiment de limite "officielle", les délais sont variables. Si vous avez soumis votre site dans une catégorie sans éditeur, il pourra attendre plusieurs mois, voire même n'être jamais traité. L'ODP francophone manque cruellement d'éditeurs actifs. Si vous soumettez votre site dans une grande catégorie avec un éditeur très peu actif, c'est la même chose."

Question 2 : "Si votre site n'est pas traité dans un délai d'environ 4 semaines, vous pouvez :

- le soumettre à nouveau dans la catégorie d'un niveau supérieur ayant un éditeur.
- envoyer un e-mail à un éditeur d'une catégorie parente. Le resoumettre est en général peu concluant, la nouvelle soumission rejoignant simplement la précédente ! Un mail à un des éditeurs du top World/Français peut être concluant. Cliquez sur le nom de l'éditeur, et ensuite sur le lien « écrire à ». Celui-ci vous répondra, ou fera en sorte que votre site soit traité par l'éditeur de la catégorie.
- devenir éditeur vous-même en cliquant sur le lien "devenir éditeur" de la catégorie dans laquelle vous avez soumis votre site. Vous n'avez cependant \*aucune chance\* d'être accepté pour une catégorie avec un grand nombre de sites (plus de 50 sites). Choisissez une petite catégorie, avec peu de sites, et soignez votre demande. (Lisez attentivement le formulaire)."

### **Guillaume Le Friant (MSN.fr, Responsable outil de recherche) :**

Question 1 : "Les sites qui nous sont soumis sont généralement évalués entre 2 à 3 semaine en moyenne".

Question 2 : "Les utilisateurs peuvent nous soumettre gratuitement leurs sites en utilisant le formulaire dispo à l'url suivante :

<http://www.msn.fr/leguide/suggestions.asp> : Ce lien est dispo sur pratiquement toutes les pages de MSN Search. Les sites soumis par des robots sont automatiquement écartés de l'évaluation. Un site déjà sélectionné (ou non) peut nous faire part de ses modifications majeures en utilisant de nouveau ce formulaire. Les sites soumis sont évalués sur la base des critères disponibles à l'url <http://www.msn.fr/services/a2/aideguide.asp>. Nous n'assurons pas de retour vers les soumetteurs pour leur signaler leur éventuelle intégration ou les raisons de leur non-intégration dans notre sélection."

### **Thomas Chaumont (Yahoo ! France, Lead Surfer) :**

Question 1 : "Les utilisateurs désirant soumettre un site à Yahoo! doivent remplir un formulaire en ligne. Ces demandes sont ensuite examinées, dans un délai variable qui dépend de notre charge de travail, par notre équipe de surfeurs, qui visitent chaque site et le classent dans la ou les catégories appropriées de l'annuaire."

Question 2 : "Si un site est refusé, le surfeur de Yahoo! qui l'a visité envoie un courrier électronique à la personne l'ayant soumis, l'informant de la raison précise du refus. Ainsi, si un site est refusé parce qu'il contient des liens ne fonctionnant pas, il suffit de réparer ces liens. Si la version française d'un site multilingue est incomplète, il suffit de la finir. Une fois le site fini et en état de marche, il faut alors le resoumettre. C'est pourquoi nous conseillons aux utilisateurs de ne jamais soumettre un site avant de l'avoir fini et testé sur plusieurs navigateurs différents. Les délais de traitement des demandes variant selon notre charge de travail, il est possible que certaines personnes s'impatientent. C'est tout à fait normal et nous y sommes habitués ;-). Dans ce cas, il faut envoyer un e-mail à l'adresse [webmaster@fr.yahoo-inc.com](mailto:webmaster@fr.yahoo-inc.com), en rappelant la date de la demande de référencement (sans cette information, nous considérons que le site n'a jamais été soumis par la voie normale). Normalement, ça débloque les choses !"

### **Marie-Claude Belier (Nomade.fr, Cyberdocumentaliste) :**

Question 1 : "Dans l'éventualité d'un référencement dès le lendemain de la soumission, il faut compter une semaine au maximum pour que le site

soit visible en ligne ET accessible à la recherche par mots-clés. En effet, il faut compter de 2 à 4 jours après acceptation et traitement du site pour que celui-ci soit publié dans l'annuaire. C'est le temps de traitement qui est variable. Il est en moyenne d'une quinzaine de jours en cas "d'affluence".

