

Recherche et Référencement - Numéro 26 -- Avril 2002

-- Lettre d'actualité mensuelle sur la recherche d'information et le référencement de sites Web --

***Au sommaire ce mois-ci :***

- > [Comment référencer un site web réalisé avec des \*\*frames\*\* ou en \*\*Flash\*\*](#)
- > [Lancement du nouveau moteur d'\*\*AOL France\*\*](#)
- > [Lycos France : un point sur l'outil de recherche](#)
- > [Différences culturelles sur les \*\*mots clés\*\*](#)
- > [Kartoo : la V2 est pour bientôt](#)
- > [Bruits et chuchotements](#)
- > [En bref](#)
- > [Les nouveaux entrants dans l'annuaire des outils de recherche régionaux](#)
- > [Cherchez, Référenciez-vous](#) (nouveaux outils de recherche)
- > [Contenu](#) : sites proposant du contenu ou des fonctions intéressants
- > [Outils](#) : logiciels et sites Web qui aideront les webmasters et les chercheurs d'information dans leur travail quotidien
- > [Revue d'URL](#)

Le contenu de cette lettre est accessible sur la zone "Abonnés" du site Abondance, à l'adresse :

<http://abonnes.abondance.com/archives/acturech/0402.html>

Pour tout renseignement : © Olivier Andrieu, [oa@abondance.com](mailto:oa@abondance.com)

## Comment référencer un site web réalisé avec des frames ou en Flash

[Retour au sommaire de la lettre](#)

**Note** : cet article actualise et complète certains articles parus précédemment sur le site Abondance et concernant le sujet traité...

### Sites réalisés avec des frames

De façon assez générale, les frames sont très souvent considérées comme un réel obstacle pour les moteurs de recherche et donc pour le bon référencement d'un site qui prendrait en compte cette technique de subdivision de l'écran en différentes fenêtres indépendantes. Et... c'est vrai. Enfin, c'est souvent vrai, serait-on tenté de dire, car nous verrons plus loin que les frames peuvent également être utilisées pour obtenir un meilleur positionnement...

Avant de voir comment les moteurs réagissent lorsqu'ils arrivent sur une page ainsi bâties, il est nécessaire de dire deux mots de la réalisation de ce type de page en HTML. Imaginons une page web imaginaire qui aurait pour nom Frames.html et dont le code HTML serait le suivant :

```
<FRAMESET ROWS=20%,80%>
<FRAME SRC="FH.html" NAME="haut">
<FRAMESET COLS=*,2*>
<FRAME SRC="FGB.html" NAME="gauchebas">
<FRAME SRC="FDB.html" NAME="droitebas">
</FRAMESET>
<NOFRAMES>
Cette page a été réalisée avec des frames.
</NOFRAMES>
</FRAMESET>
```

La balise `<FRAMESET> ... </FRAMESET>` permet de définir les cadres qui constitueront l'écran du navigateur et d'indiquer quels fichiers HTML (ici FH.html, FGB.html et FDB.html) seront affichés à l'intérieur de ces zones (= cadres = frames). Notez bien que le fichier Frames.html, que nous appellerons "fichier mère", sert uniquement à la description des zones de découpage de l'écran et ne contient aucune indication sur le texte ou les images qui y seront affichées.

Les fichiers FH.html, FGB.html et FDB.html, que nous appellerons "fichiers filles", contiennent pour leur part les informations à afficher dans chaque partie d'écran. Le fichier Frames.html (fichier "mère") décrit donc la façon dont les fichiers "filles" seront affichés sur l'écran.

Mettons nous maintenant à la place du *spider* qui arrive sur une page, obligatoirement de type "mère" dans un premier temps. Il peut avoir quatre réactions différentes :

- \* Premièrement, il ignore complètement la page Web et ne l'indexe pas, car il a décidé (enfin, ses concepteurs ont décidé pour lui) de ne pas prendre en compte les pages avec frames. ce type de cas n'existe quasiment plus aujourd'hui sur le Web.

- \* Deuxièmement, il indexe les fichiers "mère" et "filles" et fonctionne exactement comme pour une page sans frames. La situation est idéale : si un mot clé est trouvé dans l'une des pages "filles", il peut revenir au fichier "mère" et affiche la page comme elle doit l'être normalement, restaurant ainsi le contexte d'affichage initial. Très peu de moteurs (et pour tout dire aucun) ne savent faire cela. Dommage...

- \* Troisièmement, le spider indexe uniquement la page "mère", et notamment le contenu de la balise `<NOFRAMES>...</NOFRAMES>` et ignore délibérément les fichiers "filles" en ne suivant pas les liens présents dans les balises `<FRAME>`. Bon nombre de moteurs fonctionnent ainsi. Et d'autres, censés bien prendre en compte les frames, n'en font souvent pas plus.

- \* Quatrièmement, le spider indexe les fichiers "mère" et "filles". Puis il les considère tous comme des pages Web distinctes, sans rapport les unes avec les autres. Si un mot clé est trouvé, par

exemple, dans la page FDB.html, le moteur proposera un lien direct vers ce document et non pas vers Frames.html. FDB.html s'affichera alors seul sur l'écran du navigateur. Le moteur n'a pas pu reconstituer le lien entre le fichier "fille" FDB.html et le "mère" Frames.html. Le contexte des frames est ainsi perdu. L'internaute, qui a cliqué sur un lien dans la page de résultats du moteur, se retrouve avec une page sortie de son contexte "framé". Pour tout dire, c'est assez gênant et, malheureusement, très courant sur les moteurs de recherche...

Conclusion simple, mais bien souvent irrémédiable : réfléchissez bien avant d'utiliser des frames dans vos pages. Leur contenu risque de ne pas être visible dans les moteurs qui, en règle générale, gèrent peu ou mal ces documents. Et comme il existe de moins en moins de site réalisés avec des frames (visualisez les cinquante premiers sites mondiaux en termes d'audience et vous verrez vite qu'aucun n'utilise cette technique, abandonnée même par son concepteur, Netscape, sur son site), il y a peu de chances que les moteurs fassent quoi que ce soit pour mieux les prendre en compte à l'avenir.

Cependant, si votre site est ainsi réalisé, il existe des solutions "miracles", grâce à la balise <NOFRAMES> et au Javascript, qui font que votre site, même réalisé avec des frames, sera bien pris en compte et affiché par les moteurs de recherche. Ouf...

Deux palliatifs peuvent atténuer le fait qu'un site est mal pris en compte lorsqu'il est bâti sur la base de frames : l'emploi, dans la page "mère", de bons titres et de balises <META>, qui s'avèrent ici absolument indispensables (même si ces balises sont aujourd'hui bien moins bien prises en compte que par le passé par les moteurs majeurs), et l'utilisation de la balise <NOFRAMES> ... </NOFRAMES>, qui permet d'indiquer un texte, à l'origine destiné aux navigateurs n'acceptant pas cette fonctionnalité.

Soignez le texte que vous allez indiquer ici, car il y a de fortes chances, si le moteur ne prend pas en compte les balises Meta, pour que seules ces lignes soient affichées dans la description du fichier "mère" sur la page de résultats du moteur. Si, comme dans l'exemple ci-dessus, votre page a été réalisée avec ce code :

```
<NOFRAMES>
Votre navigateur n'accepte pas les frames.
</NOFRAMES>
```

Cela donnera sur la page de résultats du moteur :

```
1. Chaussures de sport Stela
Votre navigateur n'accepte pas les frames.
http://www.stela.com/index.html - size 1K - 20-Mar-20002 -
```

On peut rêver mieux comme description de document, non ? Soignez donc les textes introduits dans la balise <NOFRAMES> !

Voici également une astuce qui devrait vous être d'un grand secours lors de la réalisation de vos pages : insérez un lien dans la partie <NOFRAMES> vers les documents "filles" qui affichent les liens de navigation interne à votre site.

Pour être plus explicite, reprenons la page de début, intitulée Frames.html, et adaptons ce fichier à une entreprise fictive appelée Stela. Le code de cette page est le suivant :

```
<FRAMESET ROWS=20%,80%>
<FRAME SRC="FH.html" NAME="haut">
<FRAMESET COLS=*,2*>
<FRAME SRC="FGB.html" NAME="gauchebas">
<FRAME SRC="FDB.html" NAME="droitebas">
</FRAMESET>
<NOFRAMES>
Cette page a &eacute;t&eacute; r&eacute;alis&eacute;e avec des frames.
</NOFRAMES>
</FRAMESET>
```

La zone <NOFRAMES> est alors remplie avec un texte de remplacement tout à fait commun. Modifions maintenant cette zone ainsi :

```
<NOFRAMES>
<A HREF="FH.html">Stela</a>, spécialiste de la vente de <A
HREF="FGB.html">chaussures de sport</a>, bas&eacute; &agrave; <A
HREF="FDB.html">Paris, France</A>.<BR>
</NOFRAMES>
```

Que se passe-t-il pour le robot qui, dans un grand nombre de cas, ne connaît que le contenu de cette balise <NOFRAMES> ? Il va indexer la description fournie (Stela, spécialiste... France), puis il va suivre les liens proposés vers les fichiers FH.html, FGB.html et FDB.html. Or, ces fichiers contiennent obligatoirement des liens vers les autres parties du site. Vous avez gagné. Le spider va alors visiter les pages importantes de votre site et les indexer. En revanche, elles seront enregistrées et visualisées ensuite sans les frames sous la forme de fichiers "filles", mais la situation est tout de même bien meilleure que précédemment, où votre site devenait souvent entièrement "transparent" pour le moteur de recherche.

