

■ Comment signaler à un outil de recherche la modification ou l'arrêt d'un site web

Dans cette lettre, nous parlons souvent de référencement de sites (sur les annuaires) et de pages (sur les moteurs) web. Mais il peut arriver qu'on ait besoin de déréférencer une source d'information ou d'indiquer une modification d'un site ou d'une page aux outils de recherche. Les raisons pour cela peuvent être multiples. En voici quelques unes :

- Le site a changé d'url ou de nom de domaine.
- Une page a disparu du site, elle a été éliminée car, par exemple, son contenu était devenu obsolète.
- Le site a cessé son activité (cela arrive, hélas...).
- Le site d'un ancien concurrent (qui a cessé son activité) ressort devant vous lors de la saisie d'un mot clé (sur un annuaire ou un moteur) et vous voudriez bien voir disparaître ce site obsolète pour gagner une place ;-)

Il existe deux grandes familles d'actions à mener pour prendre en compte le déréférencement ou la signalisation d'une modification sur votre site auprès des annuaires et moteurs majeurs : une première famille de points à prendre en compte sur votre site même pour que les outils de recherche prennent bien en compte les modifications demandées, ainsi qu'une deuxième famille d'actions à mener, de façon active, auprès des outils de recherche eux-mêmes.

Actions à mener sur votre site
~~~~~

Si vous désirez que votre site ou certaines de vos pages ne soient plus pris en compte par les moteurs, la première règle à adopter est, bien entendu, d'y insérer un fichier "robots.txt" adéquat ainsi que des balises Meta "Robots" qui indiqueront aux spiders des moteurs ce qu'ils doivent faire : indexer les fichiers ou non, suivre les liens ou non.

Vous trouverez toutes les infos sur ces fichiers et balises ici :

Comment écrire un fichier "robots.txt" :  
<http://docs.abondance.com/robots.html>

Explications des balises Meta "Robots" :  
[http://docs.abondance.com/meta\\_4.html](http://docs.abondance.com/meta_4.html)

En mettant en ligne ces indications sur votre site, vous indiquerez clairement au spider, lors de son prochain passage, ce qu'il doit faire ou ne pas faire. Dans ce cas, il supprimera (s'il est bien programmé ;-)) de son index les pages que vous lui demandez de ne plus indexer. Tous les moteurs majeurs prennent en compte le fichier robots.txt et la majeure partie comprennent les balises Meta "Robots" (certains, comme Google, proposent même une version spécifique de cette balise, voir plus loin). Utilisez-les donc à bon escient et, en quelques semaines (le délai pris par le robot pour revenir visiter vos pages), la situation devrait revenir à la normale.

Autre point qui peut être intéressant : si vous supprimez une page web de votre site, essayez de ne pas la supprimer complètement (en mettant le document HTML en question à la corbeille) et laissez le fichier sur le disque sous son nom original (par exemple <http://www.votresite.com/dossiers/sous-dossier/exemple.html>) en indiquant dans le code HTML une redirection (à l'aide d'une balise Meta "Redirect" ou d'un code Javascript idoïne, voir ci-dessous) vers une autre page de votre site : une page de contenu, le plan du site ou la page d'accueil. Ainsi, en attendant que le spider vienne rafraîchir son index et prenne en compte votre fichier robots.txt et/ou la balise Meta "Robots" de votre page (ce qui parfois prend quelques mois, comme sur Altavista ces derniers temps), l'internaute qui trouvera le lien en question (avant qu'il ne soit effacé de l'index par le moteur) sera redirigé vers une page de votre site et ne sera pas confronté à une erreur 404. C'est toujours plus propre...

Rappelons un code Javascript simple pour effectuer une redirection (il en existe beaucoup, que l'on peut trouver sur un bon site de scripts comme <http://www.allhtml.com/>) :

```
<Script language="javascript">
document.location.href="pagecible.html"
</script>
```

Et, puisqu'on en parle, n'oubliez pas de créer une page d'erreur 404 spécifique de votre site, comportant le plan du site, ce sera toujours ça que le spider du moteur pourra se mettre "sous la dent" (si tant est que les spiders aient des dents...) s'il tombe dessus lors de l'indexation de votre site. Il pourra ainsi, éventuellement, suivre les liens vers les zones les plus importantes de votre source d'information. Tant qu'à faire...

Les actions que nous venons de lister concernent plutôt les spiders des moteurs. En ce qui concerne les annuaires, les pages internes de votre site ne les concernent pas, puisque, en majorité, les annuaires ne vont prendre en compte que l'url de votre page d'accueil. Si celle-ci a changé, n'oubliez donc pas de mettre une redirection automatique de l'ancienne vers la nouvelle version, comme indiqué ci-dessus. Puis, avertissez l'annuaire comme indiqué ci-dessous.

Actions à mener auprès des outils de recherche  
~~~~~

Si vous avez tout prévu sur votre site (ce qui est déjà un point non négligeable...), vous pouvez également avertir les outils de recherche de tout changement sur votre site. Voici, outil par outil, comment faire (lorsque la possibilité est donnée par l'annuaire ou le moteur, bien entendu).