Question 2 : "Il ne vaut mieux pas envisager une nouvelle soumission car elle risque d'être interprétée comme une forme de spamming, mais plutôt, envoyer un mail (à l'adresse [Pbajouter@sa.nomade.fr](mailto:Pbajouter@sa.nomade.fr)) à l'équipe de référencement qui suit au cas par cas les inscriptions et qui est à même de dire s'il s'agit d'un simple retard, d'un refus ou d'un problème technique.

En général, il suffit d'être patient, car un mail est envoyé systématiquement à chaque éditeur pour l'informer du suivi de son référencement : c'est-à-dire, si le site est accepté (avec son classement dans l'annuaire) ou refusé et pourquoi. Nous sommes aussi à l'écoute de toutes les formes de réclamations et suggestions :)"

#### **Marc Andouche (Lycos.fr, Lead surfer) :**

Question 1 : "Nous répondons aux soumissions par un mail d'accord ou de refus. A l'heure actuelle, il faut compter en moyenne 1 mois avant que les soumissions soient traitées sachant que nous traitons en priorité les sites d'information et de contenus gratuits par rapport aux sites de société et qu'une soumission pour un site qui n'indique pas la bonne catégorie peut prendre encore plus de temps. (Ainsi si vous voulez référencer le site d'une société, assurez-vous que vous le soumettez pour une catégorie commençant par « Lycos Home > Économie, Entreprise > Sociétés » ). Inutile de vérifier sur l'annuaire tant que vous n'avez pas reçu un mail."

Question 2 : "Si au bout d'1 mois et demi, vous n'avez pas reçu de réponse, contactez le websurfeur en charge de la catégorie dans laquelle vous vouliez être référencé ; vous trouverez son adresse mail sur la page suivante : <http://www.lycos.fr/service/contact.html>. Indiquez l'adresse de votre site et la catégorie dans laquelle vous vouliez vous trouver. Nous espérons que vous saurez rester courtois, vous n'en serez que mieux traité ; il arrive à nos websurfeurs de prendre du retard devant la masse de travail qui leur incombe. Une FAQ sur ce sujet est également disponible à l'adresse <http://www.lycos.fr/help/>."

#### **Emmanuelle Siary (Looksmart France, Directrice Business Development) :**

Question 1 : "nous confirmons l'acceptation ou refus du site sous 5 jours. Pour l'inclusion, nos mise à jour de la base se fait en général entre les 2 et 4 semaines pour les sites que nous hébergeons. Ceci peut prendre plus de temps pour les sites que nous n'hébergeons pas comme Altavista."

Question 2 : "s'informer via notre service client [whereisit-fr@looksmart.net](mailto:whereisit-fr@looksmart.net) qui répondra au plus vite à ce type de demande".

#### **Bernard Gay (Ctrouvé , Directeur) :**

Question 1 : "Actuellement l'inscription est arrêtée. A partir de juin : compter 15 jours pour un référencement gratuit, 24 heures pour un référencement payant. Immédiat pour une modification."

Question 2 : "Pas encore déterminé complètement. On se tourne vers un contrôle étape par étape pour le webmaster du traitement de son site (ce qui est déjà un peu le cas). Il pourra ainsi savoir où en est sa soumission et savoir si elle a été éventuellement rejetée."

Merci à tous pour ces réponses !!

#### **■ Le projet "Britney" d'Altavista Europe**


Les gens d'Altavista France et moi-même avons essayé (avec moult difficultés il faut bien le dire, et cela me permet de remercier le staff d'Altavista France qui a tenté de passer par les fenêtres pour obtenir des infos ;-)) d'en savoir plus sur le projet "Britney". Rappelons que cet outil a tout d'abord été mis en place sur Altavista Grande-Bretagne : lorsque vous saisissez, des mots clés comme "ferrari" ou "bmw", par exemple, vous voyez apparaître, dans les pages de résultats, une image qui vient mettre en valeur certains liens proposés.