Mais il y a également une façon de contourner le fait que les moteurs peuvent proposer un lien vers une page "fille" devenu "orpheline". Dans le code HTML de chacun de ces fichiers "filles" (dans la balise <HEAD>), insérez le code Javascript suivant :

```
<SCRIPT Language="javascript">
<!-- // Test d'affichage sans l'environnement frames
if (parent.frames.length==0) parent.location.href="pagemere.html"
// --> </SCRIPT>
```

Ce code ne permet pas l'affichage de la page "fille" seule et recrée obligatoirement l'environnement de la page avec des frames. Indiquez, à la place du nom pagemere.html, le nom de la page "mère" (ici : Frames.html, par exemple) correspondant à chaque page "fille" pour lesquelles vous allez insérer ce bout de code.

Bon, tout ça fait un peu "bidouille" et représente un travail qui n'est pas négligeable - et que vous n'aviez pas obligatoirement planifié -, mais nécessaire si vous désirez que votre site réalisé avec des frames soit bien référencé.

**Point important** : tenez compte des mots et expressions importants pour insérer des liens dans la balise <NOFRAMES>. Dans l'exemple précédent, nous avons choisi les mots Stela, chaussures de sport et Paris, France pour y insérer les liens. Pourquoi ? Parce que les moteurs privilégient les termes situés entre les balises <A> et </A> dans leurs calculs de pertinence. Alors, autant y proposer des mots importants...

Enfin, bien que cela ne soit pas nécessaire sur les documents "filles" pour lesquels la balise <TITLE> n'est logiquement pas requise, donnez un titre explicite à la page FH.html ainsi qu'aux autres documents qui seront indexés par le spider, car cela constituera une zone de mots clés importants. De même, insérez des balises <META> spécifiques (si vous avez décidé d'en mettre dans vos pages) à chacune des pages "filles", puisqu'elles seront traitées comme des pages à part entière par les moteurs. Bref, traitez les pages "filles" exactement comme si elles n'étaient pas des parties de frames mais des pages HTML comme d'autres. Là aussi, beaucoup de travail en perspective...

Si, éventuellement, vous ne vouliez pas insérer le code JavaScript de reconstitution de l'environnement "framé", et si vous voulez éviter que les pages "filles" soient indexées par les moteurs (seules seraient prises en compte, dans ce cas, les pages "mères"), il vous faudra insérer des balises <META NAME="ROBOTS" CONTENT="NONE"> dans chacun de ces documents.

Un avantage inattendu des pages réalisées avec des frames, et notamment des documents "mères", est que le peu de mots qu'ils contiennent en font des pages souvent très optimisées. Cela les fait d'ailleurs souvent classer en tête dans les résultats des moteurs. A quelque chose malheur est bon ;-)...

La technique du "Frame 100%" est, d'ailleurs, très souvent utilisée pour pallier un facteur bloquant

sur un site web. Exemple : des urls dynamiques (avec un "?" dans l'adresse des pages), qui pourraient être refusées par les moteurs.

Pour illustrer ce fait, prenons l'exemple d'une page répondant à l'adresse :

<http://www.monsite.com/prod/art/search/?sp-a=00051053-sp00000002&sp-q=stylo>

Cette url étant considérée comme "dynamique" (un "?" y est inclus), elle risque donc de ne pas être acceptée sur certains moteurs de recherche.

Une astuce consiste alors à l'intégrer à l'intérieur d'une URL statique avec une frame 100%.  
Illustration :

```
<HTML>
<HEAD>
<TITLE>Titre de la page (a choisir avec soin)</TITLE>
<META NAME="description" CONTENT="descriptif à choisir avec soin">
<META NAME="keywords" CONTENT="mots clés à choisir avec soin">
</HEAD>
<FRAMESET COLS="100%" border=0 framespacing=0 MARGINHEIGHT=0
frameborder=no>
<FRAME SRC="http://www.monsite.com/prod/art/search/?sp-a=00051053-
sp00000002&sp-q=stylo" NAME=main MARGINWIDTH=0 MARGINHEIGHT=0 border=0
frameborder=no framespacing=0>
</FRAMESET>
<NOFRAMES>
<BODY>
Texte de la balise Noframes (à choisir avec soin)
</BODY>
</NOFRAMES>
</HTML>
```

Finalement, que contient cette page ? Un titre, des balises meta et un texte qui ne s'affichera pas sur l'écran du navigateur. Bref, tout l'attirail du parfait référenceur / positionneur, pour peu que son site ait un bon indice de popularité...

En jouant finement sur ces champs, vous devriez arriver à bien référencer et positionner la page web dynamique qui, au départ, aurait posé de nombreux problèmes, ne serait-ce que pour son intégration dans les index des moteurs.

Attention, cependant, que cela ne vous pousse pas à tenter de spammer les moteurs sous prétexte que le contenu des balises <NOFRAMES> ne s'affiche pas sur l'écran du navigateur ! Car le moteur, lui, détectera la supercherie et vous black-listera à tire-larigot...

En revanche, si vous proposez un bon titre (pas plus de 10 mots et des termes explicites du contenu du site), des bonnes balises Meta (éventuellement car, nous l'avons déjà dit, ce champ est moins efficace aujourd'hui) et surtout un contenu textuel pour la balise <NOFRAMES> très descriptif, contenant des liens, des mots importants mis en exergue (gras, balises H1, etc.) mais surtout ayant un rapport avec le contenu du site affiché, cela ne devrait pas poser de problèmes...

Bien entendu, ces astuces ne sont valables que si les moteurs de recherche acceptent et lisent le contenu de la balise <NOFRAMES>. Eh bien, bonne nouvelle, c'est le cas... Nous avons, en effet, dans le cadre de cet article, posé la question aux principaux moteurs de recherche actuels. Voici leur réponse :

**Question : "Est-ce que votre moteur prend en compte le contenu de la balise <NOFRAMES> des pages réalisées avec des frames ?"**

**Réponse :**

Google	Altavista	Fast	Voila	Inktomi
Oui	Oui*	Oui	Oui	**

\* La réponse d'Altavista n'a cependant pas été très claire sur ce point : "AltaVista has just launched its newest index the 5th of April. This new index contains almost 500 million pages. As AltaVista wants to give the best results to the Internet users, AltaVista canceled a lots of identical pages. This new index supports some "framed content", but only under certain conditions that ensure the integrity of the index is maintained. AltaVista is committed to providing the user with the best search experience, and our current level of "framed content" support ensures good results."

A vous d'interpréter leur réponse... ;-)

Je pense qu'il s'agit du fait qu'il est nécessaire que la page mère et la page fille éventuellement affichée soient présentes obligatoirement dans l'index pour être prises en compte. Mais c'est une interprétation.... Enfin, on a au moins appris qu'Altavista a un nouvel index ;-)

\*\* Inktomi avait promis de répondre de répondre à la question mais j'attends toujours... J'espère compléter cet article avec la réponse d'Inktomi le mois prochain...

Et après, on se pose des questions sur la mauvaise santé d'Altavista et d'Inktomi...

### Sites réalisés en Flash

La problématique des sites réalisés en Flash est finalement assez proche de celle des sites bâtis à l'aide de frames. Cette problématique peut se résumer en quelques mots : les moteurs ne savent pas lire le contenu des fichiers Flash.

Dans un premier temps, voyons comment est structuré un fichier Flash. Lorsque vous créez une animation, vous obtenez un fichier, nommé, par exemple, anim fla (l'extension .fla est caractéristique du format Flash). Pour afficher ce fichier dans une page web, il est nécessaire de l'exporter au format Shockwave Flash (extension .swf). C'est ce fichier, une fois exporté, que vous allez utiliser pour votre site web.

Si l'animation réalisée contient du texte, celui-ci ne sera pas pris en compte par le moteur de recherche. Seuls certains moteurs de recherche (dont Google) savent explorer en partie ces fichiers pour en suivre éventuellement les liens, mais rien de plus. Le texte d'une animation Flash n'est donc jamais pris en compte dans le calcul de pertinence des moteurs...