Pour les annuaires
~~~~~

- Si votre site traite toujours du même sujet, des mêmes thèmes, et que les modifications que vous avez effectuées touchent plutôt au contenu proposé (nouveaux articles, nouvelles fonctionnalités, etc.), il y a peu de chance que vous puissiez postuler à l'inscription auprès de nouvelles catégories des annuaires, puisque les thèmes principaux de votre source d'information sont restés globalement inchangés. Vous pouvez peut-être, en revanche, demander une modification du résumé descriptif de votre site si nécessaire et surtout si les nouveautés sont assez notables et "réelles" pour le justifier.

- Si les thèmes traités ont changé en profondeur ou en cible et que votre site, qui parlait par exemple auparavant d'agriculture au sens de l'information en ligne s'est mué en site de vente en direct de produits agricoles pour les professionnels, vous pouvez alors éventuellement demander une modification de la catégorie d'inscription actuelle ou l'ajout dans une nouvelle rubrique. Bien entendu, vous pouvez également en profiter pour demander une modification du résumé descriptif alloué à votre site.

- Si le site en question (le vôtre ou un autre ;-)) n'existe plus ou a changé d'url, c'est une raison suffisante pour le signaler aux annuaires majeurs. ;-)

- Dans les cas exposés ci-dessus (modification de descriptif, d'url, de catégorie et notification d'arrêt d'un site), une procédure est, la plupart du temps, prévue par les annuaires majeurs. Voici ce qu'il faut faire, outil par outil.

-> Annuaires anglophones :

**\*\* Yahoo.com (<http://www.yahoo.com/>)**

L'annuaire Yahoo.com propose une interface vous permettant d'indiquer toute modification survenue sur votre site, à l'adresse : <http://add.yahoo.com/fast/change>

Ce formulaire sert à indiquer toute modification d'URL, mais également une demande de changement de titre, du résumé, une demande de nouvelle catégorie ou une modification de l'adresse e-mail associée à votre fiche descriptive (notez-la précieusement lors du premier référencement). Si vous désirez indiquer que le site a disparu ou a cessé ses activités, indiquez-le dans le champ "commentaires" mis à votre disposition dans le formulaire. En règle générale, ajoutez un petit (ne racontez pas votre vie) texte de commentaires afin de bien expliquer votre démarche et pourquoi vous effectuez cette demande. S'il s'agit spécifiquement d'un lien mort présent dans la base de données de Yahoo!, un formulaire adéquat existe à l'adresse :

<http://docs.yahoo.com/docs/writes/deadlink.html>

Bien entendu, comme d'habitude, vous proposez et Yahoo! dispose selon sa ligne éditoriale. En règle générale, les modifications sont assez rapidement prises en compte. Vérifiez sinon toutes les 3 semaines et resoumettez la modification, au cas où...

**\*\* Looksmart.com (<http://www.looksmart.com/>)**

Comme la soumission d'un site est payante sur Looksmart.com, il est "logique", voire "normal", que la modification d'une fiche descriptive le soit également. Vous trouverez donc un descriptif de l'offre "Express Modify" à l'adresse :

<http://submit.looksmart.com/>

Pour en "profiter", le site doit déjà être inclus dans Looksmart. Ainsi, pour 199 US\$, vous pouvez demander une modification du résumé et pour 149 US\$, demander une catégorie supplémentaire. Dans les deux cas, vous obtenez la réponse en 5 jours ouvrés. Autre possibilité (Looksmart indique que cette possibilité n'est possible que dans les 30 jours qui suivent l'inscription d'un site, mais vous pouvez toujours tenter votre chance...), avec le formulaire "Customer Service Form" qui permet d'indiquer des modifications à effectuer ou des commentaires sur un site donné. Exemple à l'adresse (qui pourrait concourir aux Jeux Olympiques de l'URL la plus longue) :

[http://submit.looksmart.com/info.jhtml;\\$sessionId\\$JGGOSUQAAAFMHQAQGHICORA?lastpg=faq&page=form\\_cust&recipient=Suggest+a+](http://submit.looksmart.com/info.jhtml;$sessionId$JGGOSUQAAAFMHQAQGHICORA?lastpg=faq&page=form_cust&recipient=Suggest+a+)

Description+Correction&sessionParam=r00ABXNyABNqYXZhLnV0aWwSGFzaHRhYmxlE7sPJ5FK5LgDAAJGAAPsb2FkRmFjdG9ySQAjdGhyZXNob2xkeHA%2FQAAAAAAS3cIAAAAAwAAAAAR0AARzeW5kdAACVVN0AARjaGFudAAAbHNob21YnVzdAAEcGFnZXQAA2ZhcXQABmxhc3RwZ3QABm1vZGlmeXg%3D

Si cette URL ne fonctionne pas, un lien est aisément identifiable dans la FAQ (lien vers la FAQ en bas d'écran du site looksmart.com), rubrique "Changes", question "What if the URL or content of my site changes after it is listed in the directory?".

Au pire, tentez d'envoyer un e-mail à l'adresse [feedback@looksmart.net](mailto:feedback@looksmart.net) (voir <http://aboutus.looksmart.com/>) en expliquant le pourquoi du comment de votre requête.

**\*\* Open Directory US (<http://www.dmoz.org/>)**

Dans les catégories de l'Open Directory, en haut d'écran, à côté du lien "Add URL", se trouve un lien intitulé "Update URL". Cliquez dessus, indiquez l'URL du site telle qu'elle est proposée dans l'Open Directory et vous aurez accès à un formulaire de demande de modification, similaire à celui proposé par Yahoo!.

Là aussi, le formulaire doit être utilisé pour modifier une URL devenue obsolète, corriger des éventuelles fautes d'orthographe dans le titre ou le résumé ou suggérer un nouveau titre ou un nouveau descriptif si l'activité du site a changé. N'hésitez pas à lire le texte d'introduction, proposé avant le formulaire de demande de modification. Les remarques qui y sont proposées sont valables pour tous les annuaires...

-> Annuaires francophones :

**\*\* Yahoo! France (<http://www.yahoo.fr/>)**

Comme sa "maison mère", Yahoo! France propose une interface pour vous permettre de lui notifier les modifications survenues sur votre site à l'adresse :

<http://add.europe.yahoo.com/bin/change?chFR>

Les conditions d'utilisation sont les mêmes que pour la version US (voir ci-dessus). Attention : comme il est indiqué, vous ne pouvez utiliser ce service si votre site est référencé dans Yahoo! France depuis moins d'une semaine (sigle "nouveau" encore présent). Si votre site (ou celui de votre concurrent ; -)) n'existe plus, indiquez-le, là aussi, dans la zone "commentaires"... En revanche, nous n'avons pas trouvé de formulaire spécifique pour signaler un lien mort à Yahoo! France.

Vérifiez, là aussi, si les changements ont bien été pris en compte. Si, après plusieurs tentatives, rien n'a changé, envoyez un petit mail - poli - à l'adresse [webmaster@fr.yahoo-inc.com](mailto:webmaster@fr.yahoo-inc.com) en expliquant le problème et en indiquant la date à laquelle vous avez soumis pour la dernière fois le formulaire de modification. Logiquement, le problème se résout dans les heures qui suivent...

**\*\* Nomade.fr (<http://www.nomade.fr/>)**

Nomade.fr prévoyait, jusqu'à il y a peu de temps, une adresse e-mail pour tout changement sur un site web déjà intégré (Pbmodif@sa.nomade.fr), mais elle semble ne plus exister. En tout cas, elle n'est plus signalée sur le site.

A la place, Nomade propose deux adresses e-mail : [inforef@sa.nomade.fr](mailto:inforef@sa.nomade.fr) pour les sites ayant été soumis au travers de l'offre gratuite et [grandevitesse@sa.nomade.fr](mailto:grandevitesse@sa.nomade.fr) pour les sites ayant été soumis au travers de l'offre "TGV" (soumission payante). Tentez votre chance selon le cas qui vous concerne. On peut noter également que Nomade.fr propose une offre de "Modification Express". Cette offre, pour l'instant exclusivement réservée aux partenaires professionnels de l'annuaire, permet d'effectuer des demandes de modifications sur les sites déjà inscrits sur l'annuaire, et garantit un traitement de la demande en 48h et une mise en ligne dans un délai de 8 jours ouvrés. Le prix de cette offre est de 59 euros et de 179 euros pour les sites visant un public adulte ou averti. Les modifications concernent le résumé, les mots-clés et les catégories du site (les mises à jour d'URLs et de coordonnées restent cependant effectuées gratuitement sur simple demande aux adresses e-mail ci-dessus).

**\*\* Looksmart France (<http://www.looksmart.fr/>)**

La procédure est indiquée dans la FAQ à l'adresse :

[http://submit.looksmart.co.uk/submit\\_fa\\_q.asp?isp=btlsfr#content](http://submit.looksmart.co.uk/submit_fa_q.asp?isp=btlsfr#content)

dans la zone "Modifications (Visualisation) / Que faire si l'URL ou le contenu de mon site varie après son inscription dans l'annuaire ?"

Les demandes de modifications doivent être envoyées à l'adresse [expsvc-fr@looksmart.net](mailto:expsvc-fr@looksmart.net). Il n'existe pas de formulaire spécifique en ligne. Tous les échanges se font par mail.

**\*\* Guide de Voila (<http://guide.voila.fr/>)**

Comme pour Looksmart France, la procédure de demande de modification est basée sur l'envoi d'un e-mail à l'adresse [info.gui.gen@fr.voila.net](mailto:info.gui.gen@fr.voila.net). Cette procédure est indiquée sous le formulaire de soumission gratuite de site à l'annuaire.

**\*\* Annuaire de Lycos France (<http://www.lycos.fr/dir/>)**

Pour proposer ou signaler un changement à l'annuaire de Lycos France, le mieux est de contacter directement les netsurfeurs en charge des catégories, dont vous trouverez les coordonnées ici : <http://www.fr.lycos.de/service/contact.html>

Il est alors indiqué, en bas de page :

"Contactez nos surfeurs pour toute question concernant l'annuaire Lycos :

\* Merryll ([sciences@lycos-europe.com](mailto:sciences@lycos-europe.com)) pour les catégories Actualités, Médias , Art, Culture, Célébrités , Féminin , Emploi, Enseignement, Loisirs, Sciences, Techniques , Sciences humaines, Voyage, Tourisme.

\* Dominique ([informatique@lycos-europe.com](mailto:informatique@lycos-europe.com)) pour les catégories Auto-moto, Économie, Entreprise, Informatique, Institutions, Société, Jeux vidéo, Régional, Sports."

A vous d'expliquer au mieux votre demande dans le mail que vous leur enverrez.

**\*\* Annuaire de MSN France (<http://search.msn.fr/>)**

Ici, le cas est simple. Sur la page permettant de soumettre un site, à l'adresse : <http://search2.msn.fr/suggestions/>

Indiquez les modifications que vous désirez effectuer et cliquer sur le bouton-radio "modification" pour bien indiquer que la demande doit s'opérer sur un site déjà inscrit dans l'annuaire. Indiquez également dans la boîte "Remarques" qu'il s'agit bien d'une demande de modification ou d'annulation d'un site.

Pour les moteurs  
~~~~~

Comme nous l'avons déjà explicité plus haut, sur la plupart des moteurs de recherche, il suffira de bien utiliser le fichier robots.txt, la balise Meta "Robots" ainsi qu'une redirection Javascript pour obtenir les effets désirés.

La plupart du temps, il suffira de resoumettre une page obsolète au moteur au travers du formulaire "Add Url". Le spider va prendre en compte l'url demandée, va aller sur votre site et va tenter de l'indexer. S'il voit une erreur 404, il enlèvera la page de son index, tout simplement. S'il voit que la date de dernière modification a été modifiée, il ôtera l'ancienne version pour la remplacer par la nouvelle.

Mais, de toutes façons, sur les moteurs, si vous avez modifié le contenu de vos pages, le moteur, qui met à jour en moyenne tous les mois les pages qu'il détient dans son index, devrait venir effectuer ce "rafraîchissement" lors de sa prochaine mise à jour sans que vous ne fassiez rien. Attention : une nouvelle indexation de votre page suite à une soumission manuelle ("Add URL") n'est pas garantie. Lorsque le spider s'apercevra qu'il détient déjà le document soumis dans sa base de données, il n'est pas dit qu'il effectue automatiquement cette mise à jour. Il peut prendre la décision de laisser ce soin à sa procédure de rafraîchissement automatique (voir ci-dessus). Ceci dit, cela ne coûte rien d'essayer. (si ce n'est quelques minutes de votre temps)

Pour résumer, et pour ce qui est des moteurs, une fois le fichier robots.txt écrit et mis en ligne, les balises Meta "Robots" renseignées et insérées, une bonne page d'erreur 404 proposée, et des redirections adéquates mises en place sur les pages devenues obsolètes, vous pouvez vous contenter d'attendre "tranquillement" les spiders des moteurs. Si vous avez un peu de temps devant vous (cela arrive parfois ;-)), soumettez les urls obsolètes ou ayant été modifiées aux moteurs pour tenter d'accélérer les choses. Mais rien ne dit que ces tentatives seront couronnées de succès... Certains moteurs proposent cependant quelques fonctionnalités complémentaires. Petite revue de détail...

-> Moteurs anglophones :

** Altavista (<http://www.altavista.com/>)
Pas de procédure spécifique (voir <http://addurl.altavista.com/sites/addurl/newurl>).

** Hotbot (<http://www.hotbot.lycos.com/>)
Pas de procédure spécifique (voir <http://www.hotbot.lycos.com/help/addurl#11>).

** Google (<http://www.google.com/>)
Google propose une zone spécifique sur ce point à l'adresse :
<http://www.google.com/remove.html>
Une procédure d'urgence est notamment proposée pour enlever au plus vite une page web de l'index de Google, à l'adresse :
<http://services.google.com/urlconsole/controller/>
Par ce biais, la page en question sera enlevé de l'index de Google pendant 90 jours. Une demande par e-mail (avec création de compte) est effectuée auparavant. Un fichier robots.txt (ou une balise Meta "Robots") adéquat doit se trouver à la racine du site, bien sûr (désolé si vous pensiez, ainsi, enlever les sites web de vos concurrents de l'index de Google ;-)).
Google permet également l'emploi d'un balise spécifique sous la forme :
<META NAME="GOOGLEBOT" CONTENT="NOINDEX,NOFOLLOW">
Elle est identique à la balise Meta "Robots", mais sous une forme spécifique à Google. Rappelons que, si vous ne désirez pas voir votre page archivée par le "cache mémoire" de Google, vous devez indiquer dans votre page la balise :
<META NAME="GOOGLEBOT" CONTENT="NOARCHIVE">

** Lycos (<http://www-english.lycos.com/>)
Pas de procédure spécifique (voir <http://home.lycos.com/addasite.html>).

** Alltheweb (<http://www.alltheweb.com/>)
Pas de procédure spécifique (voir http://www.alltheweb.com/help/FAQ_webmaster.html#changeremove). Les délais assez courts de rafraîchissement de l'index de Fast (9 à 12 jours théoriques) résolvent pas mal de problèmes le plus souvent.
D'autres informations sont disponibles ici :
<http://www.fastsearch.com/support/crawler.asp>

** Northern Light (<http://www.northernlight.com/>)
Pas de procédure spécifique (voir http://www.northernlight.com/docs/gen_help_faq_webmaster.html#remove).