L'image est, bien entendu, extraite de la page en question. Ce projet est également disponible depuis peu sur la version française du moteur. Tapez un mot clé comme "Renault" ou "Peugeot" et vous verrez apparaître une image, issues des pages web proposées, qui met en valeur certains liens affichés comme résultat. Le choix de l'image n'est pas calculé en temps réel car certaines pages qui aboutissent sur une erreur 404 se voient quand même mises en valeur par une vignette.

Le système n'est pas neutre, car il consiste en une mise en valeur du lien et donc à une incitation plus forte à cliquer dessus pour l'utilisateur du moteur. Nul doute que les référenceurs malins vont se pencher sur la question pour réaliser des pages compatibles avec ce nouveau système :-)

Voici quelques informations que nous avons pu glaner. J'essaie d'en savoir plus pour le mois prochain, mais cela semble assez difficile. Mais on a la foi ;-)

Les critères pris en compte pour afficher une image seraient les suivants :

- Présence d'une image dans la page (ben oui, quand même ;-))
- Format Jpeg quasi obligatoire, même si, dans certains cas très rares, on a pu déceler quelques images au format Gif.
- A priori, l'image contient une relation par rapport au mot clé demandé, et ce dans trois cas différents : soit dans le nom de l'image elle-même (pour le mot clé "bmw", l'image s'appellera [bmw.jpg](#) ou [logo-bmw.jpg](#)), soit dans l'option ALT de la balise IMG (exemple : ) soit dans un texte visible très proche de l'image.
- Image en couleur obligatoirement et de taille inférieure à 40 Ko.
- Eventuellement, l'image doit plutôt se trouver plutôt en haut du code HTML de la page (information à confirmer).

Il semble également qu'une solution de paiement pour afficher cette image (sorte de référencement payant) pourrait être envisagée à moyen terme par Altavista. Affaire à suivre...

D'autre part, voici un petit point sur les projets immédiats d'Altavista Europe et France :

- Le projet "Gladiator" est en train d'être mis en oeuvre. Il s'agit de la refonte du back office (noyau et serveur) du moteur qui permettra d'obtenir des temps de réponse comparables à ceux du site américain. A environnement graphique plus lourd que Google, les temps de réponses obtenus sont, du coup, très proches de ceux de ce moteur, très connu pour être parmi les plus rapides du marché. La mise en production devrait être effective au mois de juin prochain.

- Référencement et positionnement payant : pas la priorité numéro un pour l'instant pour Altavista France. Il y aura sans doute des nouveautés dans ce domaine pour septembre 2001 mais rien de sûr encore. Le principe parfois évoqué en interne : préserver un système d'indexation gratuit mais ajouter un système payant qui garantisse un délai d'indexation (système proche de ce que propose Inktomi). Pas de système de positionnement payant envisagé, ni vente aux enchères des premières places de type Goto sur Altavista.com.

- Prise en compte des pages dynamiques (ASP, CFM, PHP, JSP), avec ou non des "?" dans les intitulés des urls : Les 4 extensions sont indexées. Les paramètres posent problème la plupart du temps. Bien qu'Altavista indexe ces pages, la pertinence ne suit pas. En d'autres termes, les documents ayant une adresse "exotique", bien que normalement présents dans l'index, ne sont que rarement trouvés sur une requête par mots clés.

- Soumission de pages : pas de système anti-spam et anti-logiciel de référencement automatique tel que celui que propose Altavista.com depuis quelques jours (code graphique à saisir dans le formulaire) pour l'instant. Ce n'est pas une priorité. Pas avant la rentrée de septembre en tout cas.