Une page (voire un site !) entièrement réalisée à l'aide de la technologie Flash passera donc quasi inaperçue pour les moteurs de recherche, qui n'indexent que du contenu HTML. Mais le fichier HTML qui "lance" l'animation Flash est, lui, pris en considération. Dans ce cas, le remède sera très proche de celui déjà indiqué pour les frames (voir ci-dessus) : utilisation optimisée des balises <TITLE>, <META> et, dans notre cas, <NOEMBED> (de la même façon qu'on emploie <NOFRAMES>). Dans la balise <NOEMBED>, essayez d'insérer le plus possible de texte (sans spammer), afin que celui-ci soit "aspiré" par les spiders des moteurs. Exemple :

```
<HTML>
<HEAD>
<TITLE>Titre de la page (a choisir avec soin)</TITLE>
<META NAME="description" CONTENT="descriptif à choisir avec soin">
<META NAME="keywords" CONTENT="mots clés à choisir avec soin">
</HEAD>
<BODY bgcolor="#000099" leftMargin=0 topMargin=0 marginheight="0"
marginwidth="0">
<OBJECT classid="clsid:D27CDB6E-AE6D-11cf-96B8-444553540000"
codebase="http://download.macromedia.com/pub/shockwave/cabs/flash/swflash.c
ab#version=6,0,23,0"
WIDTH="550" HEIGHT="400" id="myMovieName">
<PARAM NAME=movie VALUE="myFlashMovie.swf">
<PARAM NAME=quality VALUE=high>
<PARAM NAME=bgcolor VALUE=#FFFFFF>
<EMBED src="myFlashMovie.swf" quality=high bgcolor=#FFFFFF WIDTH="550"
HEIGHT="400" NAME="myMovieName" ALIGN="" TYPE="application/x-shockwave-
flash" PLUGINSPPAGE="http://www.macromedia.com/go/getflashplayer">
<NOEMBED>
```

Texte de la balise Noembed (à choisir avec soin)

```
</NOEMBED>  
</EMBED>  
</OBJECT>  
</BODY>  
</HTML>
```

Comme pour les frames, les zones à optimiser sont le titre, les balises Meta et la zone <Noembed>. Si c'est le site en entier qui est concerné, vous ne pouvez jouer que sur les seules (rares ?) pages HTML présentées pour en optimiser le contenu.

En tout état de cause, retenez toujours un point essentiel : un *spider* (donc un moteur de recherche) ne sait lire que du code HTML ! Le contenu d'une animation Flash en est donc banni.

En revanche, là aussi, cette astuce n'est utile que si les spiders vont lire le contenu de la balise <Noembed>. Nous avons donc posé la question aux moteurs, comme pour les frames. Voici leur réponse :

**Question : "Est-ce que votre moteur prend en compte le contenu de la balise <NOEMBED> des pages réalisées en Flash ?"**

**Réponse :**

<b>Google</b>	<b>Altavista</b>	<b>Fast</b>	<b>Voila</b>	<b>Inktomi</b>
<b>Oui</b>	<b>Oui</b>	<b>Oui</b>	<b>Non*</b>	<b>**</b>

\* La version actuelle de Voila ne prend que rarement en compte le contenu de la balise <Noembed>. Mais la prochaine version du moteur (sortie prévue au plus tard en milieu d'année 2002) le prendra, elle, en compte.

\*\* Inktomi avait promis de répondre de répondre à la question mais j'attends toujours... J'espère compléter cet article avec la réponse d'Inktomi le mois prochain...

Il est donc tout à fait possible d'utiliser cette balise pour palier le fait que le format Flash n'est pas pris en compte par les moteurs. Sinon, il reste la possibilité d'insérer la page réalisée en Flash dans une frame 100%... Ca ferait une excellent révision ;-)

**Lancement du nouveau moteur d'AOL France : AOL.FR**[Retour au sommaire de la lettre](#)

Après de longues semaines d'attente, AOL France (<http://www.aol.fr/>) lance le 17 avril 2002 son nouvel outil de recherche, développé conjointement avec la société parisienne Exalead (<http://www.exalead.com/>). Une bonne raison pour faire un tour complet de l'outil (moteur + annuaire) avec Frédéric Mahé, Directeur Portails et Recherche d'AOL France. Dans la prochaine lettre R&R (datée du mois de mai 2002), nous reviendrons plus en détail sur la partie "moteur" de ce nouvel outil.

**AOL.FR**

Accueil | E-mail | Recherche | AIM | AOL Contacts | Shop@AOL | Mobile | Découvrir AOL

AOL Mail | Messagerie instantanée | Portefeuilles Boursiers

**Recherche**

Rechercher

Web français | Web mondial

**Actualités et médias**  
Journaux, Radios, TV...

**Autos et motos**  
Achat & vente, Presse...

**Bourse et finance**  
Actu boursière, Banque...

**Carrière et emploi**  
CV, Formation, Offres...

**Cinéma**  
Acteurs, Films, Presse...

**Immobilier**  
Crédit, Location et vente...

**Jeux**  
Actu, Jeux vidéo, Download...

**Multimédia**  
Web, Téléphonie, Webcams...

**Musique**  
Actu, Artistes, MP3...

**Santé**  
Forme, Grossesse, Maladie...

**Sciences**  
Astronomie, Généalogie...

**Shopping**  
Informatique, Livres & CD...

**Sports**  
Foot, F1, Rugby, Tennis...

**Tourisme et Voyage**  
Bons plans, Hébergement...

**AIM, la messagerie instantanée**  
Télécharger AIM maintenant | Envoyer AIM à un ami

- Annuaire téléphonique
- AOL Emploi
- E-cartes
- AOL Auto
- AOL Finance
- Logithèque
- AOL Contacts
- AOL Sports
- Plans et cartes
- AOL Voyage
- Téléphonie

**Aujourd'hui,**

**Music non stop sur AOL**  
Et aussi...  
Places de concerts  
Salons & forums musicaux  
Top 10 Warner Music

**Infos**  
Médecins, accord entre Médecins Généralistes  
> [A la Une](#) | [France](#) | [Monde](#) | [Sports...](#)

**Météo**  
Clermont-Fer. : Min. 7°C Max. 14°C  
> [Autres villes](#)

**Horoscope**  
Signe du mois : [Capricorne](#)  
> [Autres signes](#)

**Programmes TV**  
20h50 TF1 : [Navarro](#) et les méchants bandits  
> [Autres chaînes](#)

**La sélection Shop@AOL**  
[Jeu GTA III pour PSX2 : 59,90 Euros](#)  
> [Toutes les affaires](#)

**Bourse**  
CAC40 : 447,74 Pts -4,7%  
Rechercher une valeur  
OK  
> [Toute la bourse](#) | [Portefeuilles](#)

Découvrez AOL | Ecrivez-nous | AOL recrute | A propos d'AOL

AOL dans le monde : [Allemagne](#) | [Argentine](#) | [Australie](#) | [Brésil](#) | [Canada](#) | [Espagne](#) | [GB](#) | [Japon](#) | [Mexique](#) | [USA](#)  
[Service presse](#) | [Publicité](#)  
Tous droits réservés © AOL France

**- Question : Bonjour et merci d'accepter de répondre à nos questions. Parlons stratégie pour commencer. Pourquoi avoir changé d'outil de recherche sur AOL France et développé vos propres fonctionnalités ? L'ancien outil ne donnait pas satisfaction ?**

- Réponse : La recherche est une des fonctions les plus utilisées par les internautes, conscients du potentiel d'informations disponibles sur Internet. Le choix d'AOL France a été d'innover et de se différencier de ses concurrents afin de proposer des outils de plus en plus performants pour répondre aux attentes des internautes. AOL.FR est un outil de recherche novateur et puissant : une synergie parfaite entre un moteur de recherche et un annuaire. Le but était de répondre aux



attentes du marché en prenant en compte le retour de nos abonnés. Il aurait été possible d'essayer d'améliorer l'outil existant, mais il ne nous permettait pas de prendre en compte certaines problématiques spécifiques de notre cahier des charges. Nous avons alors décidé de nous orienter vers un nouvel outil afin d'innover dans ce domaine.

***- Pourquoi avoir choisi Exalead pour ce nouvel outil ? Avez-vous testé d'autres solutions disponibles sur le marché ?***

- Avant de choisir Exalead, nous avons regardé l'ensemble des solutions du marché, comme celles de Google ou d'Inktomi, mais nous avons pensé que la capacité d'innovation d'Exalead était plus forte et pouvait mieux répondre à l'ambition que nous avons pour ce projet.

***- Ce nouvel outil a-t-il pour vocation d'être uniquement disponible sur AOL France ou pourrait-il être, à l'avenir, être proposé à d'autres sites du groupe AOL, voire à des portails ne faisant pas partie du groupe ?***

- Pourquoi pas ? Pour le moment AOL.FR est un outil franco-français proposé aux internautes français. Les autres filiales regardent ce produit de près.