-> Moteurs francophones :

** Altavista France (<http://www.altavista.fr/>)
Pas de procédure spécifique (voir http://help.altavista.com/adv_search/ast_haw_avoiding, en anglais...).

** Hotbot France (<http://www.hotbot.fr/>)
Pas de procédure spécifique (voir <http://www.hotbot.lycos.fr/help/webmasterfaq.html#13>).

** Voila (<http://www.voila.fr/>)
Pas de procédure spécifique (voir http://www.voila.fr/Informations/Recherche/?faq_moteur).

** Lycos France (<http://www.lycos.fr/>)
Pas de procédure spécifique (voir <http://www.lycos.fr/content/help/faq/zoom5.html>).

** Lokace (Antisearch) (<http://www.antisearch.net/>)
Pas de procédure spécifique (pas d'indication sur le site).

■ Comment signaler une tentative de spam à un moteur

Lors de vos pérégrinations sur le Web, vous vous êtes certainement rendu compte de certaines tentatives de spam de la part de certains sites (notamment de la part de vos concurrents, bien sûr, cela ne vous étonne pas de leur part, les vils ;-)). Certains outils de recherche ont mis en place des solutions pour que vous puissiez leur signaler ces tentatives de spam si la page en question est présente dans leur index. En voici une liste.

Pour les annuaires
~~~~~

Les annuaires, référençant des sites sous la forme de fiches descriptives écrites par leurs netsurfeurs, ne sont pas très en prise avec des problèmes de spam. En cas de gros problème dont vous auriez pu vous rendre compte, il est toujours, cependant, possible de le leur signaler en utilisant les voies indiquées dans l'article ci-dessus ("Comment signaler à un outil de recherche la modification ou l'arrêt d'un site web") et en indiquant bien, dans la zone "commentaires" des formulaires ou dans l'e-mail envoyé, la raison du signalement. Attention à ne pas faire de la délation pure et simple de "délits" qui n'en vaudraient pas la peine...

Pour les moteurs  
~~~~~

Dans ce cas, les outils de recherche sont beaucoup plus exposés au spam. Voici une liste de ceux qui indiquent une procédure spécifique de

signalement sur leur site :

-> Moteurs anglophones :

**** Altavista (<http://www.altavista.com/>)**

Plus d'informations sont données aux adresses suivantes :

http://help.altavista.com/adv_search/ast_haw_spam

<http://www.altavista.com/sites/about/termsfuse>

(rubrique "Manipulation of Search Results")

Une adresse e-mail est proposée pour signaler les tentatives de spam à Altavista ou pour obtenir plus d'informations sur la politique du moteur à cet égard : spam-support@av.com. Vous pouvez également envoyer un e-mail à l'adresse search@support.altavista.com avec comme titre "Search Results Manipulation", en expliquant pourquoi vous pensez que les résultats d'Altavista ont été manipulés par des webmasters sans scrupules. Le tout en anglais, bien sûr...

**** Hotbot (<http://www.hotbot.lycos.com/>)**

Rien sur ce point, a priori, sur le site de Hotbot, mais celui d'Inktomi (qui fournit les résultats "moteur" à Hotbot) nous en dit plus ici :

http://www.inktomi.com/products/search/content_policy.html

La page décrit bien ce qu'Inktomi considère comme étant du spam, et vous pouvez, par la suite, signaler tout problème à l'adresse spamcrusader@inktomi.com.

**** Google (<http://www.google.com/>)**

Vous pouvez signaler tout problème à Google en envoyant un e-mail à l'adresse spamreport@google.com en explicitant la raison de ce signalement.

Vous pouvez également effectuer un signalement au formulaire adéquat à l'adresse : <http://www.google.com/contact/spamreport.html>

Ce formulaire est intéressant, car, là aussi, il permet d'indiquer quel type de spam vous dénoncez : texte invisible, répétition de mots clés, etc. On a ainsi une meilleure idée de ce que Google considère comme étant du spam.

**** Lycos (<http://www-english.lycos.com/>)**

Lycos propose un formulaire de signalement de spam ici :

http://help.lycos.com/LycosHelp/help/watchdog/htdocs/lycos_watchdog_form.htm

Mais il semble plutôt être spécialisé dans le spam par e-mail par des personnes issues de ses sites de communautés, moins le spamdexing du moteur.

Sinon, essayez le formulaire général de la recherche d'information, à l'adresse :

http://help.lycos.com/LycosHelp/help/search/htdocs/search_form.htm

En choisissant le thème "Comments/Suggestions" dans le menu déroulant intitulé "Your Topic".

**** Alltheweb (<http://www.alltheweb.com/>)**

Pas de procédure spécifique a priori. Tentez un message sur le formulaire général à l'adresse :

<http://www.alltheweb.com/feedback.php>

**** Northern Light (<http://www.northernlight.com/>)**

Pas de procédure spécifique a priori. Tentez un message sur le formulaire général à l'adresse :

http://www.northernlight.com/docs/gen_help_comments.html

-> Moteurs francophones :

**** Altavista France (<http://www.altavista.fr/>)**

Le site Altavista France renvoie vers les mêmes pages et la même adresse e-mail que le site US (voir ci-dessus).

**** Hotbot France (<http://www.hotbot.fr/>)**

Rien à ce sujet sur le site HotBot France, mais les résultats étant fournis, là aussi, par Inktomi, voir ce qui était dit sur HotBot US, ci-dessus.

**** Voila (<http://www.voila.fr/>)**

Rien à ce sujet sur Voila. Le mieux est certainement d'envoyer un e-mail à l'adresse webmaster@fr.voila.net en explicitant de façon précis el'objet de votre signalement.

**** Lycos France (<http://www.lycos.fr/>)**

Rien à ce sujet sur Lycos. Le mieux est certainement d'envoyer un e-mail à l'adresse webmaster@lycos.fr en explicitant de façon précis el'objet de votre signalement.

**** Lokace (Antisearch) (<http://www.antisearch.net/>)**

Rien à ce sujet sur Antisearch. Le mieux est certainement d'envoyer un e-mail à l'adresse antibot@antidot.net en explicitant de façon précis de l'objet de votre signalement.

■ Nouveaux formats indexés sur Google

Vous en avez certainement entendu parler (la nouvelle a fait grand bruit...), Google indexe depuis peu de nouveaux formats de fichiers, dont notamment :

- PDF (.pdf), mais c'était déjà le cas depuis pas mal de temps (à ce jour, 22 millions de documents sont indexés par Google sous ce format).
- Word (.doc)
- Excel (.xls)
- Powerpoint (.ppt)
- Rich Text Format (.rtf)
- Postscript (.ps)

Ainsi, il est possible de visualiser des documents autres qu'au format HTML grâce à Google. Ces documents sont visualisés de deux façons différentes par le moteur :

- Soit en indiquant, entre crochets, le format en question, à gauche du titre :

[PDF] Investment Strategy

ou :

[PPT] \$6 Million Common Stock Offering

- Soit sous la forme d'une phrase affichée sous la ligne du titre :

File Format: PDF/Adobe Acrobat - View as HTML

ou :

File Format: Unrecognized - View as HTML

Cette dernière indication ("View as HTML") est intéressante : elle permet de lire un fichier Powerpoint, par exemple, au format HTML, car Google les

convertit auparavant en HTML, comme les fichiers PDF et autres.

Google permet, par la fonction "filetype:", d'effectuer une recherche spécifique sur un format donné. Exemple :

secret filetype:pdf
confidential filetype:ppt

Bien entendu, si vous désirez restreindre votre recherche uniquement aux fichiers HTML, tapez :

faq filetype:html

Rappelons : - Que l'espace est considéré comme un ET par Google. Pas la peine, donc, d'utiliser le "+" (+faq +filetype:html), qui est ici inutile.
- Que les mots clés ci-dessus ("secret" et "confidential") ne sont pas innocents puisque vous avez certainement entendu parler des petits problèmes qu'ont eu certains sites, américains notamment, qui ont vu leurs documents Word plutôt confidentiels disponibles du jour au lendemain sur Google, car ils avaient oublié de les protéger, un lien vers ces fichiers demeurant disponibles sans sécurité sur une de leurs pages web. Pas très malin... Si cette mésaventure vous arrive (on ne sait jamais...), prenez en compte la procédure d'urgence proposée par Google pour éliminer un lien de son index (voir premier article de cette lettre : "Comment signaler à un outil de recherche la modification ou l'arrêt d'un site web").

Si vous voulez savoir les fichiers aux formats "exotiques" référencés sur votre site par Google, utilisez la fonction "site:", par exemple :

site:www.abondance.com filetype:pdf
soit l'url :
http://www.google.com/search? hl=en&q=site%3Awww.abondance.com+filetype%3Apdf

Et vous aurez accès à la liste de tous les fichiers PDF, présents, dans cet exemple, sur le site www.abondance.com et indexés par Google. Intéressant pour savoir si vous devez faire le ménage ou non ;-)

Vous pouvez également demander plusieurs formats d'un seul coup. Exemple :

filetype:ppt OR filetype:doc site:www.defense.gouv.fr
soit l'url :
http://www.google.com/search? hl=en&q=filetype%3Appt+OR+filetype%3Adoc+site%3Awww.defense.gouv.fr

Mais Google ne semble pas indexer que les formats "officiellement" déclarés ci-dessus. Si l'on fait des tests, on s'aperçoit que les formats suivants sont également pris en compte :

- Lotus 1-2-3 (.wk1, .wk2, .wk3, .wk4, .wk5, .wki, .wks, .wku)
Exemple : http://www.google.com/search?q=analysis+filetype%3Awk1
- Lotus WordPro 97 (.lwp)
Exemple : http://www.google.com/search?hl=en&q=analysis+filetype%3Alwp
- MacWrite (.mw)
Exemple : http://www.google.com/search?hl=en&q=internet+filetype%3Amw
- Microsoft Works (.wks, .wps, .wdb)
Exemple : http://www.google.com/search?hl=en&q=internet+filetype%3Awks
- Microsoft Write (.wri)
Exemple : http://www.google.com/search?hl=en&q=internet+filetype%3Awri
- Texte brut (.ans, .txt)
Exemple : http://www.google.com/search?hl=en&q=tools+filetype%3Atxt
- Mais également Flash (.swf)
Exemple : http://www.google.com/search?hl=en&q=faq+filetype%3Aswf

Cette nouvelle est quand même d'importance et représente un nouveau pas dans l'indexation des fichiers Flash, notamment, qui pullulent actuellement sur le Web. Bonne nouvelle donc...

