

■ Looksmart France : naissance d'un nouvel annuaire francophone

Présentation



Depuis quelques jours, l'annuaire Looksmart (ou plutôt BTLooksmart) France, dernier né des annuaires francophones est officiellement en ligne. BTLooksmart est une joint-venture entre Looksmart et BT (British Telecom), créée en 1999 (<http://www.blooksmart.com/>) et couvrant le marché européen et asiatique.

Pour information, Looksmart a été fondé en 1995 en Australie par Evan Thornley (ancien consultant chez McKinsey) et sa femme Tracey Ellery. Le site américain a ensuite été créé en 1997. Looksmart revendique un taux de pénétration de 84% aux Etats-Unis et de 65% en Grande-Bretagne. Son taux actuel en France serait de 35% (source Mediametrix) avec pour objectif à court terme un chiffre similaire à celui observé chez nos voisins britanniques. Looksmart est aujourd'hui présent dans 26 pays avec 31 annuaires disponibles en 13 langues différentes.

BTLooksmart France compte 30 personnes, dont 22 travaillent pour le pôle éditorial : un manager, deux "team-leaders", deux correcteurs, un responsable de l'arborescence (taxinomiste) et 16 éditeurs-netsurfeurs qui sont chacun spécialisés dans un domaine particulier et sont soumis à des objectifs en termes de nombre de sites intégrés dans la base chaque jour. Pour la petite histoire, les netsurfeurs travaillent sur place (pas de télétravail) et sont soumis aux 35 heures. Y a pas de raison... Looksmart France propose une base de données de 125 000 sites web francophones, classés dans plus de 10 000 catégories, et est déjà présent depuis quelques semaines sur les sites suivants :

- Altavista France (<http://www.altavista.fr/>)
- Club Internet (<http://www.club-internet.fr/selection/>)
- Worldonline (<http://www.worldonline.fr>)
- Freesurf (<http://www.freesurf.fr/>)

Il propose maintenant une présentation de son annuaire sur le site : <http://www.looksmart.fr/>

Sans que ce site ait pour objectif de servir d'outil de recherche particulier, mais plutôt de "vitrine" pour l'entreprise. En effet, l'approche de Looksmart est basée sur une offre B2B et non grand public. Ses clients sont les FAI, portails et sites média. Ils intègrent leur solution de recherche dans les couleurs et charte graphique du site partenaire en partageant par la suite les revenus publicitaires garantissant ainsi une nouvelle source de revenu à son client.

Les serveurs de l'annuaire francophone sont basés chez BT à Londres. Les pages proposant les données émanant de Looksmart France sur Club-Internet, Freesurf et Worldonline viennent en fait en direct des serveurs de BT. Seul Altavista a intégré l'outil (il reçoit en fait un fichier de mise à jour toutes les semaines de la part de Looksmart) sur ses serveurs (également présents à Londres d'ailleurs), comme Excite France le fait avec l'annuaire de l'Île des Médias (voir lettre du mois dernier).

Si le nombre de grandes rubriques est le même d'un portail à l'autre (13), Altavista France n'en propose que 4 sur sa page d'accueil. En fait, on a bien accès sur cet outil de recherche à l'annuaire Looksmart au grand complet, mais seulement à partir de cette adresse : <http://fr.altavista.com/cgi-bin/query?pg=dir&cn=fr&cl=fr>

Cette situation n'est que temporaire, Altavista France s'orientant certainement vers une interface proche du nouveau look du site de la maison-mère (<http://www.altavista.com/>), faisant la part belle à l'annuaire Looksmart.



Description de l'annuaire

Looksmart France propose 13 grandes rubriques :

- Achats et services
- Actualité et médias
- Arts et culture
- Economie et commerce
- Education et emploi
- Famille et maison
- Informatique et Web
- Loisirs et Discussion
- Santé et mise en forme
- Sciences et références
- Société et politique
- Sports et plein air
- Tourisme et voyage

Une autre arborescence, sur 10 grandes rubriques, sera également disponible dans peu de temps, Looksmart France y travaille. Cette nouvelle segmentation, qui sera mise en place en complément de l'actuelle, proposera exactement le même nombre de sites, mais répertoriés de façon différente.