***- Quelles sont les ambitions et objectifs de ce nouvel outil ?***

- Nous souhaitons proposer une véritable alternative aux moteurs actuels et devenir un acteur incontournable en termes de recherche. Notre priorité est de répondre aux attentes des internautes.

***- Pouvez-vous nous présenter le nouvel outil de recherche d'AOL France ? Quelles sont ses principales fonctionnalités ?***

Le nouveau moteur d'AOL France est ultra rapide, novateur et vise à répondre parfaitement aux attentes des internautes. Ainsi, il propose l'un des plus gros index francophones, des règles de ranking conjuguant pertinence et popularité, ainsi que deux fonctionnalités inédites permettant à l'internaute de trouver de manière efficace selon ses propres besoins et selon la richesse réelle du Web. Ce moteur ouvre la voie à une nouvelle manière de chercher sur Internet, notamment grâce à la fonction d'affinage avec des mots automatiques remontés par l'analyse statistique des documents trouvés.

***- Quand sera-t-il mis en ligne ?***

- Le nouveau moteur de recherche sera accessible à tous, sur le service abonnés AOL ainsi que sur [www.aol.fr](http://www.aol.fr), dès le 17 avril 2002.

***- Les résultats ("liens sponsorisés") d'Overture seront-ils proposés dès le lancement ?***

- Non, nous y travaillons à l'heure actuelle, mais pour l'instant il n'y a pas de date fixée pour l'affichage de ces liens sponsorisés, qui seront, par ailleurs, clairement identifiés dans les pages de résultats.

***- Sur quel index le nouveau moteur travaille-t-il ?***

- L'index est scindé en deux : une partie francophone développée en collaboration avec Exalead, contenant 50 millions de pages (issues d'un crawl spécifique effectué par Exalead) et une partie mondiale (pour les recherches portant sur le Web mondial) développée par Inktomi, riche de plus d'un milliard de pages dans toutes les langues. Le nouveau moteur concerne la partie francophone, mais nous gardons la solution Inktomi pour le Web mondial. Il sera toujours temps, par la suite, d'étendre la solution développée avec Exalead à l'ensemble du Web mondial. Pour l'instant, nous nous focalisons plutôt sur les requêtes des internautes francophones, ce qui est bien logique.

***- Quelles sont les principales fonctions de recherche (opérateurs booléens, recherche avancée, etc.) ?***

- Sur AOL.FR, il est nullement besoin d'utiliser les opérateurs booléens. Le moteur décide de l'opérateur le mieux adapté à la recherche. Il n'est pas nécessaire d'utiliser les requêtes complexes

et fastidieuses conjuguant les AND, OR, +, "" et autres et qui souvent donnent des résultats décevants et déroutent les internautes. Sur AOL.FR, l'internaute n'a plus besoin de penser de manière compliquée sa requête, le moteur lui suggère des pistes de recherche et d'affinage grâce à des groupes de mots pertinents issus de l'analyse statistiques des documents trouvés. Grâce au côté magique de ces mots suggérés, la recherche devient intuitive.

The screenshot shows the AOL.FR search engine interface. At the top, there is a navigation bar with links for Accueil, E-mail, Recherche, AIM, AOL Contacts, Shop@AOL, Mobile, and Découvrir AOL. Below this is a promotional banner for "50H D'ESSAI" (50 hours of free trial) for AOL services. The search bar contains the word "Louvre" and a "Rechercher" button. Below the search bar, there are radio buttons for "Web français" and "Web mondial". The search results are displayed in a blue header: "Vous cherchez : louvre" and "Triez par rubrique". There are three main categories: "Musées et centres culturels", "Hébergement", and "Art et antiquités". Under "Musées et centres culturels", there are sub-links for "Musées d'art et d'histoire" and "Musées nationaux". Under "Hébergement", there are sub-links for "Paris" and "12<sup>e</sup> arrondissement". Under "Art et antiquités", there are sub-links for "Île-de-France" and "Paris". The search results show "77 245 documents trouvés" in "0.64 seconde". Below this, there is a section "Affinez avec les mots :" with a grid of suggested terms: Musée du Louvre à Paris, Carrousel du Louvre, Louvre Museum, Jardin des Tuileries, Hôtel Louvre, Grand Louvre, Centre Georges Pompidou, and Opéra Garnier. The first result is "Musée du Louvre - Site officiel" with a description: "Découvrez l'histoire, les collections, les salles et les différentes activités du Louvre. Vente de billets et d'articles de boutiques." The second result is "Musée du Louvre - Collection égyptienne" with a description: "son savoir au service du Louvre. Le 15 décembre 1827, le musée égyptien du Louvre est inauguré par le roi..." The third result is "Louvre des antiquaires, le - Paris, ler" with a description: "B C Les galeries du Louvre des sont classées ici par ordre alphabétique. Vous pouvez les visiter une à une..." The fourth result is "Hôtel Louvre Piémont - Paris, ler" with a description: "PRES DE LA PYRAMIDE DU LOUVRE... Au cœur du Paris... deux pas du Musée du Louvre et du Jardin des..." At the bottom, there is a footer with "©2002 AOL France" and "Powered by Excite".

**- Comment peut-on se référencer sur ce moteur (délais, critères bloquants, etc) ?**

- De la même manière que les moteurs de dernière génération, AOL.FR s'appuie sur un crawl, un robot qui navigue de lien en lien, de site en site et en retient le contenu. Mais si être référencé dans un moteur est une chose, celle d'être visible ou de remonter dans les résultats en est une autre. C'est justement une des raisons pour laquelle, AOL.FR propose l'affinage avec les mots. Cette fonctionnalité permet ainsi de faire remonter des résultats intéressants qui jamais ne seraient visibles sur la page de résultats d'un moteur classique. Ils seraient noyés sous une masse colossale de résultats et pourtant répondent au mieux à la requête de l'internaute et au sens qu'il lui donne. La fonctionnalité d'affinage avec les mots offre ainsi une bien meilleure visibilité à tout le Web dans toute sa profondeur et permet ainsi, souvent, de trouver l'aiguille dans la botte de foin. Il n'y a donc pas de formulaire de soumission d'un site web dans notre index. Mais si un site est présent dans les annuaires principaux et s'il est déjà "pointé" sur le Web, il y a de fortes chances qu'il apparaisse rapidement sur notre moteur.

**- Quels sont les côtés les plus originaux de l'outil, par rapport à ce qui existe déjà sur le Web ?**

- AOL.FR propose deux fonctionnalités uniques sur le marché qui aident l'internaute à trouver le plus rapidement possible la réponse à sa question. D'une part, l'affinage de la requête, d'autre part, le tri par rubriques. Grâce à des mots pertinents suggérés par le moteur, l'affinage permet à l'internaute de préciser le sens de sa recherche, de lever des ambiguïtés et de sélectionner uniquement les résultats qui correspondent au sens qu'il choisit de donner lui-même à sa recherche. Quant au tri par rubriques, il permet de sélectionner les sites incontournables qui répondent le mieux à la requête. Cette nouvelle approche de navigation, intuitive, au sein d'un annuaire de sites devient une véritable alternative à la navigation fastidieuse proposée par les corpus monolithiques que sont les annuaires traditionnels.

**- Comment l'annuaire de sites Web d'AOL France a-t-il été constitué ? Certains disent qu'il a été construit sur la base de celui de Looksmart France ?**

- L'annuaire d'AOL France est effectivement issu, à l'origine, de celui de Looksmart. Depuis un an et demi, une équipe de documentalistes de Cybion, dédiée à AOL France, l'adapte et l'enrichit au quotidien de manière à en faire le coeur de pertinence du moteur AOL.FR. Il ne se veut pas exhaustif mais plutôt qualitatif, restreint pour s'adapter au mieux aux demandes des internautes.

**- Beaucoup de gens se posent la question suivante : pourquoi avoir développé son propre annuaire alors qu'AOL est propriétaire de l'Open Directory ?**

- Le choix d'AOL France a été une fois de plus, d'innover et de se différencier de ses concurrents. En étant l'unique utilisateur de son annuaire, AOL.FR s'impose de facto comme une alternative à tous les moteurs qui utilisent le même directory, l'ODP.

**- Combien de sites Web cet annuaire comprend-il aujourd'hui ?**

- Environ 90 000 sites.

**- Combien de documentalistes travaillent-ils à son élaboration ?**

- 5 documentalistes et 2 taxonomistes.

**- Comment peut-on référencer un site sur cet annuaire ?**

- La possibilité de suggérer un site a été supprimée il y a peu, mais elle sera à nouveau proposée dans les prochaines semaines. Il s'agit bien d'une fonctionnalité de suggestion de sites plus que de référencement.