La mauvaise nouvelle, en revanche, est le fait que, on le sait bien, de nombreux fichiers Word et Excel, notamment, sont infestés de virus. Donc, donner un "libre" accès à des milliers de fichiers de ce type n'est-il pas dangereux ? Si Google ne dispose pas d'un anti-virus (et rien ne dit, à notre connaissance, qu'il passe dans un premier temps, les fichiers .doc et .xls indexés à ce type de contrainte technique), n'y a-t-il pas à un risque évident de voir des virus se propager rapidement au travers du Web ? Le danger est en tout cas réel. L'avenir dira si cette crainte était fondée ou non...

■ Bilan de l'année 2001, par les acteurs et experts du domaine de la recherche d'information sur le Web

L'année 2001 se termine. Elle a connu de nombreux bouleversements dans le monde des outils de recherche. Pour effectuer un premier bilan des mois qui viennent de s'écouler, nous avons demandé à plusieurs acteurs du domaine (et pas obligatoirement ceux dont vous avez l'habitude de lire les réactions ou indications dans cette lettre) de répondre à trois questions :

Question 1 : L'année 2001 a vu la disparition de certains outils de recherche, des difficultés pour d'autres mais aussi l'apparition de nouveaux acteurs. Quel regard portez-vous sur cette année qui se termine, dans le domaine des annuaires et moteurs sur le web ? Comment qualifieriez-vous cette année : transition ? Evolution ? Révolution ? Ou autre ?

Question 2 : Le modèle économique basé en grande partie sur la publicité semble aujourd'hui marquer le pas, du moins en partie. Les outils de recherche sont actuellement en quête de nouvelles voies pour être moins dépendants de l'unique marché de la pub en ligne. Quelle est votre opinion sur ce point en 2002 ? Quelles possibilités imaginez-vous de la part des outils de recherche pour capter de nouveaux revenus, de nouveaux marchés dans les mois qui viennent ?

Question 3 : Un peu de prospective : Quel sera, d'après vous, le paysage des outils de recherche dans les mondes anglophones et francophones dans les deux à trois ans qui viennent ? Quelle sera, à votre avis, l'évolution la plus marquante dans ce domaine ? Et les freins éventuels les plus importants ?

Voici leurs réponses (merci à tous ceux qui ont joué le jeu et nous ont fait parvenir leur témoignage) :

Eileen Rodriguez, Google (<http://www.google.com/>)
~~~~~

1) 2001 a été une année d'évolution pour l'industrie des outils de recherche. Les moteurs qui ont connu le succès ont continué à évoluer en proposant de plus en plus de possibilités de recherche. En tant que plus important, quantitativement parlant, moteur de la planète, Google propose actuellement un index de plus de 1,6 milliards de pages. Cette année, Google a également ajouté de nouveaux formats et de nouvelles sources d'information à ses résultats de recherche (NDLR : voir article ci-dessus), comme les documents électroniques aux formats PDF, Microsoft Office et bien d'autres. De nouvelles versions linguistiques de l'outil ont également vu le jour (Google est aujourd'hui disponible en 66 langues en tout),

incluant le grec, l'arabe, l'hébreu, etc. Nous avons également concentré notre effort sur de nouvelles possibilités d'accès à l'information, notamment le "sans fil" grâce à des partenariats avec des sociétés comme Vodaphone, Handspring, Sprint PCS et AT&T Wireless. Et nous continuons à effectuer des recherches dans de nouvelles voies, notamment dans le cadre de la reconnaissance vocale. Notre but est toujours de délivrer des possibilités efficaces de navigation et de recherche sur le Web et de maintenir notre position de leader dans le monde.

2) Avant tout, il est important de comprendre comment Google draine ses revenus. Notre business model est basé sur deux types de revenus : la publicité en ligne et les services de recherche (en fournissant notre technologie à des clients comme Yahoo!, AOL/Netscape, Sony ou Cisco). Nos revenus sont équitablement répartis à 50/50 entre les deux sources. L'un n'est donc pas dépendant de l'autre et réciproquement. Google a également récemment annoncé que l'entreprise dégagait des profits depuis le début de l'année 2001. Ainsi, les deux sources de revenus se portent bien et connaissent une croissance continue chaque trimestre. En 2002, notre but sera de continuer sur cette voie et de proposer, encore et toujours, un meilleur outil de recherche aux internautes du monde entier et à nos clients. L'approche de la publicité en ligne vue par Google a prouvé son efficacité, avec des taux de clic qui sont, le plus souvent, cinq fois supérieurs à la moyenne constatée sur d'autres sites. Bien que notre modèle économique ait été au départ géographiquement basé sur le marché américain, nous prévoyons une expansion beaucoup plus mondiale en 2002. Nous avons déjà obtenu quelques succès sur des marchés internationaux clés, comme le Japon, l'Allemagne et la Grande-Bretagne, qui nous encouragent à persévérer dans cette voie.

En continuant de maintenir un business sur ces deux voies de revenus, nous sommes optimistes sur le fait que 2002 soit une année pleine de succès pour Google.

3) D'ici deux ou trois ans, nous essaierons d'intégrer de la recherche d'information dans tout ce que nous faisons à tous les niveaux. L'objectif le plus important est de faire en sorte que les outils de recherche soient accessibles à tout le monde, dans tous les pays et sur tous les supports. Il faudra, pour cela, tenir compte de la croissance rapide de l'Internet et amener de nouvelles sources d'information à nos utilisateurs tout en délivrant un outil simple à utiliser et efficace en nous servant de notre expérience du domaine.

François Bourdoncle, PDG d'Exalead (<http://www.exalead.com/>)

~~~~~

1) Je dirais qu'il s'agit d'une transition vers un monde où le "métier" de la recherche et le "métier" des portails sont enfin clairement séparés. L'un relève de l'édition logicielle, l'autre d'un métier de média. Mélanger les deux a conduit des gens comme AltaVista à ne faire bien ni l'un ni l'autre.

2) Pour le modèle de revenus, je crois beaucoup à la recherche comme "produit d'appel", c'est-à-dire à un produit sur lequel on ne gagne pas d'argent mais qu'il est indispensable d'avoir. Seuls les portails avec une vraie source de revenus autre que le "search pur" pourront survivre (peut-être des gens de type Google pourront survivre, mais survivront-ils à une guerre des prix sur l'offre de search par exemple? ou à une offensive d'un grand acteur décidant de faire du dumping?). Les éditeurs logiciels, d'autre part, ont un "business model" plus traditionnel.

3) Je crois que les deux métiers portail/éditeur logiciel vont se consolider et que cela va avoir des conséquences sur le coût de production des services de recherche, avec une plus grande rationalisation et une recherche d'économies. La course au nombre de documents indexés va rendre ce point particulièrement important (et placer Exalead en bonne place, of course ;-). Les fonctionnalités de recherche et de navigation, de cross-selling, de ciblage de la pub vont d'autre part devenir aussi de plus en plus importants, en particulier pour les méga-portails dont la concentration va se poursuivre à vive allure. Là aussi, Exalead devrait avoir son mot à dire...

En terme d'évolution marquante, je crois que le "prêt à porter" de la recherche verra sa suprématie mise à mal par de nouvelles technos plus ciblées, plus "sur mesure" prenant en compte la politique éditoriale des portails et leur galaxie de partenaires. Et là aussi.... ;-))

Jean-François Marti, DG Adjoint de Nomade.fr (<http://www.nomade.fr/>)

~~~~~

1) Je qualifierais cette année de "période de structuration". Le ralentissement du marché publicitaire aura été fatal à plusieurs outils de recherche. Les outils ayant su, ces dernières années, conquérir des parts de marché et proposer donc l'audience la plus élevée se sont partagé l'essentiel des recettes publicitaires et ont donc pu assurer leur survie.

2) La bataille aujourd'hui ne porte plus en priorité sur les parts de marché mais sur l'atteinte au plus vite de l'équilibre financier. Atteindre la rentabilité passe obligatoirement par la réussite de deux objectifs :  
- Prouver l'efficacité de l'outil de recherche en tant que support publicitaire puisque la publicité en ligne demeurera la principale source de revenus.  
- Développer de nouvelles sources de revenus en dehors de la publicité. Tout d'abord auprès des entreprises avec les services de référencement payant. Et auprès des utilisateurs avec le développement de services en ligne payants. Autant le premier modèle est en train de faire ses preuves, autant le second est aujourd'hui encore un point d'interrogation.

3) A mon avis, l'Europe verra la coexistence de 2/3 portails "internationaux", d'origine nord-américaine, présents dans chaque pays, avec celle de 2/3 acteurs locaux de la recherche.

La tendance, aujourd'hui favorable aux moteurs nord-américains, tendant à s'inverser au profit des acteurs locaux dans les années à venir. Les acteurs locaux profiteront d'une plus forte réactivité et d'une spécialisation sur leur marché. Donc d'une meilleure "localisation" de leur offre. Par ailleurs, je pense que les outils de recherche "humains" prendront le pas sur les "robots" en proposant des réponses plus éditoriales. Ces réponses "éditoriales" correspondant mieux aux demandes d'une population d'internautes plus large dans les années à venir (mais peut-être ne me trouverez vous pas très objectif sur cette réponse ;))

Thibaut Deshayes, Webmaster du site Asktibbs (<http://www.asktibbs.com>)

~~~~~

1) Je pense que l'année 2001 est une année sans pitié (une de plus ?) où les outils de moindre importance auront connu de réels problèmes. Seuls subsistent ceux qui sont ancrés depuis des années dans le monde de la recherche (type Altavista, Yahoo!) ou d'autres qui ont vraiment innové et qui poursuivent dans ce sens (Google).