Le "déroulé" des rubriques et sous-rubriques est assez classique. La version d'Altavista France propose cependant, là encore, une différence par rapport aux autres clients de Looksmart : elle propose une liste de sites intermédiaires, dans des rubriques proposant des liens vers des sous-rubriques, alors que, sur les autres moutures, seule une liste finale de sites est proposée, sans aucune indication à l'intérieur des catégories intermédiaires.

Côté Club-Internet et Worldonline (25 résultats affichés par page), à l'intérieur d'une catégorie finale, les sites sont classés de façon alphabétique. Cependant, les netsurfeurs de Looksmart France ont la possibilité de placer en tête de liste les sites qui leur semblent les plus pertinents, sans tenir compte de cet ordre alphabétique. Ceci correspond aux "chiens d'or" de Lycos, aux lunettes de Yahoo! ou aux palmes de Nomade.fr. Une légère confusion vient du fait que ces sites mis en exergue ne sont identifiés par aucun pictogramme. Cependant, il se pourrait que cela change prochainement.

Les sites qui sont ainsi "remontés" manuellement sont soit les sites proposant un contenu très pertinent, soit les sites reconnus comme étant "de référence", soit les sites officiels (par exemple pour les clubs de football).

Looksmart permet également à ses clients portails ou FAI un système de type "coup de coeur" qui permet à ce dernier de faire remonter dans les listes, à l'intérieur d'une catégorie, un ou plusieurs sites, alors qu'il serait logiquement classé par ordre alphabétique à l'intérieur de la liste "classique". Il peut alors s'agir d'un partenaire, d'un annonceur, etc.

Altavista France (10 résultats affichés par page), pour sa part, prend en compte les résultats proposés par Looksmart mais les reclasse selon des critères qui lui sont propres. Il est très possible qu'un système de type indice de clic ou de popularité soit pris en compte par le moteur, en

complémentarité des critères de Looksmart.

Freesurf, Altavista, Club Internet et Worldonline proposent également un module de saisie de mots clés sur l'annuaire (Altavista propose, en fait, les résultats de l'annuaire avant ceux de son moteur de recherche, ce qui est une petite révolution chez cet outil qui avait, jusqu'à maintenant, toujours privilégié les résultats de type "moteur" sur ceux des annuaires).

Chez Freesurf, Club-Internet et Worldonline, les classements sont identiques d'un outil à l'autre. Altavista France propose un autre classement des sites de l'annuaire suite à la saisie d'un mot clé. Sont présentés, de façon classique : le site, un résumé (les résumés des sites présents dans Looksmart font un maximum de 145 caractères) ainsi que la catégorie à laquelle appartient la source d'information. Pour des mots clés génériques, les catégories les plus pertinentes sont également affichées. Bon, il reste encore quelques petites imperfections à régler (la mention "Next 10" sur un outil de recherche francophone... ;-)) mais l'annuaire semble aujourd'hui bien au point.

Recherche sur Looksmart France

L'outil de recherche ne propose pas d'opérateur booléen (ET, OU, AND, OR, +, -, guillemets, etc.) pour effectuer les recherches. Même si la situation est moins "grave" sur un annuaire que sur un moteur (une recherche vierge d'opérateur sur 125 000 sites génère moins de "bruit" que sur 20 millions de pages), on peut cependant regretter de ne pouvoir mieux maîtriser ses recherches. Dommage également que le nombre de résultats trouvés dans l'annuaire ne soit pas indiqué. Peut-être dans une future version de l'outil ? Les fonctionnalités de la recherche sont les suivantes :