**- Quels sont les critères d'acceptation et de refus sur cet annuaire ?**

Les documentalistes répondent à un cahier des charges bien précis qui prend en compte la qualité du site et de son information, sa fréquence d'actualisation, son aspect incontournable mais aussi et

surtout la demande des internautes en matière de recherches. L'objectif de cet annuaire n'est pas de référencer la totalité du Web francophone mais plutôt de classer de manière intelligente, pérenne et la plus naturelle possible, les sites essentiels, ce qui peut aller de certaines pages persos jusqu'aux sites institutionnels. Les sites répondant à ce cahier des charges sont intégrés dans l'annuaire et ceux qui n'y répondent plus sont supprimés.

**- Quel est le délai entre le moment où un site Web est accepté et celui où il est visible sur l'annuaire ?**

- Difficile à dire, c'est vraiment au cas par cas.

**- Quel est le délai entre le moment où un site est soumis et celui où il est évalué par un documentaliste ?**

- Là aussi, c'est vraiment du cas par cas. En fait, le gros du travail des documentalistes consiste en un travail de veille. Le formulaire de soumission de site est plutôt considéré comme une source d'informations complémentaires, en plus de la veille effectuée quotidiennement. Difficile donc de donner un délai précis voire même approximatif...

**- Envisagez-vous de mettre en place une procédure de soumission payante sur l'annuaire, comme les autres annuaires à l'heure actuelle (Yahoo!, Looksmart...) ?**

- Non, ce n'est pas prévu...

## Lycos France : un point sur l'outil de recherche

[Retour au sommaire de la lettre](#)

Au moment où Lycos France connaît une vague de licenciements sans précédent, il nous a semblé important de faire un point sur l'outil de recherche qu'il propose aux internautes et qui se place le plus souvent dans le "Top 5" des annuaires et moteurs les plus utilisés par les internautes francophones à l'heure actuelle. Arnaud Jonglez, Directeur Portail et Service, et Davy Lor, Responsable Communication du portail, ont accepté de répondre à nos questions.

### **- Question : Sur quelles bases de données s'effectuent les recherches d'information sur Lycos France à l'heure actuelle ?**

- Réponse : Nous utilisons actuellement trois sources d'informations : notre annuaire de sites web, maintenu par nos équipes de netsurfeurs et qui propose environ 75 000 sites francophones. Puis Fast, notre partenaire "moteur de recherche", qui fournit son index de 900 (ca augmente!!) millions d'objets (pages web, images, sons, vidéos). Enfin, Espotting fournit cinq liens sponsorisés (trois en haut à droite et deux en bas de la page de résultats).

### **- Lycos tente-t-il d'améliorer continuellement la qualité de ses résultats de recherche ?**

- Tout à fait, nous testons très souvent de nouvelles possibilités pour afficher nos résultats de recherche. Les résultats issus de l'annuaire sont, par exemple, intégrés en début de page de résultats. Pour ceux-ci, nous avons enlevé, dans un premier temps, la mention de la catégorie à laquelle appartient le site, mais nous l'avons finalement rajoutée. Aujourd'hui, les résultats issus de liens sponsorisés (Espotting) sont également clairement identifiés, à droite des liens "naturels" et en bas de page. Ceci dit, nous testons encore d'autres possibilités pour l'emplacement de ces liens sponsorisés.

Nous avons également essayé de simplifier le plus possible notre page de résultats, en limitant à deux colonnes (une pour les liens, une pour la publicité à droite) le nombre d'informations proposées. Nous proposons également moins de couleurs, d'éléments graphiques, de polices de caractères différentes, une meilleure navigation, etc. Le souci de l'internaute et de son confort visuel est quasi quotidien chez nous...

Fin mars, les pages de résultats de Lycos ont également été mises en place sous la forme d'un format compressé (tous les caractères "inutiles" sont enlevés dans le code HTML et on utilise un format spécifique de compression) pour accélérer le temps de chargement en diminuant d'un tiers la taille de ce document... Nous espérons que les internautes apprécieront !

### **- Proposez-vous de nouvelles fonctionnalités de recherche ?**

- La recherche avancée de Lycos France propose maintenant une fonction de recherche sur la date de dernière modification des documents ainsi que sur leur taille en octets. Cela permet d'éliminer d'une recherche des documents qui seraient trop lourds à charger et à lire ou trop anciens pour correspondre à la requête demandée.

De même, nous proposons maintenant un système de "clustering" en ne retenant qu'un lien par nom de domaine. Une mention "Autres résultats de ce site" apparaît alors pour obtenir les autres pages du même site répondant à la demande. Notons cependant que, pour Lycos, les noms de sites "www.abondance.com", "www.abondance.net" et "outils.abondance.com" sont considérés comme trois sites différents et seront bien affichés dans la page de résultats, même avec l'option de clustering.

Nous avons également rajouté des zones d'aide en ligne (pictogramme indiquant un point d'interrogation en bleu sur fond jaune) qui, nous l'espérons, vont aider les internautes à mieux trouver l'information qu'ils recherchent en utilisant au mieux notre outil.

### **- Comment se passe le partenariat avec Fast ?**

- Bien, merci :-). L'index est actuellement rafraîchi de façon assez fine sur une base de 9 à 12 jours. Des algorithmes spécifiques ont été mis en place pour crawler de façon très rapprochée certains sites et/ou certaines pages qui sont très souvent mises à jour (sites d'actualité notamment). Depuis février 2002, les pages dynamiques commencent également à être indexées. L'index de Fast est passé récemment de 625 à 900 millions d'objets (pages web, images, sons,

etc.). Fast a également récemment développé de nombreux outils pour lutter contre le spam. Enfin, notre partenaire a diminué d'un tiers le nombre de liens morts et de risques d'erreurs 404. Un "Dead link crawler" spécial a été développé pour lutter contre ce fléau. Dès qu'une page est trouvée en erreur 404, elle est éliminée de l'index dans les 48 heures. En dehors de notre travail avec Fast, nous travaillons également, à l'heure actuelle à une meilleure détection des requêtes en allemand et en français et la lemmatisation (identification d'un mot par son lemme, c'est-à-dire sa forme "de référence", exemple : le lemme de "seront" est "être") est en cours, notamment pour la langue française, et sera disponible d'ici peu.

***- Quelques nouvelles de l'annuaire de Lycos France ?***

- L'annuaire de Lycos France recense aujourd'hui 75 000 sites dans 7 500 catégories et est maintenu par 8 documentalistes et taxinomistes. En ce qui concerne l'inscription de nouveaux sites, priorité est donnée aux soumissions de sites de contenus, institutions, associations, etc. Nous avons très peu de retard pour ce type de soumission. Pour les entreprises "lambda", en revanche, les délais sont plus longs... Des "ménages" (retravail sur certaines catégories et remises à jour - modification ou retrait - des sites inscrits) sont également effectués de façon permanente sur l'outil.

***- De nombreux outils de recherche ayant développé, dans un premier temps, leur propre annuaire, réfléchissent actuellement à l'externalisation de cette fonctionnalité, pour des raisons de coût. Une telle réflexion existe-t-elle chez Lycos France ?***

- Non, pas pour le moment. Travailler avec des partenaires nous forcerait à payer soit une somme par lien fourni (très couteux!!) soit à partager le revenu de la soumission payante. Nous croyons être mieux outillé financièrement avec notre annuaire propre.

***- Comment testez-vous la qualité de vos résultats ?***

- Chez Lycos, nous testons régulièrement les résultats de notre moteur, mais également en comparant leur pertinence par rapport à ceux proposés par nos concurrents. Nous disposons d'un outil, que nous avons développé nous-même, et qui permet de proposer des "blind tests" à des panels d'utilisateurs chargés de noter chaque résultat. Le système est assez simple : certains mots clés (sélectionnés au hasard parmi les termes demandés sur notre site par les visiteurs) sont proposés et, pour chacun d'eux, les résultats de Lycos et de la plupart des outils de recherche principaux sont proposés, sans aucune mention indiquant la source de ces liens. Le "cobaye" note alors le degré de pertinence qu'il estime pour chaque lien proposé sans savoir s'il est fourni par Lycos, Google ou Voila. Cela nous permet, en analysant ces résultats, d'avoir une vision objective de notre pertinence par rapport à la concurrence. Cela nous permet également d'évaluer de nouvelles fonctionnalités potentielles et d'améliorer l'affichage des résultats en demandant l'avis de "beta-testeurs" dans un premier temps.

***- Qu'en est-il des projets de référencement et de soumission payante sur Lycos ?***

- L'Allemagne est notre pays "pilote" pour ces projets actuellement (voir l'adresse [http://www.lycos.de/service/add\\_site\\_all.html](http://www.lycos.de/service/add_site_all.html)). Des offres similaires devraient voir le jour dans le deuxième trimestre en France et plus globalement en Europe.