Des outils majeurs d'il y a quelques années (Altavista, Yahoo, NorthernLight, Lycos, Excite, Webcrawler, Infoseek, Hotbot), on peut remarquer que ceux qui se trouvent aujourd'hui en difficulté ou ont disparu sont ceux qui n'ont pas su innover d'un point de vue technologique ou qui n'ont pas su offrir des index en proportion avec la taille grandissante du Web.

Du côté des annuaires, même Yahoo a connu des problèmes malgré son avantage d'avoir été le premier à proposer ce type de répertoire en ligne. On assiste à un monopole de la part de Yahoo, de Looksmart et l'Open Directory, qui outre leurs services propres, fournissent les moteurs de recherche avec leurs annuaires.

Cette année n'a pas été vraiment révolutionnaire. La dernière vraie révolution est à mon avis l'arrivée de Google, qui tient la palme de l'innovation. Je qualifierais cette année comme une année de transition où seuls les problèmes financiers ont, malheureusement, alimenté l'actualité des outils de recherche.

2) La publicité en ligne est apparemment trop peu génératrice de revenus. Les outils imaginent les moyens de récolter de l'argent comme ils peuvent. Le référencement payant, le positionnement payant, l'achat de mots clés sont à mon avis dommageables pour l'utilisateur. De nombreux sites perso ou à petit budget ne pourront pas se permettre de telles dépenses et connaîtront probablement des chutes de trafic. C'est un peu contre l'esprit du net mais il faut bien vivre. Personnellement je n'apprécie pas trop le fait que l'argent influence les résultats que peuvent fournir les outils de recherche, surtout que cela est souvent fait de manière maladroite. On voit également les problèmes juridiques que cela peut éventuellement poser au niveau du droit de la concurrence.

Le succès de l'Open Directory Project est un beau pied de nez à toutes les offres payantes qui fleurissent ici ou là.

3) Je pense que la tendance actuelle et les difficultés que connaissent les outils de recherche vont se poursuivre. Depuis Google, on n'a pas vu grand chose de nouveau et je pense que le salut peut venir des milieux universitaires. On nous annonce l'arrivée imminente d'Hyperbee (<http://www.hyperbee.com/>), outil sur lequel on peut fonder beaucoup d'espoir.

Issu d'un projet universitaire on se dirige de plus en plus vers des outils basés sur l'interactivité homme / machine. Si l'on se rappelle que Yahoo et Google sont eux aussi issus du génie de quelques étudiants de l'université de Stanford...
Seule l'innovation technologique ou de gros moyens financiers peuvent garantir le succès d'un outil de recherche.

Laurent Baleydière, fondateur et PDG du métamoteur Kartoo (<http://www.kartoo.com>)

1) C'est l'année du retour du système D. Le contexte économique a favorisé ceux qui ont misé sur la technologie et le bouche à oreille. L'internaute débutant n'est plus influencé par les spots TV et entend dire "va faire tes recherches avec Google, ça marche bien". L'internaute moyen a pu tester différents moteurs et a choisi Google parce que c'est rapide, simple et souvent satisfaisant. L'internaute expérimenté utilise Google depuis sa sortie. Mais attention, on sait aussi maintenant que le paysage des moteurs de recherche peut changer très rapidement. Pour l'instant, Google gère bien sa petite avance : le programme est bien conçu, il supporte à la fois une grosse base de pages et un trafic monstrueux, on peut lui greffer des innovations facilement et en assurer la maintenance.

2) Je crois que la publicité en ligne a toujours un bel avenir. Le bandeau publicitaire classique ne trompe plus l'utilisateur moyen qui a souvent été déçu en cliquant dessus. Du coup, on propose des liens textuels un peu noyés dans la liste des résultats des moteurs... Les moteurs de recherche restent de véritables mines d'or : des sites web très visités sur lesquels les gens laissent facilement leurs centres d'intérêt. On peut faire des milliers de choses avec ça.

3) Chez KartOO, ce n'est pas une surprise, on mise sur l'interface et la convivialité. Google ressemble encore trop à un outil fait par les informaticiens pour les informaticiens. La solution pour le rendre accessible est celle du grand vide : on supprime tout ce qu'il y a autour de la case de recherche. C'est efficace, mais on peut peut-être faire mieux, non ?

Emmanuelle Siary, Directrice du Développement, Looksmart France <http://www.looksmart.fr/>

1) Il s'agit d'une évolution. Le passé a montré que de nouvelles technologies apparaissent chaque année et qu'un nouvel entrant totalement inconnu pouvait détrôner un grand acteur bien établi. 2001 a vu Google détrôner Inktomi. En 2002, il se peut très bien qu'un nouveau moteur détrône Google. Le grand défi de chaque acteur n'est pas d'offrir une technologie qui offre des résultats plus pertinents que le précédent mais bien d'avoir aussi un modèle économique qui tienne la route, lui permettant de générer des revenus et des profits rapidement. En conclusion, il s'agit bien d'une évolution :

- Des nouvelles technos.
- Des mentalités : il ne s'agit pas uniquement de prendre des parts de marché mais de trouver un équilibre financier et une rentabilité comme toute entreprise qui se respecte.

2) Il est très clair que les produits publicitaires graphiques ne suffisent plus à faire vivre un service de recherche. LookSmart a été très précurseur sur le sujet et a été le premier à lancer en France :

- Le référencement payant (avec notre offre Express).
- Le référencement de sous-page avec un modèle au clic (Subsite).

Le seul moyen de vivre aujourd'hui est de s'orienter vers un modèle économique similaire à celui des pages jaunes. C'est-à-dire pouvoir allier en harmonie des informations pertinentes à la requête de l'internaute et en même temps offrir un positionnement au site qui désire une prépondérance.

3) 2002 sera décisif. Beaucoup de "petits" moteurs disparaîtront. Ceux qui resteront auront su monétiser leur trafic. Les critères clés de succès étant :

- La couverture du marché
- La pertinence des résultats
- Les produits publicitaires adaptés à la recherche orientés vers un modèle de type "Pages Jaunes".

Julien Pansu & Christophe Raynaud, Fondateurs de netboussole (<http://www.netboussole.com>)

1) 2001 fut...

- Une rupture financière avec la chute de nombreux acteurs asphyxiés par un marché publicitaire fuyant et la recherche incessante de financements voire de nouveaux modèles économiques. A côté de géants comme AOL (avec sa plate-forme propriétaire), Yahoo!, MSN, Lycos qui semblent avoir les reins assez solides pour passer "l'hiver numérique", beaucoup d'acteurs ont disparu ou ont connu de graves difficultés (Go, Excite, Altavista...) 2001 restera une année sombre, d'autant plus qu'un certain nombre d'outils de nouvelle génération (P2P notamment), même s'ils apportent un nouvel élan et de nombreuses innovations, n'ont pas encore pu faire leurs preuves.

- Une rupture de positionnement des outils traditionnels avec l'avènement de nouveaux outils focalisés uniquement sur la technologie de recherche d'information. Google fonde son succès sur trois avantages clés pour les outils de recherche : pertinence, simplicité et rapidité. Le modèle de "portail" inventé par Yahoo! et repris par tout le monde semble ainsi révolu. Malgré le succès en termes de trafic des plates-formes multi-services (shopping, actualité, services pratiques, agenda en ligne, e-mails...), la rentabilité économique n'est pas évidente.

2) Il est évident que l'élément marquant de cette année 2001 fut l'effondrement brutal du marché de la publicité en ligne.

Une question fondamentale est de savoir si ce ralentissement provient du support lui-même ou de campagnes publicitaires malheureusement au service de produits à faible valeur ajoutée et qui, par un taux de transformation dérisoire, ont décrédibilisé ce marché. Pendant de longs mois, de nombreuses sociétés ont uniquement vécu par la diffusion de campagnes publicitaires elles-mêmes lancées par d'autres sociétés financées par la publicité. Ainsi, ces sociétés se sont mutuellement et artificiellement fait vivre et cela sans la moindre recherche de taux de transformation, de service apporté et encore moins de satisfaction du client... seul comptait le sacré PAP ! A ce titre, et en vue des années à venir, cet effondrement apparaît bénéfique et même nécessaire pour l'assainissement et le rebond de ce marché. Animée par des sociétés et des produits à forte valeur ajoutée, il est fort probable que la publicité en ligne retrouve progressivement crédibilité et efficacité.

Mais au-delà des campagnes elles-mêmes, il est indispensable que les supports et principalement les outils de recherche concentrent leurs efforts sur un ciblage et une personnalisation efficaces. Dans ce sens, l'intégration de liens sponsorisés en haut de certaines pages de résultats - sous réserve d'une indispensable corrélation entre les deux - semble être une possibilité intéressante tant pour l'annonceur que pour l'utilisateur.

Personne n'a jamais été choqué par l'insertion d'encarts de grande dimensions au sein des pages jaunes et nombreux sont les utilisateurs portant leur choix sur ces derniers...

A cette première évolution, s'ajoute également le système de référencement payant devenu en quelques mois le nouvel Eldorado de pratiquement tous les outils de recherche. Si le système apparaît judicieux, financièrement prometteur et intéressant pour de nombreux professionnels en quête de rapidité, il est indispensable que les outils de recherche continuent de proposer les deux systèmes.

Dans le cas contraire - et la tentation pour les outils de recherche de privilégier l'attribution des ressources financières au référencement payant risque d'être grande - de nombreuses thématiques risquent de décrépir à très grande vitesse au profit d'autres sur-médiatisées et très courtisées. Or, l'utilisateur reste le moteur de ce business. Si ce dernier ne trouve plus l'information, il ira très vite voir ailleurs (Open Directory par exemple) et l'outil de recherche, focalisé uniquement sur une offre de référencement payant, verra très rapidement son trafic chuter et son avenir s'assombrir à la vitesse d'un bouche à oreille désastreux!

3) Au-delà d'une logique de concentration propre à tout marché en phase de maturation, il est fort probable d'assister prochainement à une fusion logique entre navigateurs et outils de recherche. Si le navigateur permet à l'heure actuelle de surfer à l'intérieur d'un site, il intégrera prochainement toute la panoplie des outils de recherche habituels permettant ainsi de naviguer de site en site. Depuis quelques mois, tous les grands annuaires et moteurs (Voilà, Google, Nomade, Yahoo, ...) proposent des barres de navigation directement intégrées au navigateur. Ces barres se limitent pour l'instant à rediriger l'utilisateur vers les pages du site web mais une voie s'est ouverte. Le "tout en un" semble prochainement incontournable. A titre d'exemple, MSN Explorer évolue rapidement dans ce sens, les dernières versions de Netscape intègrent certaines fonctions de recherche, Yahoo! lance un navigateur Internet Explorer intégrant directement ses services et la dernière version de Netboussole s'intègre complètement dans le

navigateur.