- Casse des lettres : non prise en compte : une recherche sur "IBM" est identique à une requête sur "ibm".
- Accentuation : l'outil a semblé avoir pas mal de problèmes, dans un premier temps, avec les lettres accentuées : une recherche sur les mots "électricité" et "bébé" ne donnait aucun résultat, alors que "bebe" et "electricite" fournissaient des réponses. Vestiges d'un système d'origine américain ne gérant pas les accents ? Le bug est cependant quasiment entièrement corrigé à l'heure actuelle et ne devrait rapidement plus être qu'un mauvais souvenir.
- Ordre des mots : non pris en compte. Une recherche sur "paris dakar" est identique à une autre sur "dakar paris".
- Opérateur par défaut : ET : "paris france" correspond à "paris ET france". Il ne semble pas possible d'indiquer un "OU" entre deux mots. :-(

Looksmart France ne propose pas non plus de zone de recherche avancée. En fait, au niveau de la syntaxe de saisie des requêtes, la version francophone de l'outil se calque exactement sur les (faibles) possibilités du site de sa maison-mère (voir https://www.abondance.com/outils/l_s_syntaxe.html). Les responsables de la filiale francophone n'ont certainement, d'ailleurs, aucune possibilité n'y changer quoi que ce soit...

Soumission sur Looksmart France

Une soumission de site est possible sur Altavista France, le site de Looksmart France et bientôt sur celui de Club Internet.

Pour 159 euros (soit 1043 FHT, et donc 1247,43 FTTC), Looksmart France propose l'offre "Soumission Express", assez similaire aux offres américaines et anglaises, déjà disponibles. Voici quelques informations complémentaires sur cette offre :

- Il n'existait pas, dans un premier temps, d'offre gratuite, seule l'offre payante a été tout d'abord disponible. Une offre gratuite a ensuite vu le jour (http://submit.looksmart.co.uk/submit_non.asp?isp=ut) pour les associations et organismes à but non lucratif. Lorsqu'un site associatif est soumis (et qu'il répond, bien sûr, à la charte de Looksmart), les délais ne sont pas garantis pour la prise en compte. Un e-mail est cependant envoyé, que le site soit accepté ou refusé, notifiant la décision de Looksmart.
- Contre cette somme, il est offert une garantie que le site sera évalué par les netsurfeurs sous les 5 jours ouvrés suivants.
- S'il est accepté, le site sera classé dans les catégories les plus appropriées de l'annuaire.
- Il n'est possible d'effectuer une soumission que dans une seule catégorie. Mais les netsurfeurs peuvent décider, par la suite, d'inscrire le site dans plusieurs d'entre elles, éventuellement.
- Les sites acceptés le sont pour une durée illimitée et sont visibles dans le répertoire deux semaines après leur indexation dans la base de données. Que le site soit accepté ou refusé, un e-mail est envoyé à la personne ayant soumis le site.
- Le fait d'être inscrit au travers de l'offre proposée sur Altavista France inscrit le site sur l'annuaire fourni à tous les autres clients de Looksmart France (Club-Internet, Worldonline, etc.).
- Il ne s'agit en rien d'une garantie d'inscription, mais bien d'une prime à la rapidité de prise en compte du site de la part de l'outil de recherche.
- De même, il ne s'agit pas d'une proposition de positionnement du site dans les résultats d'une recherche.
- La somme versée n'est pas remboursée si le site est refusé. Cependant, dans un délai de 30 jours, il est possible d'effectuer une nouvelle soumission par mail à l'adresse correction-fr@looksmart.net, après avoir corrigé certains "points noirs" du site ayant entraîné le refus (liens cassés ou zones en cours de travaux par exemple).

Les critères d'acceptation (http://submit.looksmart.co.uk/submit_guide.asp?isp=ut mais également développés à l'adresse http://submit.looksmart.co.uk/submit_fa_q.asp?isp=ut) sont très classiques pour un annuaire. Tous les sites doivent, notamment, être "family friendly", donc pas de site à caractère pornographique, violent ou illégal). Il est également indiqué que Looksmart France fera régulièrement des "ménages" sur son outil de recherche et qu'un site ne répondant pas à la charte de qualité de l'annuaire peut être enlevé de sa base de données.