***- Quel est le premier bilan de ces offres en Allemagne ? (même si elles sont assez récentes). Quel est le résultat par rapport aux objectifs initiaux ?***

- On est très heureux du succès obtenu pour l'instant puisqu'on a pu convaincre plusieurs grands clients d'utiliser ce programme. De plus, les webmasters de "petits" sites ont bien accepté l'offre de la soumission payante dans l'annuaire.

***- Quand ces offres seront-elles disponibles en France ?***

- Nous avons quelques problèmes avec un partenaire extérieur (pas FAST, je vous rassure :-)), ce qui explique le retard pris jusqu'à maintenant, mais on estime pouvoir offrir ce service vers la fin du second trimestre de cette année.

**- Y aura-t-il une offre de soumission payante sur l'annuaire ou une offre de référencement payant sur Fast ou les deux ? En même temps ?**

- Les deux et en même temps.

**- La page de référencement d'un site sur Fast par l'intermédiaire de Lycos France (<http://www.lycos.fr/content/help/faq/zoom2.html>) aboutit, une fois le site soumis, sur une page blanche depuis de nombreuses semaines (auparavant, on arrivait sur un message d'erreur). Est-ce un bug ?**

- ?? Certainement. On va corriger cela au plus vite.

**- L'index de Fast utilisé par Lycos est-il le même que celui d'Alltheweb ?**

- Non, mais selon nos tests, la différence est minuscule.

**- Comment sont organisés les index de Fast pour les sites européens de Lycos ?**

- Nous utilisons un "production cluster" situé aux Etats Unis pour la recherche web. Il existe également des index spéciaux (multimedia, la soumission sur FAST et les pages les plus populaires). Les détails de l'architecture FAST ne sont malheureusement pas publiées. En Europe, nous utilisons un seul index pour tous les pays. L'organisation interne du "production cluster" est confidentielle. Afin de gérer un tel volume des données d'une manière efficace, je crois bien que tous les moteurs sont en train d'optimiser la hiérarchie des index utilisés.

**- Fast a-t-il plus particulièrement travaillé certains aspects de son algorithme de pertinence depuis quelques mois (en dehors des points indiqués auparavant) ?**

- Non, je ne crois pas, pas à ma connaissance.

**- Quels sont les projets de Lycos France en ce qui concerne la recherche d'information et le référencement au cours de l'année 2002 ?**

- Nous cherchons continuellement à améliorer l'expérience utilisateur en ajoutant un côté ludique à la consultation de notre outil de recherche, par exemple via la mise en place de "roulettes thématiques" (fonction "Un site au hasard" dans la partie gauche de certaines pages de l'annuaire). Nous cherchons également à animer l'annuaire en faisant découvrir aux internautes de nouvelles catégories par des mises en avant graphiques et des sélections thématiques ou événementielles réalisées par nos surfeurs.

**- Finissons par une question pas obligatoirement très agréable : Est-ce que les réductions d'effectifs chez Lycos.fr, décidées il y a quelques jours de cela, auront des répercussions sur la partie "search" du portail ? Et si oui, sait-on déjà lesquelles ?**

- Il est encore un peu tôt pour en parler, toutes les décisions ne sont pas prises à ce niveau. Mais il est certain que le département "portail et services" dont fait partie l'équipe "search" sera touchée mais nous ne pouvons vous donner plus de précision pour l'instant.

## Différences culturelles sur les mots clés

[Retour au sommaire de la lettre](#)

Spotting nous a fourni dernièrement la liste des 49 mots clés et expressions le plus souvent demandés sur son réseau d'affiliés (Lycos, HotBot, Altavista, yahoo!, etc.) en Grande-Bretagne, en France et en Allemagne. Une bonne occasion pour regarder s'il existe des différences notables entre les trois pays au niveau des termes le plus souvent demandés sur les outils de recherche du Web. Et, effectivement, il y a de grandes différences...

Rang	Français	Anglais (Grande-Bretagne)	Allemand
1.	sexe	casino	anal sex
2.	caramail	anal	routenplaner ( <i>calculs d'itinéraires</i> )
3.	google	online casino	sex
4.	pages jaunes	sex	morpheus
5.	sex	gambling	chat
6.	sncf	roulette	ebay
7.	anpe	travel	Erotik
8.	annuaire	google	sms
9.	mp3	slots	google
10.	porno	finance	telefonbuch ( <i>annuaire téléphonique</i> )
11.	yahoo	jobs	porno
12.	hotmail	music	Grußkarten ( <i>carte de voeux</i> )
13.	wanadoo	poker	musik
14.	meteo	online roulette	Wetter ( <i>météo</i> )
15.	impots	online slots	spiele ( <i>jeux</i> )
16.	emploi	slots online	Gebrauchtwagen ( <i>voitures d'occasion</i> )
17.	jeux	viagra	mp3
18.	Musique	hotmail	free sms
19.	lycos	shopping	Prügelpause ( <i>jeu de combat</i> )
20.	immobilier	Car Insurance	cheats
21.	chat	PlayStation 2 Cheats	Reisen ( <i>voyages</i> )
22.	sexe gratuit	yahoo	madonna
23.	fnac	loan	songtexte ( <i>paroles de chansons</i> )
24.	voila	cars	test
25.	horoscope	porn	horoskop ( <i>horoscope</i> )
26.	sms	cheap flights	auto
27.	Humour	test	treiber ( <i>drivers</i> )
28.	divx	hardcore	sven
29.	sms gratuit	flights	bildschirmschoner ( <i>écran de veille</i> )
30.	rencontre	life insurance	arbeitsamt ( <i>agence pour l'emploi</i> )
31.	voissa	ask jeeves	Bilder ( <i>images</i> )
32.	pages blanches	games	immobilien
33.	michelin	career	telefonauskunft ( <i>info par téléphone</i> )
34.	air france	estate agents	intel
35.	TF1	chat	grusskarten ( <i>carte de voeux</i> )
36.	sfr	friends reunited	praline ( <i>chocolat, magazine érotique</i> )
37.	zoophilie	Holidays	kontaktanzeigen ( <i>annonces contact</i> )
38.	crack	electronics	gmx ( <i>portail SMS</i> )
39.	gay	mobile phones	suchmaschinen ( <i>moteurs de recherche</i> )
40.	argus	hotels	referate ( <i>exposés</i> )
41.	dictionnaire	maps	gedichte ( <i>poèmes</i> )
42.	voyages	health	last minute



<b>43.</b>	impot	loans	Hintergrundbilder ( <i>fond d'écran</i> )
<b>44.</b>	le monde	software	klingeltöne ( <i>sonneries</i> )
<b>45.</b>	morpheus	free ringtones	stadtplan ( <i>plans de quartiers</i> )
<b>46.</b>	Harry Potter	pop idol ( <i>émission TV</i> )	web.de
<b>47.</b>	Britney Spears	Blackjack	cliparts
<b>48.</b>	loto	bbc	wörterbuch ( <i>dictionnaire</i> )
<b>49.</b>	cul	property	rezepte ( <i>recette</i> )