Karl Gregory, European Marketing Director, Altavista (<http://www.altavista.com>)

1) L'année 2001 a été une année difficile pour toutes les sociétés de technologie, pas seulement les moteurs de recherche. Cependant, l'industrie Internet est encore jeune, quand on la compare avec des médias plus établis comme la télévision. Je crois que nous sommes encore dans une étape d'évolution. Du point de vue d'un moteur de recherche, le "search" est la deuxième activité la plus populaire des internautes sur le Web. La quantité de contenu disponible sur le Web croît exponentiellement et les utilisateurs auront de plus en plus accès à cette information. Les outils de recherche comme Altavista devraient bénéficier de cette croissance à l'avenir.

2) La publicité en ligne représente encore une part importante de notre business, et Altavista a récemment proposé de nouveaux et innovants formats publicitaires, comme le "skyscraper", qui permettent à l'internaute de chercher directement dans les bases de données de l'annonceur depuis le bandeau publicitaire lui-même. En plus du fait que les internautes continueront à rechercher l'information au travers de mots clés, la force d'Altavista est d'offrir des campagnes basées sur des mots clés ciblés et très efficaces.

En termes de nouveaux revenus, nous proposons également des programmes de référencement payant comme Express Inclusion, qui ont été développés par notre équipe marketing et lancés dans un premier temps aux Etats-Unis. Ce type de programme de référencement payant est intéressant car ils accélèrent le délai de prise en compte d'un site web dans l'index du moteur.

3) Altavista investit dans la recherche d'information depuis 1995. Nous sommes des vétérans du domaine. Nous pensons que la quantité de contenu sur le Web continuera à croître, comme continueront à grandir les moteurs de recherche qui se nourriront de ce contenu. Malheureusement, nous ne pouvons révéler nos plans sur les deux ou trois prochaines années. Cependant, nous pensons que les moteurs de recherche offriront de nombreux, nouveaux et puissants outils pour aider les internautes à trouver de façon pointue l'information qu'ils désirent. La difficulté sera de prendre en compte les différences dans les cultures locales. Par exemple, une recherche sur le mot "football" aux Etats-Unis amènera (et devra amener) des résultats très différents pour la même recherche en France. Les outils de recherche qui reconnaîtront de façon intelligente ce type de différences auront gagné un point par rapport à leurs concurrents.

Henrik Hansen, Marketing Director Search, Inktomi (<http://www.inktomi.com/>)

1) Nous voyons une transition dans le sens où le but du vendeur de solutions de recherche sur le Web est de mettre en place un passage depuis les solutions strictement web vers le monde des intranets car de plus en plus d'applications basées sur les technologies internet sont déployées par les entreprises sur leurs intranets. En même temps, le marché pour les outils de recherche sur le web lui-même ("public") stagne depuis quelques temps.

2) Les sociétés et leurs personnels voudront de plus en plus une information "utile". Cela signifie qu'ils vont s'efforcer d'aller vers des outils d'accès simple, efficaces et homogènes qui puissent apporter une information précise à la fois aux personnes et aux applications d'entreprise. Le capital d'information à l'intérieur d'une entreprise ne génère pas de valeur ajoutée tant qu'il n'est pas identifié et appliqué. Ceci signifie qu'il faut prendre en compte les données émanant des employés, des partenaires commerciaux et des fournisseurs tout en s'assurant que ces données nourrissent de nouvelles applications d'entreprise.

3) Plus de sécurité, de personnalisation et de catégorisation.

Béatrice Foenix-Riou, responsable de la revue Netsources (<http://www.fla-consultants.fr/>)

1) Pour ma part, je choisirai le terme de "transition" pour qualifier l'année 2001. En fait, si je devais dégager les "tendances de l'année 2001" qui me semblent les plus intéressantes pour les outils de recherche, je reprendrais pratiquement celles que j'avais données dans Netsources pour conclure l'année 2000 ! A savoir :

- La course à "l'index le plus grand" continue : les moteurs de recherche semblent mettre ce point dans leurs priorités ; Google se démarque désormais avec une indexation d'une partie du web invisible, à savoir les documents dans d'autres formats que html : après le PDF, les documents Word, Excel et Powerpoint sont aujourd'hui indexés, ce qui a suscité quelques émois chez certains serveurs découvrant sur le web des documents qui n'auraient pas dû y être... On notera que Google ne se contente pas d'enrichir sa couverture. S'il a conquis nombre de netsurfers par sa simplicité, il a notablement amélioré ses possibilités de recherche cette année. Mais où s'arrêtera-t-il ?

- Le principe du référencement payant (incluant à la fois le soumissionnement, l'indexation et le positionnement payant - pudiquement appelé emplacement publicitaire par certains outils -) se généralise : seuls les outils anglophones le proposaient il y a un an ; les outils de recherche francophones suivent le mouvement aujourd'hui. Il faut que le netsurfer prenne conscience de cette évolution, surtout s'il utilise un métamoteur (Copernic, Dogpile...); les premiers résultats ne sont pas forcément les plus pertinents...

- Les portails thématiques (ces fameux "vortals") se multiplient, dans tous les domaines, et constituent des sites de référence sur de multiples sujets : chimie (www.chemindustry.com), business (www.business.com), ingénierie (www.eevl.ac.uk)... C'est aujourd'hui dans cette catégorie que ce créent le plus souvent les nouveaux outils...

2) Les systèmes de référencement payant (au sens large) risquent fort de se développer et de se généraliser. L'offre de Yahoo.com semble avoir du succès, puisque le prix de la soumission est passé de 199 \$ à 299 \$...

D'après CNet, Google réfléchit à des versions "verticales" de son moteur accessibles sur abonnement ; de telles solutions seront peut être adoptées en 2002 et, si Google s'y met, on peut penser que les autres outils suivront le pas rapidement.

3) Pour le moment, les grands annuaires et moteurs cherchent à avoir la "meilleure" couverture du Web. Dmoz recense plus de 3 millions de sites, soit plus de 30 % de l'offre, si l'on se fie aux estimations d'OCLC. Le pourcentage des pages indexées est du même ordre de grandeur pour Google. La question que je me pose est : les grands annuaires et moteurs vont-ils pouvoir suivre dans les années qui viennent le développement du Web ? et surtout, est-ce souhaitable ?

Est-ce qu'un annuaire de 6 millions de sites, par exemple, aurait un sens ? Le visiteur aurait toute les chances de s'y perdre ou de n'identifier qu'une petite partie des sites qui répondent à sa requête.. Si les outils de recherche souhaitent maintenir leur couverture du Web, ils devront enrichir leurs possibilités de recherche. Mais le succès de Google semble démontrer que c'est surtout la simplicité qui plaît aux internautes.

Aussi, on risque fort d'aller vers un développement des outils "thématiques", qui recenseront tous les sites (ou toutes les pages des sites) dans un domaine. Cette fragmentation des portails permettra aux annuaires d'offrir une arborescence plus précise et donc plus détaillée.

D'autre part, "l'internet payant" se développe, et les sites professionnels accessibles sur abonnement se multiplient. Les outils de recherche vont peut être suivre la tendance et l'exemple de Northern Light (version "light" en accès libre, et version complète sur abonnement par exemple...); ils pourraient alors offrir, en complément de l'accès libre à une partie de leur base, un certain nombre de services payants (envois d'alertes, possibilités de recherche plus sophistiquées...)

Christelle OTT, Directrice Marketing et Commerciale, Antidot - AntiSearch (<http://www.antisearch.net/>)

1) Selon nous, il ne s'agit pas vraiment d'une révolution mais d'une phase importante d'évolution. Après une première vague de moteurs (Inktomi, Altavista...), on ne peut faire qu'un constat d'insatisfaction tant du point de vue des professionnels que des utilisateurs, qui a entraîné l'explosion de Google.

Parallèlement on assiste à l'émergence de nombreux projets de recherche débutés pour la plupart il y a 2 ou 3 ans, et qui commencent à sortir maintenant. Il est extrêmement clair que le marché n'a jamais été aussi peu structuré dans la mesure où les "incontournables" de ces dernières années, n'ont a priori pas su évoluer avec le volume croissant d'informations et la complexité des nouveaux formats. Aux nouveaux acteurs de proposer à présent des outils de recherche performants, sémantiques ou (et) statistiques.

La prochaine étape sera peut-être la prédominance de l'une des solutions ? Du rapprochement des uns et des autres... Bref une étape de structuration ?

2) Il est clair que les portails de recherche "Grand Public" ne peuvent plus avoir un modèle de revenu uniquement basé sur la publicité. Cependant l'offre faite aux annonceurs doit à mon sens s'étoffer et ne plus mettre en avant uniquement le volume de visiteurs mais également le ciblage à travers une exploitation dynamique des mots clés, développer des offres de profiling, etc.
Côté utilisateurs, le développement de services payants est en cours de par l'apparition des offres payantes de référencement dans les annuaires, je pense que là il y a encore beaucoup de possibilités : le référencement actif par exemple (consistant à indexer à une fréquence donnée les sites dit dynamiques permettant d'avoir une actualité beaucoup plus fraîche). Développer des "moteurs thématiques", les thématiques étant définie par la catégorisation des mots clés, et ensuite les sites se définissant comme pouvant faire partie de la catégorie soient également indexés plus fréquemment etc. Ce type d'offre pourrait être payante tant du côté utilisateur que du côté site souhaitant être référencé.
Mais notre seule problématique est d'éviter pour nos clients le service de recherche qui remplira tant les attentes de leurs internautes du point de vue de la qualité des réponses, que celles de l'opérateur du point de vue de sa capacité à être rentable.

3) Les intervenants proposant des outils de recherche dans les 2 ou 3 années à venir seront caractérisés par leur grande flexibilité, la capacité de leur solution à répondre à la spécificité du client (la pertinence et la rapidité du temps de réponse étant les 2 critères incontournables et non différenciant !).