- Catégories de mots clés le plus souvent saisis sur Altavista France en ce moment :

\* Sexe (cette catégorie regroupe toutes les requêtes dont la finalité est d'atteindre des sites à caractère pornographique (on y arrive très vite!). Ainsi, des recherches comme voyeur, charme, britney spears ou gay y ont été incluses) : 25%

\* Téléchargement gratuits (cette rubrique regroupe toutes les demandes concernant les logiciels gratuits, du type : warez, mp3, crack, gamez, moretv, astalavista, napster, gratuit, hacking, winzip...) : 11%

\* Divertissement (concernant loftstory, toutes les formes d'écriture ont été prises en compte exceptées celles à caractère sexuel. Le divertissement concerne également toutes les requêtes qui touchent les secteurs comme : M6, skyrock, humour, horoscope, loto, manga, jeux, infos sur les célébrités (madonna par exemple), multimédia (vidéo, photo)) : 48% dont 32% pour le seul LoftStory

\* Utile (manger, dormir) (toutes les requêtes considérées comme pratiques y sont répertoriées. Exemples : emploi, sncf, traduction, voiture, aifrance... ) : 7%

\* Informations (actualités, recherches...) (il est difficile de savoir quelle est la finalité d'une requête comme "lycos" par exemple. Par conséquent, on retrouvera ici des recherches assez générales et variées telles : météo, pages jaunes, france, euro, bourse, adsl ou encore internet) : 7%

\* Communication (rencontre, chat...) (incluant les requêtes sms, chat et rencontre, cette rubrique semble être sous-représentée. En effet, ces services étant souvent proposés par des sites comme lycos ou voilà classés dans la rubrique information) : 2%

## ■ Des nouvelles de Kartoo

Voici quelques nouvelles du site Kartoo, le métamoteur à interface graphique dont il a été largement question dans les pages du site Abondance ces dernières semaines.

Suite au lancement de l'outil de recherche, le trafic généré au bout de 4 jours était celui que ses concepteurs espéraient atteindre 1 an après le lancement dans leur business plan. Le trafic actuel est supérieur à 3 millions de pages vues par mois et ce à 95% sur le web francophone. L'objectif est, à court terme, d'atteindre un trafic beaucoup plus élevé grâce notamment aux autres versions linguistiques qu'ils vont proposer afin d'internationaliser l'outil.

Prochaines étapes du développement de KartOO pour le grand public :

-----

-> Semaine 22 :

- Installation de nouveaux serveurs qui vont permettre de réduire le temps de chargement des cartes et des résultats associés (car à certaines périodes de trafic important, correspondant à 20 000 requêtes par heure, le temps de chargement est trop long).

- Lancement de KartOO Perso, programme d'affiliation qui permet de proposer gratuitement une case "rechercher avec KartOO" à des sites web (pages personnelles, sites de pme, d'associations, portails communautaires, etc...). L'installation se fera en ligne en remplissant un formulaire grâce auquel on peut choisir l'environnement graphique de "son" KartOO.

- Lancement de KartOO en version anglaise (début de l'internationalisation de la technologie).

- Lancement de KartOO à télécharger qui permet de :

\* Gagner beaucoup de temps dans le chargement de l'interface flash,

\* Visualiser les cartes en plein écran,

\* Enregistrer les cartes sur son PC : on pourra repartir de ces cartes quand on le souhaitera pour compléter sa recherche, ce qui évite de devoir toujours repartir à zéro quand on effectue une recherche d'information... Objectif : arriver à gérer la recherche d'information pour des thèmes récurrents et pouvoir partager ces informations avec des amis ou des collègues (échange de cartes). Ils se prennent pour des Pokemon, chez Kartoo ? ;-)

-> Fin juin : mise en ligne de la deuxième version de KartOO

- Lancement de la V2 en Java qui va permettre de proposer des cartes en 2 à 3 secondes au maximum, y compris pour un trafic très important !

-> Cet été :

- Lancement d'autres versions linguistiques (espagnol, portugais, allemand, brésilien,...).

- Lancement de KartOO thématiques (KartOO freeware & shareware....).

- Proposition de nouvelles interfaces cartographiques très différentes de celles proposées actuellement.

L'objectif, dans un moyen terme, est de développer de nouvelles innovations complétant la technologie déjà en ligne dans les domaines suivants :

\* La recherche d'information.

\* Le travail collaboratif et l'échange d'informations.

\* La gestion des connaissances ou KM.

\* La veille sur le Web.

-> Lancement de l'offre professionnelle (multi-supports, multi-canaux) :

-----

Les offres commerciales de Kartoo, à l'heure actuelle, sont les suivantes :

\* Solutions commercialisées actuellement :

Licence de la technologie + prestations de services spécifiques : \*\* KartOO Intranet (outil de recherche d'informations au sein d'un intranet d'entreprise).