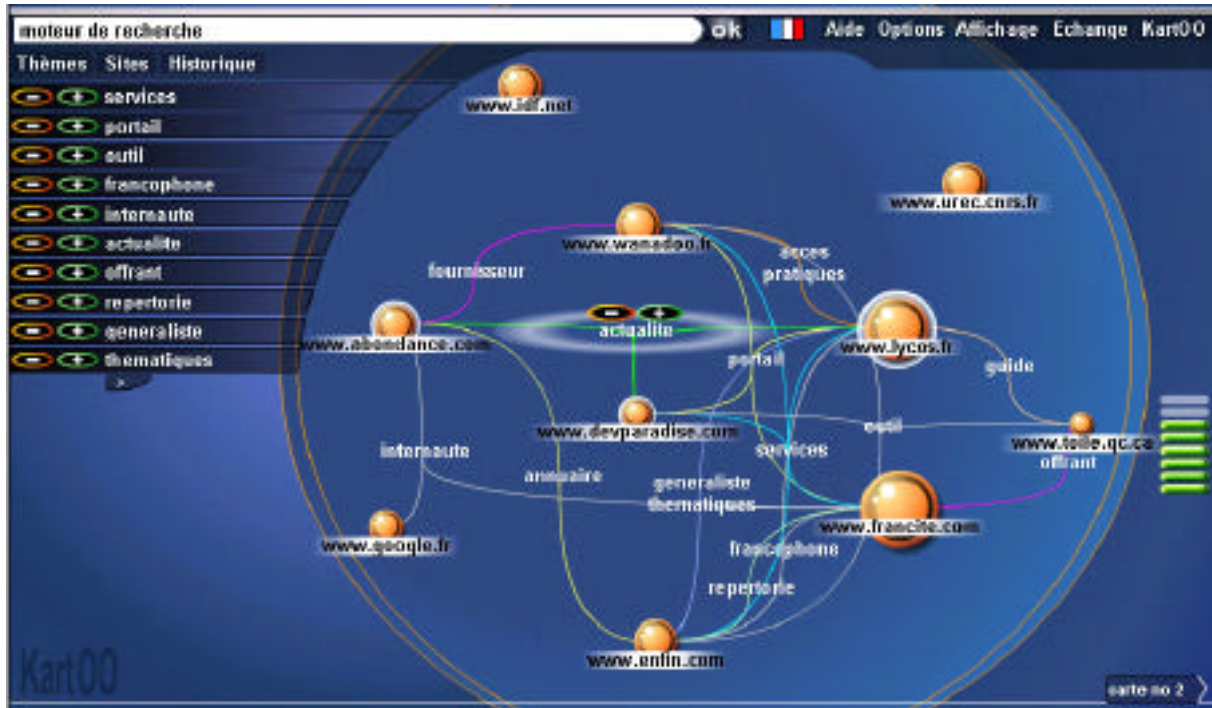
Côté français, pas vraiment de surprises avec 8 termes ayant rapport au sexe (aucune des listes ci-dessus n'est censurée), 7 à des portails ou sites d'outils Internet (caramail, google, yahoo...), 6 à des sites dits "utiles" ou de commerce électronique (snCF, fnac, michelin, tf1...), 5 à des outils de recherche "traditionnels" (pages jaunes, annuaire, dictionnaire...), 11 à la vie pratique et au divertissement (meteo, impots, jeux, musique...), 3 aux télécoms (sms, sfr...), 2 à l'emploi (anpe, emploi) et, finalement, assez peu de termes du domaine informatique (chat, divx, morpheus). Seuls Harry Potter et Britney Spears apparaissent en fin de classement en tant que personnalités ou personnages. On peut également noter que peu de gens ont encore le réflexe de taper "www.tf1.fr" pour aller sur le site de la chaîne télévisée mais passent encore par un outil de recherche avec le mot clé "TF1", ce qui est assez étonnant. Le cas de TF1 est encore plus étonnant puisque son adresse est publiée un peu partout sur l'écran de télévision, notamment à la fin des émissions...

Les anglais, pour leur part, semblent obnubilés par les jeux de casinos en ligne. Pas moins de 10 termes et expressions y ont rapport (casino, gambling, online roulette...). Bien évidemment, le sexe est présent également, mais avec seulement 5 termes (dont "viagra", mot clé absent des listes française et allemande, mais ne comptez pas sur moi pour en tirer des conclusions hâtives ;-)). Point intéressant : les anglais semblent moins saisir de noms de sites (5 seulement : google, hotmail...) sur les outils de recherche. Les mots clés demandés, sinon, sont assez "classiques" et généralistes (finance, jobs, music, shopping, flights...). Là aussi, comme chez les français, assez peu de termes ayant trait à l'informatique. On peut noter cependant quelques particularités étonnantes : l'assurance (thème absent des requêtes françaises) semble très présente dans les préoccupations de nos amis anglais (car insurance, life insurance), ainsi que la recherche d'amis perdus de vue (friends reunited).

Les allemands sont, également, moins portés que les français sur le sexe (5 termes "seulement"). L'informatique (6 requêtes : morpheus, drivers, écran de veille...) semble plus présente et la téléphonie (4 requêtes : sms, free sms, gmx...) est également très demandée, comme ailleurs. Notons que seuls 3 noms de sites sont demandés (ebay, google et web.de), ce qui pourrait indiquer une maturité plus grande des utilisateurs du réseau outre-Rhin (dans ce cas, on peut se poser la question de la présence du mot clé "test" en 24ème position ???). Quelques particularités se font jour quand même : Madonna a encore la cote chez nos amis allemands et l'Internet semble servir aux écoliers pour faire leurs devoirs ("exposés" et "poèmes").

**Kartoo : la V2 est pour bientôt**[Retour au sommaire de la lettre](#)

La "V2" du métamoteur Kartoo (<http://www.kartoo.com/>) va bientôt être mise en ligne par l'équipe qui développe l'outil à Clermont-Ferrand. Voici une première vision de cette interface :



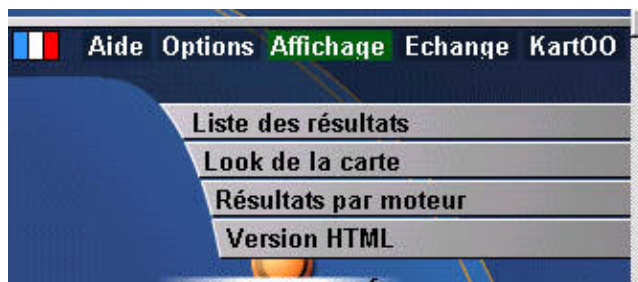
Au niveau "look", l'interface a un côté beaucoup plus professionnel, moins "BD" qu'avant. Un choix logique, puisque la cible de Kartoo est avant tout institutionnelle. Mais rassurez-vous, le génie est toujours là ! (il suit toujours l'actualité et change tous les mois en fonction de celle-ci, une galerie d'images d'archives est d'ailleurs prévue pour bientôt...). Au niveau des fonctionnalités ajoutées à l'outil, la principale est la taille des "boules" qui représentent les sites et qui ont maintenant une taille variable. Plus la boule est grande, plus sa pertinence est forte. Auparavant, cette taille était calculée par l'interface Flash en fonction du nombre de liens vers ce site (donc son indice de popularité). Maintenant, la taille est calculée côté serveur et dépend directement de la pertinence du site. Lorsque l'on saisit des noms de marques comme mots clés (Renault, EDF, BMW, Fnac), les sites officiels ressortent ainsi à chaque fois avec une boule bien plus importante, ce qui est très pratique et efficace. Avec le nouveau look et le temps de chargement accéléré des cartes (l'interface est plus légère : 12Ko pour la page principale et 20ko pour les menus qui se chargent progressivement au lieu de 105ko pour l'ancienne version), il s'agit là de la principale innovation de cette nouvelle version. Rajoutons quand même à cela le temps de réponse du métamoteur (code optimisé, meilleure bande passante, serveur supplémentaire) et la liste des outils de recherche interrogés qui va être revue progressivement (suppression des moteurs disparus, regroupement des moteurs qui interrogent les mêmes sources, interrogation directe de DMoz, tri en amont des liens sponsorisés, etc.)

Plusieurs options sont disponibles en haut et à droite de l'écran :

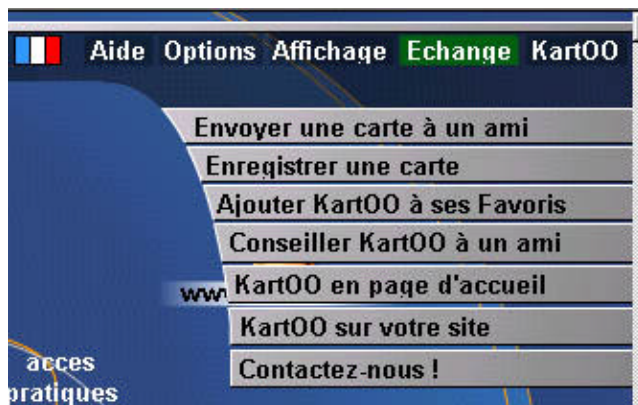
\* Le choix "Option" permet de configurer l'outil de recherche selon plusieurs critères :



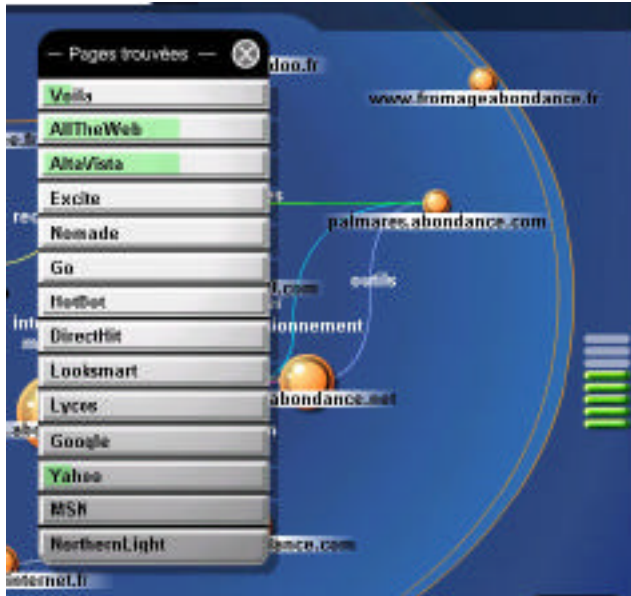
\* Le choix "Affichage" permet de visualiser les résultats selon des interfaces configurables (look de boules, fond d'écran, couleurs, etc.)... :



\* Le choix "Echange" permet d'effectuer plusieurs manipulations d'envoi de cartes ou autres, comme indiqué ci-dessous :



Enfin, le baromètre, à droite (les barres vertes verticales), indique le nombre de résultats fourni par chaque outil interrogé pour le mot clé demandé :



Sur ce baromètre, plus le niveau est haut, plus il y a de résultats trouvés, donc plus on a intérêt à préciser sa recherche.