Sur le plan technologique, on pourrait espérer une normalisation de l'accès aux informations à crawler. Le crawling "http" montre déjà ses limitations et une des évolutions majeures et nécessaires est la standardisation de moyens d'indexation dédiés. Peut-être les avancées actuelles sur XML/SOAP et WSDL/UDDI permettront d'avancer dans cette direction. Les plus difficiles à convaincre ne seront pas les opérateurs de services de recherche, mais bien les sites de contenu.

Greg R. Notess, responsable du site Search Engine Showdown (<http://searchengineshowdown.com>)

~~~~~

1) 2001 a été une année de test, pendant laquelle les moteurs de recherche, ou les compagnies qui les ont mis en place, ont déterminé si les dépenses engagées pour les maintenir étaient rentables ou non. Malheureusement, certains d'entre eux ont décidé que le jeu n'en valait pas la chandelle et ont arrêté leur offre de recherche d'information. Heureusement, d'autres ont fait ce qu'il fallait pour survivre à cette année difficile et, particulièrement pour Google et Fast, se sont même permis d'introduire de nouvelles fonctionnalités et possibilités de recherche. Encore plus surprenant, deux nouveaux outils de recherche (Teoma et Wisenut) ont même fait leur apparition malgré le contexte économique difficile. Alors, c'était peut-être une année de transition, mais il est important de noter que la recherche d'informations sur le Web s'accélère, est disponible sur de nombreux sites, qu'elle est toujours aussi populaire et qu'elle se développe de jour en jour.

2) La publicité en ligne continue de jouer un rôle dans les revenus des outils de recherche. Mais les ventes de technologie pour la recherche d'information sur les sites web ou les intranets ont un réel potentiel pour devenir une source plus importante de revenus. Les offres de référencement payant peuvent aider également, mais je n'ai aucune idée de leur potentiel économique futur. Je pense que de nouvelles idées vont également émerger d'ici peu.

3) Les moteurs de recherche sur le Web vont continuer à répondre à la demande, mais je ne suis pas sûr que quiconque soit prêt à payer pour obtenir des résultats de recherche. Dans le domaine de l'évolution des technologies de recherche, il est très complexe de prédire le futur. Cela dépend de beaucoup de facteurs. Personnellement, je voudrais voir des fonctionnalités de recherche avancée de haute tenue, une plus grande couverture du Web ainsi que de nombreuses et efficaces possibilités de recherches booléennes.

Leslie Ray, VP, International Markets, NorthernLight (<http://www.northernlight.com/>)

~~~~~

1) Comme toutes les années précédentes, en 2001, Northern Light s'est continuellement focalisé sur l'évolution des besoins de nos clients qui incluent maintenant des petites et grandes entreprises comme Cisco et Mastercard International entre autres. 2001 a été une année dure pour les outils de recherche, beaucoup d'entre eux disparaissant du Web. La situation de Northern Light est un peu différente dans le sens où nous sommes là depuis six ans maintenant. Un dinosaure comparé à certains, et nous en sommes fiers. Nous sommes encore forts et nous venons d'ailleurs de boucler un nouveau tour de table de 20 millions de \$ US. A contrario d'autres outils de recherche, Northern Light est plutôt spécialisé dans l'intégration de solutions de recherches et de contenus et tient à rester dans cette voie. Nous qualifierions 2001 d'année de transition. Ce ne fut pas une révolution car nos clients n'ont pas soudainement réalisé qu'ils avaient un réel besoin d'accès rapide, fiable et pertinent à l'information. Cela fait longtemps qu'ils ont émis ce souhait. La transition vient du fait que nos clients ont maintenant des besoins précis pour maintenir, mettre à jour, et exploiter les innombrables systèmes (intranets, extranets, portails externes) déjà installés, en s'aidant d'un environnement homogène. En 2001, Northern Light a vu ses clients opérer cet état de transition et a aidés certains d'entre eux à migrer leurs multiples sources d'informations en un système unique et plus simple.

2) Oui, l'environnement économique est complexe. Cela a toujours été la raison pour laquelle Northern Light ne s'est jamais uniquement focalisé sur la publicité en ligne comme modèle économique. En fait, 85% de notre "business" vient de la vente de notre technologie à des sociétés et seulement 15% de notre chiffre d'affaires provient de la publicité. Nous continuerons en 2002 sur cette voie en nous concentrant sur le "Global 2000" (les 2000 sociétés les plus importantes dans le monde) comme clientèle principale pour fournir des solutions d'intégration d'outils de recherche et de contenus. Nous continuerons, bien sûr, à vendre de la publicité, mais ce n'est pas là notre objectif majeur.

3) Il est très difficile de prédire le futur actuellement. Cependant, dans deux ou trois ans, nous pensons qu'il y aura un besoin encore plus grand pour des solutions de recherche d'information et d'intégration de contenu. Imaginez la masse de contenus créés chaque jour par les utilisateurs, les entreprises et les éditeurs. Dans deux à trois ans, nous aurons besoin de technologies encore plus efficaces et fiables et d'outils encore plus précis pour résoudre le problème de la croissance de l'information disponible.

Thomas Bertrand, passionné d'outils de recherche et ancien co-webmaster du site "Chasseurs de moteurs" (<http://www.chasseurs-de-moteurs.net/>)

~~~~~

1) L'économie de marché est sans pitié avec les mauvais élèves. La sanction est donc à la mesure des mauvaises stratégies choisies ainsi que de l'état d'esprit déplorable qui transpirait de la "home page" de la plupart des outils de recherche. En misant tout sur l'orientation "portail" et non sur la satisfaction de l'internaute, ils sont arrivés à faire fuir leurs visiteurs en les martyrisant de "pop up" et autres "skyscraper" en oubliant que la qualité des algorithmes robots (ou la vitesse et la qualité de traitement des netsurfers pour les annuaires), la pertinence à l'interrogation et le rafraîchissement d'une base de données ainsi que sa taille, restent les fondations sans lesquelles un outil de recherche n'est rien.

Cette année marque la fin d'un règne. Celle des dinosaures du Web.

Après un bon festin, il s'agit maintenant de payer l'addition. Et ce n'est pas fini. En lisant entre les lignes, on peut avancer que l'outil phare du moment, Google, cherche activement de nouveaux modèles financiers qui lui permettraient de continuer à assumer son service. Ce qui pourrait signifier que Google suit le même chemin que les Excite, AltaVista et consorts.

2) Le modèle financier basé sur la publicité ne sera pas viable tant que le commerce électronique ne décollera pas. Même en récupérant de nombreux visiteurs par l'achat publicitaire (bandeaux, liens sponsorisés et autres), les internautes n'achètent pas ou peu en ligne. Quel est l'intérêt de se retrouver avec des personnes qui passent devant votre rayon et qui ne remplissent pas leurs caddies ?

Pour gagner de l'argent, je pense qu'il y a certaines choses à tenter :

- Louer la base de données des mots clés tapés par les internautes.

- Louer la technologie du moteur (voir Fast, Inktomi, Exalead, Kartoo...).

- Développer un service sérieux destiné aux adultes (X) car avec ce thème, les internautes ne regardent pas à utiliser leurs cartes bancaires (Lycos est sur le plan d'ailleurs).

- Mettre en place des services qui permettront l'identification (profiling) puis la personnalisation de l'outil afin de mieux servir les visiteurs et de

mieux les démarcher.

3) Afin que le web fonctionne parfaitement et partout, l'éclatement de ces systèmes de recherche est nécessaire à son développement. J'imagine des milliers ou millions d'annuaires spécialisés, régionaux et autres interrogés par des interfaces "métas". Maintenant se trouve devant nous une difficulté majeure. La toile arrivera à son épanouissement le jour où sur la terre les richesses seront partagées par l'ensemble et non aux mains d'une minorité...

Brent Winters, FirstPlace Software (<http://www.webposition.com>)  
~~~~~

1) Cette année a connu un bouleversement pendant lequel les outils de recherche, comme beaucoup d'autres business en ligne, ont dû se focaliser sur le chiffre d'affaires et la rentabilité pendant que leurs cours en bourse chutaient. Il y a de fortes chances qu'Excite disparaisse cette année ou devienne un nouvel outil de recherche proposant les résultats de GoTo (Overture), de la même façon que Disney a stoppé Go.com plus tôt cette année. Les autres outils de recherche, comme Google, devraient en profiter pour affermir leur suprématie.

2) Les outils de recherche continueront d'envisager d'autres voies pour gagner de l'argent. Je les vois bien proposer plus d'options payantes aux webmasters et aux marketeurs pour répondre à leurs besoins. Les personnes choisissant le référencement payant et obtenant une indexation plus rapide de leurs pages dans les index voudront s'assurer d'obtenir un bon retour sur investissement. C'est pourquoi des logiciels comme WebPosition Gold continueront d'être indispensables dans les mois qui viennent. Pourquoi payer des centaines de dollars dans la mise en place de campagnes marketing auprès des outils de recherche pour être uniquement présent dans l'index, sans apparaître en bonne position ? Le fait de suivre et de rester en veille sur les derniers outils et technologies disponibles restera critique pour les webmasters. Ces outils leur permettront de rester compétitifs sur les outils de recherche.

3) Je ne détiens pas de boule de cristal, malheureusement. Cependant, le paysage des outils de recherche s'est modifié de façon significative depuis 1994 et continuera de se modifier. Il y aura toujours beaucoup d'outils de recherche sur le Web, mais on peut penser que les parts de marché vont se consolider pour les outils majeurs dans les prochaines années. En même temps, des millions de nouveaux venus vont arriver sur l'Internet et représenteront un nouveau marché potentiel encore plus intéressant pour les entrepreneurs en ligne.