\*\* KartOO Portail (outil de recherche de contenus au sein d'un portail web).

\*\* KartOO Bases de données (permet de rechercher des informations au sein de bases de données relationnelles complexes, pouvant contenir des millions de documents indexés de tous types de formats - word, excel, powerpoint, pdf, html, texte, authorware, flash, director, streaming audio & vidéo -).

\*\* KartOO Partner (commercialisation de la technologie de recherche en marque blanche).

\* Solutions qui seront développées à moyen terme (fin 2001- début 2002)

\*\* KartOO numérique (web TV).

\*\* KartOO Vigile (veille).

\*\* KartOO Knowledge (gestion des connaissances). etc.

Pour plus d'informations :

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

## ■ Divertissement

Un petit divertissement offert par Webhelp France sous la forme d'une discussion - véridique - entre un internaute et un "Webwizard" de Webhelp :

- Question de Germaine : "quand elle se glisse dans le poulailler c'est la panique. 7 lettres, mot commence par b "

- Un WebWizard sera à votre disposition dans moins de 3 minutes.

- Christelle : "Bienvenue chez Webhelp.fr et merci pour votre question, Germaine ! Veuillez me reformuler votre question de manière plus précise, s'il vous plaît."

Germaine : "nom d'un animal de 7 lettres commençant par b, la définition est ..quand elle se glisse dans le poulailler c'est la panique"

Christelle : "Merci. Veuillez patienter pendant que j'effectue votre recherche."

Germaine : "merci à vous"

Christelle : "En attendant, n'hésitez pas à visiter notre site ou surfer sur le web ; la réponse à votre question apparaîtra automatiquement sur votre écran."

Germaine : "merci beaucoup"

Christelle : "J'ai trouvé un site qui pourrait vous intéresser. Je vous l'envoie tout de suite. Je vous envoie le site sous forme d'adresse Internet. Pour l'afficher, merci de cliquer sur celle-ci. Pour nous retrouver, réduisez cette fenêtre. "

Christelle : <http://perso.libertysurf.fr/dleroy/BELETTE.html>

Germaine : "merci beaucoup vous êtes gentille"

Christelle : "Trouvez-vous ce site utile ?"

Germaine : "oui merci"

[Christelle - Déconnexion de l'utilisateur]

Les outils de recherche humains, ça sert à tout, non ? ;-)

## ■ Délais actuels des moteurs de recherche entre une soumission manuelle et la prise en compte effective d'une page

Au tableau d'honneur : les moteurs qui indexent de nouveaux sites assez rapidement (moins de 15 jours) : Altavista, Altavista France, Google, Voila, Fast (donc Lycos et Lycos France).

Les moteurs qui sont "moyens" (certains sites indexés de façon plus ou moins rapide, d'autres non pris en compte, bref des délais assez variables mais qui restent raisonnables) : Inktomi (donc Hotbot), Excite France.

Les moteurs à la peine ce mois-ci : Ecila, Excite.

## ■ Bruits et chuchotements

Une rubrique qui regroupe tous les bruits et rumeurs dans le (petit) monde des outils de recherche mondiaux et francophones. Rien n'est obligatoirement vérifié, mais toutes les infos sont données... de source sûre ;-)

-> Un bruit a couru sur le Web francophone sur la prochaine fermeture de l'annuaire de l'île des médias (annuaire d'Excite France notamment) en septembre prochain. Renseignements pris auprès de la société en question, il n'en serait rien. Ce bruit aurait même fâché quelques personnes au sein de l'île des médias.. ;-)

-> Les fonctionnalités "Britney" (ajout d'images émanant des pages web dans les résultats d'Altavista) seraient bientôt intégrés sur la version américaine de l'outil de recherche. A ce sujet, TOUS les sites d'Altavista (US et non-US) convergeraient, à moyen terme, vers la même interface (mais laquelle ? : l'américaine ou la non-américaine ? A priori, on pencherait plutôt vers la deuxième solution). Pour l'instant, il y a encore du boulot... ;-)

-> Dans la prochaine et nouvelle version du Guide de sites de Voila, il a été prévu l'envoi d'un mail au référenceur du site lors du traitement de la soumission pour lui indiquer la prise en compte ou non de sa source d'informations. Mise en place de cette nouvelle version courant juin normalement (voir article "Deux questions aux annuaires francophones" ci-dessus).