Le lancement de cette nouvelle interface est prévu pour le courant de la semaine 17 (entre le 22 et le 27 avril, aucune date précise n'étant encore arrêtée au moment où cet article était écrit).

## Bruits et chuchotements

[Retour au sommaire de la lettre](#)

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 ;-)

-> Overture arrive en France. Une personne, déjà nommée, démarcherait actuellement les référenceurs de la place avec des conditions très intéressantes, semble-t-il... Difficile de venir après Espotting, qui a raflé une bonne partie du marché !

-> Multimania connaissant pas mal de problèmes techniques en ce moment, il semblerait que les sites hébergés chez lui aient également quelques soucis de référencement. En effet, depuis de nombreuses semaines, des dysfonctionnements nombreux (accès impossible, mise à jour progressive des pages etc.) semblent perturber l'hébergeur gratuit. Du coup, les spiders des moteurs n'arriveraient plus à crawler les pages des sites et enlèveraient petit à petit les pages Multimania de leur index. Google semblerait notamment assez touché par ce problème à l'heure actuelle.

-> Overture et Espotting devrait bientôt (juin ?) proposer un outil de "bid optimisation tool" (outil comparatif des enchères sur certains mots clés auprès de différents fournisseurs). Espotting devrait également proposer prochainement un outil de "push" pour que les annonceurs suivent plus facilement leurs enchères sur leur site (rapports quotidiens, état du compte, etc.)

-> Fast serait en train de faire le "forcing" pour fournir les résultats de HotBot aux US. En Europe, HotBot ne semblerait pas menacé pour l'instant et continuerait de travailler avec Inktomi...

-> HotBot France serait un peu "laissé à l'abandon" en entendant les mesures prises et les conséquences des récents licenciements chez Lycos, propriétaire du moteur de recherche.

-> Altavista devrait proposer d'ici peu des systèmes de catégorisation des urls proposées, un peu à l'image des "Custom Search Folders" de Northern Light ou de ce que proposent Teoma ou Vivisimo. L'outil devrait s'appeler "AltaVista ParaPhrase".

-> Google aurait, dans son index, environ 3 millions de pages rafraîchies sur une base quotidienne... Un lien intéressant à ce sujet :  
<http://www.searchengineshowdown.com/features/google/dbanalysis.shtml>

## En bref...

[Retour au sommaire de la lettre](#)

-> Google permet maintenant de soumettre des requêtes contenant plusieurs termes en syntaxe avancée, comme :  
allintitle:altavista site:www.abondance.com  
Ce qui n'était pas possible auparavant, où seule une syntaxe avancée (allintitle: OU site:) pouvait se trouver dans la requête.

-> Inktomi a dernièrement changé sa façon d'afficher les résultats issus de son index. Si le site proposé est sur Looksmart, c'est la description de Looksmart qui est prise en compte. Si le site n'est pas sur Looksmart mais sur l'Open Directory, c'est le résumé de ce dernier qui est affiché. Si le site n'est dans aucun des deux annuaires, c'est Inktomi qui se base sur le texte de la page pour créer un résumé (balise Meta description si elle existe ou 200 premiers caractères de la page).

-> Après Jean-François Marti, Dg de Nomade.fr, c'est Pauline Tourneur, Responsable Marketing Recherches, qui est partie de l'outil de recherche francophone. Pour l'instant, c'est Frédéric Saint Sardos qui s'occupe du portail en attendant la suite des événements...

### Les nouveaux entrants dans l'annuaire des outils de recherche régionaux

[Retour au sommaire de la lettre](#)

**Pour obtenir tous les sites :**

<http://annuaire.abondance.com/>

-> Portail Auvergne-Net (région Bourgogne)

<http://www.auvergne-net.com>

-> Marseille-annuaire.com (région PACA)

<http://www.marseille-annuaire.com>

-> Ariège Pyrénées (région Midi-Pyrénées)

<http://www.ariège-pyrenees.org/rechercher.php3>

### Cherchez, référencez-vous (nouveaux outils ou rappel d'outils existants)

[Retour au sommaire de la lettre](#)

-> Articles Sciences

<http://articlesciences.inist.fr/>

Moteur de recherche et de commande d'articles scientifiques.

-> 36 15

<http://www.36-15.fr/>

Nouveau moteur de recherche francophone

-> Recherche Web

<http://www.recherche-web.com/recherche.shtml>

Nouveau moteur de recherche francophone

-> Yoopla

<http://www.yoopla.net>

Nouvel annuaire francophone

-> Tu Cherches

<http://www.tucherches.ch/>

Nouvel outil de recherche sur le web suisse

### Contenu (sites proposant du contenu ou des fonctions intéressantes)

[Retour au sommaire de la lettre](#)

-> Lettres types

<http://www.lentreprise.com/ASP/home/Ttesleslettreslibrary.asp>

[http://www.apce.com/index.php?rubrique\\_id=500000000&type\\_page=IH&simple=1&rech=1&mtc=modele](http://www.apce.com/index.php?rubrique_id=500000000&type_page=IH&simple=1&rech=1&mtc=modele)

Deux bases de données de lettres types, deux véritables cavernes d'Ali-Baba.

-> DoubleVeille

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

Nouveau service de veille proposé par la société Mediaveille.

-> Search Engine World

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

Nombreuses infos sur les outils de recherche anglophones.

## Outils (logiciels et sites web proposant des fonctionnalités utiles)

[Retour au sommaire de la lettre](#)

-> Memotoo

<http://www.memotoo.com>

Barre de navigation qui permet de mémoriser ses favoris de façon simple

-> Web Net Visa Services

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

De nombreux applicatifs pour webmasters (livre d'or, forums, faq, etc.)

-> Mes-statistiques.com

<http://www.mes-statistiques.com/>

Outil de suivi de statistiques de site web

-> Portails pour webmasters

[http://www.bonweb.com/p\\_1\\_439.html](http://www.bonweb.com/p_1_439.html)

Développeurs, ne réinventez pas la roue. Dès que vous rencontrez un problème, cherchez la solution sur le Web. Découvrez les sites de développement qui répondent à vos questions.

-> Simulateur de spider

[http://www.searchengineworld.com/cgi-bin/sim\\_spider.cgi](http://www.searchengineworld.com/cgi-bin/sim_spider.cgi)

Simule un navigateur ou un spider de moteur.

-> Search On Line

<http://www.searchonline.info/main.jsp>

Métamoteur Web.

## Revue d'URL

[Retour au sommaire de la lettre](#)

-> Google Asked to Delist Scientology Critics

<http://www.chillingeffects.org/dmca512/notice.cgi?NoticeID=232>

Copie de la lettre adressée par l'église de Scientologie à Google pour bannir de l'index du moteur certaines pages hostiles à la secte.

-> Veille sur Internet

<http://www.journaldunet.com/rubrique/veille/veille020412.shtml>

Les astuces pour optimiser la recherche d'informations, par Carlo Revelli (Cybion).

-> Teoma

<http://www.researchbuzz.com/articles/teoma0402.html>

Revue de détail du nouvel util de recherche Teoma.

-> Pages dynamiques

[http://developpeurs.journaldunet.com/tutoriel/php/020407php\\_urlrewriting.shtml](http://developpeurs.journaldunet.com/tutoriel/php/020407php_urlrewriting.shtml)

Un article très intéressant sur la façon de référencer des pages web dynamiques

-> Search Engines : What's next

<http://europemedia.net/showfeature.asp?ArticleID=9784>

Revue de détail des versions anglaises des sites Google, Ask Jeeves, MSN et Lycos.

-> Référencement payant

<http://www.journaldunet.com/0204/020402soumissionstableau.shtml>

Récapitulatif des offres actuelles de référencement payant.

-> Improve Your Rankings by Improving Your Links

[http://www.internetday.com/article/0,,1381\\_998451,00.html](http://www.internetday.com/article/0,,1381_998451,00.html)

Quelques pistes pour améliorer votre indice de popularité...

-> Booléens

<http://library.albany.edu/internet/boolean.html>

Une explication claire des différents opérateurs booléens disponibles dans la syntaxe des outils de recherche.

Merci pour votre lecture... Pour toute suggestion : [oa@abondance.com](mailto:oa@abondance.com)