Olivier Nérot, fondateur du site Human Links (<http://www.human-links.com/>)
~~~~~

1) Evolution certainement. Je pense que l'on assiste à la fin des moteurs de recherche de type "Las Vegas", clignotant de tous cotés et offrant des pubs, l'horoscope et la météo. L'essor de Google est représentatif de cette évolution : interface sobre, requêtes simples, mises à jour rapides et pertinence...

2) Si un moteur de recherche est très pertinent, il peut proposer des sites commerciaux dont il aura fait payer le référencement. Mais il faut absolument que ces sites correspondent à la requête de l'utilisateur, sinon c'est autant de bruit qui rebute. Je pense qu'il est néanmoins possible de faire payer l'indexation de sites qui souhaitent augmenter leur visibilité, à partir du moment où il y a un travail à réaliser, un service rendu (indexation d'un domaine entier, de bases de données, ou indexation journalière).

3) A mon sens la première évolution portera sur le contenu indexé. Actuellement, seule une page web complète (et sans trop de frames...) peut être indexée. Avec de nouveaux supports à l'information, l'information indexée sera de plus en plus "minimale" et multimodale : le paragraphe d'une revue, la photo d'un article, les paroles d'une chanson, les références d'un article...

Une telle précision devra être accompagnée d'une évolution des requêtes, permettant de faire des recherches précises, sans alourdir leur syntaxe. Cela ne pourra passer que par une analyse approfondie de l'attente de l'utilisateur, et/ou des requêtes par affinage. Je suis donc persuadé que les méthodes étudiées aujourd'hui pour affiner des requêtes de façon statistique (Exalead ou Vivissimo), et des méthodes de profilage "non intrusif" (celles utilisées dans Human-Links), seront nécessaires.

Il y a beaucoup à attendre des méthodes neuronales ou génétiques pour les algorithmes de recherche d'information dans les années à venir. Par contre, je ne crois pas qu'il y ait de méthodes en langage naturel suffisamment avancées pour aller au delà d'une simple démo...

Dernier point, je continue à croire que la tâche d'indexation est impossible par une seule structure, et qu'il est nécessaire que chaque fournisseur de contenu participe à ce travail. L'approche P2P fait partie des solutions intéressantes, pour garantir une indexation large et réactualisée du web. Cela offre plusieurs avantages pour le site : contrôler son indexation, analyser les requêtes, répondre de façon personnalisée à chaque demande, et adapter son contenu dans une relation client privilégiée.

Je crois donc que les moteurs de recherche évolueront vers des solutions de type "architecture distribuées" permettant à chaque grand site (hébergeur, journaux, grandes sociétés...) d'indexer eux mêmes et chez eux, l'information contenue dans le moteur. Cela peut d'autant plus faire un bon modèle économique, puisque le moteur de recherche vend sa technologies aux sites.

Philippe Crancee - P.D.G du site Indexa (<http://www.indexa.fr/>)  
~~~~~

1) Je qualifierai l'année 2001 de repositionnement économique. Devant la crise qui touche Internet, la plupart des annuaires et moteurs ont vu une baisse des revenus publicitaires, liée aux investissements bloqués. Les "start-up" ne peuvent plus se lancer dans des campagnes publicitaires fort coûteuses car la notoriété ne suffit plus à valoriser un site. On revient dans une économie plus classique ou la valorisation est déterminée par la rentabilité, critère désormais indispensable pour convaincre un capital-risqueur.

Au niveau d'Indexa, notre audience continue d'augmenter et nous avons dépassé le million de PAP en Octobre pour 137 000 visites (92500 visiteurs uniques). Cette augmentation nous a permis de conserver notre CA publicitaire alors qu'il aurait dû augmenter. Notre positionnement B-to-B nous a permis de maintenir un bon niveau de prix alors qu'il chute dans le domaine B-to-C.

2) Les revenus publicitaires sont trop aléatoires, ce qui gêne dans la gestion de l'entreprise. L'idéal est de bâtir un modèle économique basé sur des produits ou services payants, la part publicitaire représentant alors la part de bénéfice.

Nous avons déjà en mis en place d'autres sources de revenus sur Indexa et nous serons sans doute copiés d'ici quelques mois. Nous avons lancé l'inscription sous 48h dans notre annuaire en Mai 2000, maintenant tous les annuaires et moteurs proposent un service équivalent. Nous offrons déjà un service de positionnement du site par l'intermédiaire de la Fiche Flash (229 euros HT/an). Sur 10 mots clés, le site apparaît en première page des résultats de recherche. Pour garantir un bon service il faut prendre quelques précautions. Par exemple, il faut vérifier que les mots clés choisis correspondent vraiment au site (sinon tout le monde voudra acheter le mot clé "sexe"). Mais ces services ne représentent pour l'instant que 15% de nos revenus, le marché n'est donc pas encore mûr.

3) Le marché publicitaire est capricieux mais il ne s'arrêtera pas pour autant. Mais pour convaincre, il faudra prouver le retour sur investissement car c'est un critère que l'on peut mesurer sur Internet (dommage), alors que personne ne peut estimer le retour sur d'autres supports. Le taux de clics est-il meilleur sur Internet que dans une revue ? Il faudra que la publicité affichée soit ciblée, et pas seulement sur les mots clés tapés, mais aussi sur le profil de l'utilisateur. On pourrait même afficher les résultats de recherche en fonction de ce profil. Bien sûr, la difficulté est d'obtenir le profil d'un utilisateur et de respecter les conventions de la CNIL...

Olivier Roumieux, revue Archimag (<http://www.archimag.com/>)
~~~~~

1) Je qualifierai l'année 2001 d'année de maturité pour le marché des outils de recherche. Peu de nouveautés bouleversantes (à l'exception de Kartoo ou de Human Links). Suprématie de Google (les nouvelles fonctionnalités arrivent : plus de types de docs indexés, les news...). Stabilisation d'Altavista. En espérant que cette accalmie permette aux usages de se développer plus sereinement.

2) Comme vous le notez toutes les semaines dans votre lettre, l'année 2002 sera l'année de la diversification des sources de revenus pour les out...

de recherche, et donc notamment du référencement payant. On peut imaginer des prestations du type de celle de Nomade.fr, avec une réelle "écoute du client".

La diffusion en intranet semble être un incontournable, peu de sociétés semblent avoir encore les reins solides pour le prouver au travers de moteurs ouverts au public, comme ce fut le cas par le passé (Digital et Dell). De nouvelles solutions, comme Influx, se développent donc à l'usage exclusif des entreprises qui peuvent éventuellement les utiliser sur leur site (ex : 01). Dans cette gigantesque époque de "normalisation", il semble que des outils payants à forte valeur ajoutée devraient logiquement apparaître, et les acteurs déjà bien établis continuer à se développer à l'intention du grand public (ou se segmenter).

3) Je l'ai déjà évoqué dans la précédente réponse : à mon sens, le paysage se normalise considérablement. Quelques outils vont demeurer gratuits à l'intention du grand public et des services à valeur ajoutée cibleront des communautés professionnelles. Peut-être cela permettra à la profession de documentaliste de retrouver sa place d'intermédiaire dans la chaîne de l'information. Une chose me semble sûre : on ne reviendra pas complètement à la situation antérieure : même si beaucoup de sources vont devenir payantes, leur gratuité temporaire aura permis d'aiguiser le sens de la recherche chez de nombreux internautes. Revers de la médaille : ce qui m'effraie un peu, c'est le syndrome du Push et la volonté des grands éditeurs de pousser de l'information vers les utilisateurs, quitte à ce que ces derniers se cantonnent à une dizaine de canaux (je crois que le dernier produit dans le genre s'appelle le Telezap).

#### ■ 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 ;-)

-> L'Annuaire Blanc, réalisé par la société "L'île des médias" et présent sur Lokace (il était également, auparavant, sur Altavista France et Excite France entre autres), s'arrêtera fin décembre 2001. Les évolutions du marché ne permettent pas à ses concepteurs de continuer à le maintenir. Un nouvel outil de recherche qui s'arrête... Pour l'instant, aucune décision n'est prise quant à l'outil qui remplacera l'Annuaire Blanc sur Lokace.

-> La balise "Doctype" (exemple sur la page d'accueil d'Altavista France : <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">), qui est en première ligne d'un code HTML normalement, a une certaine importance sur Altavista. Ne l'oubliez pas dans vos documents HTML.

-> Une refonte de la page d'accueil de Voila serait en cours de discussion.

-> La possibilité de soumission payante sur l'annuaire de Lycos France devrait être disponible d'ici à la fin du mois de novembre 2001.

-> Altavista aurait enfin "reconstruit" ses index en un seul fichier (la nouvelle était attendue depuis de nombreux mois), mais un système de "compartiments" serait quand même utilisé, par pays, ou par version d'Altavista (altavista.fr par exemple). Difficile d'en savoir plus pour l'instant, mais je cherche des infos sur ce point...

#### ■ En Bref

-> L'Open Directory (<http://www.dmoz.org/>) a passé le cap des 3 millions de sites indexés.

-> Voici quelques statistiques sur le nombre de mots clés saisis sur le métamoteur Kartoo (<http://www.kartoo.com/>) :

29% des requêtes sur KartOO contiennent 1 seul mot clé.

71% des requêtes contiennent 2 mots ou plus.

45% des requêtes contiennent 3 mots ou plus.

26% des requêtes contiennent 4 mots ou plus.

25% des requêtes contiennent 5 mots ou plus.

8% des requêtes contiennent 6 mots ou plus.

5% des requêtes contiennent 7 mots ou plus.

3% des requêtes contiennent 8 mots ou plus.

2% des requêtes contiennent 9 mots ou plus.

On trouve encore une trentaine de requêtes par jour avec plus de 15 mots !

Les gens semblent adorer le fait de jouer avec les boutons "plus" et "moins" de l'interface ;-)

Ces statistiques sont plutôt plus importantes que sur les outils de recherche "traditionnels", puisque sur ceux-ci, on trouve à 95% un ou deux mots clés dans les requêtes.

-> Lycos US propose ses "Lycos Elite" (<http://50.lycos.com/elite.asp>), qui rassemblent les mots clés phares et les plus demandés dans l'histoire du moteur de recherche. Amusant...

-> Northern Light et Profusion proposent des "alertes" par e-mail lorsqu'un document nouveau a été indexé sur les mots clés que vous avez indiqués.

Pour Northern Light :

[http://www.northernlight.com/cgi-bin/cl\\_cliplist.pl](http://www.northernlight.com/cgi-bin/cl_cliplist.pl)

Pour Profusion :

<http://www.profusion.com/alerts.asp>