-> Sur l'Open Directory, comme sur de nombreux annuaires, un test est réalisé de façon automatique pour vérifier que le site soumis est disponible en ligne (pas d'erreur de type 404 ou autre). Ce test est encore effectué, une fois le site enregistré. Si une erreur est détectée lors de ce crawl, le site est mis dans un fichier tampon et retesté fréquemment. Si l'erreur persiste, une demande est faite auprès d'un éditeur pour qu'il teste manuellement le site afin de vérifier que l'erreur est légitime. La procédure complète peut prendre 2 à 3 semaines.

-> Looksmart (.com et .fr) serait en train de concocter une prestation de positionnement "au clic" sur son annuaire. La tarification serait en fonction du nombre de visites générées chaque mois depuis l'annuaire. Renseignements pris, une nouvelle offre publicitaire, très innovante dans le domaine, serait en train d'être mise au point, sans qu'on en sache plus. Lancement imminent cependant. Restez connectés...

## ■ En Bref

-> Le métamoteur Profusion (<http://www.profusion.com/>), en version Beta depuis plusieurs mois, est enfin lancé de manière officielle dans sa nouvelle mouture.

-> Sur Altavista.com, dans la section "Featured Sites", en début de page de résultats, il peut apparaître soit les résultats de GoTo, soit des liens émanant de sites déjà clients avec la régie publicitaire d'Altavista. Dans ce cas, la préférence sera donnée à ces derniers.

-> NetBooster annonce la nomination de Arnaud Defrenne au poste de Chief Technical Officer. Dans le cadre de cette fonction, Arnaud Defrenne sera responsable de tous les aspects technologiques et production de NetBooster. Arnaud Defrenne, 30 ans, était jusqu'à présent Directeur Technique au sein du Groupe Liberty Surf. De 1997 à 1999, aux côtés de Gilles Ghesquière et Jean Postaire, il participe à la création de Nomade.fr ainsi qu'à l'enrichissement des services proposés par l'annuaire. En 1999, lorsque Nomade est racheté par Liberty Surf, Arnaud Defrenne contribue à la mise en place de la plate-forme technique de la régie publicitaire du groupe et au développement du système d'information des portails. Fort d'un Magistère de Gestion à l'Ecole Supérieure des Affaires de Grenoble et d'un DESS Management des Systèmes d'information, Arnaud Defrenne a

débuté sa carrière dans l'informatique financière à la direction financière de Schneider puis au sein de la banque Barclays. Arnaud Defrenne rejoint la société de Jean-Pierre Eskenazi et Michel Fantin, avec pour mission de développer de nouvelles technologies et des outils destinés à renforcer la visibilité des sites sur Internet. Son rôle sera également de développer le système d'information cohérent avec ces nouvelles technologies ainsi qu'avec le développement international de NetBooster.

-> Ecila a temporairement interrompu la fonctionnalité de soumission d'urls sur son moteur de recherche. En effet, ils sont en cours de changement du parc de serveurs pour pouvoir passer à l'étape suivante (montée en puissance, nouvelles fonctionnalités, meilleure qualité de service, etc.). Pendant quelques temps, il sera donc impossible de faire des ajouts (mises à jour). Puis la fonctionnalité sera remise en service, en même temps qu'un crawl+indexation plus fréquent.

-> Looksmart France a interrompu le système de soumission gratuite pour les sites à but non lucratif. Mais il ne s'agit que d'une suspension temporaire, la fonctionnalité devrait revenir d'ici peu.

-> Excite France a déménagé. Nouvelle adresse : 14 Rue Cadet , 75009 Paris - France -> Lycos France également : Réseau Lycos France, 19 cité Voltaire, 75133 Paris Cedex 11

-> Alltheweb (Fast) affiche maintenant la balise Meta "Description" dans ses résultats, mais ne semble toujours pas prendre en compte ce champ dans son algorithme de pertinence.