L'outil utilisé par les netsurfers a été créé spécifiquement en 1995 par Looksmart et est directement disponible sur le web.

L'objectif de la personne chargée de la taxinomie de l'outil est de créer une nouvelle catégorie dès qu'une somme de 25 à 30 sites sont déjà présents dans une rubrique unique, afin d'éviter des listes interminables de sites qui ne riment à rien. Bien entendu, chaque netsurfeur dispose d'un "guideline" qui indique la structure de l'arborescence de l'outil ainsi que le style éditorial à adopter pour rédiger les résumés.

Les netsurfeurs de Looksmart ont la possibilité, comme chez Nomade, de rajouter des mots clés "invisibles" à la fois sur une fiche descriptive de site et sur un intitulé de rubrique. Ces mots clés ne s'affichent pas dans les résultats mais sont pris en compte comme critères de pertinence lors de la recherche.

Lorsqu'un site est inscrit, un pictogramme "nouveau" est affiché devant son nom pendant 7 jours, mais le site reste "à sa place" dans la liste proposée à l'intérieur d'une catégorie. Il n'est pas mis en tête de liste pendant cette période, comme sur Yahoo!. Enfin, l'un des objectifs de Looksmart France est d'avoir un taux de liens morts inférieur à 3%. Un "ménage" régulier des adresses proposées est donc planifié afin d'arriver à cet objectif.

Pour info, la FAQ sur l'offre "Soumission Express" de Looksmart France : http://submit.looksmart.co.uk/submit_fa_q.asp?isp=ut Et le sommaire de la lettre "Recherche et Référencement" spéciale "Référencement payant" : <http://actu.abondance.com/lettres/sommaire-RR-122000.html>

L'offre de référencement payant est une solution qui a été mise en place pour gérer les priorités et éviter de recevoir, comme Yahoo! France, plus de

1 000 soumissions par jour, ce qui aurait été ingérable selon les responsables de l'annuaire. Pourtant, certains sites, très intéressants, mais réalisés par des "amateurs", n'ont pas la possibilité de payer pour être pris en compte par Looksmart. Et, quel que soit le degré de spécialisation des neturfers de l'annuaire (et l'on peut penser qu'il est excellent), il est impossible, au vu de la croissance du Web actuel, même francophone, d'être sûr de ne pas avoir oublié un site intéressant, voire même majeur. Ceci dit, rien n'est gravé dans le marbre et, selon la demande, la procédure de soumission payante unique pourrait être revue à l'avenir. Plusieurs solutions seraient éventuellement envisageables. Peut-être en saurez-vous plus d'ici quelques temps... ;-)

Pour information, les chiffres pour la Grande-Bretagne, qui a mis en place l'offre de soumission payante avant la filiale française, est d'environ une cinquantaine de demandes par semaine. L'offre a été lancée il y a 3 mois et les chiffres augmentent toutes les semaines.

Offre Looksmart France

Selon Emmanuelle Siary, Directrice du Développement de BTLooksmart, "après avoir implémenté notre solution de recherche, le trafic des sites partenaire augmente en général de 10% à 15% et parfois jusqu'à 40%". Le but est de proposer un annuaire de recherche en "semi marque blanche" (le nom de Looksmart apparaît mais l'annuaire est proposé dans la charte graphique du client). La régie publicitaire est également intégrée à l'offre.

Dans le monde, Looksmart représente aujourd'hui un CA de 110 millions de \$ pour 50 millions de visiteurs uniques par mois et 15 millions de pages vues par jour. Le réseau LookSmart couvre 41% des internautes dans le Monde.

Il est à noter que Looksmart France propose non seulement un annuaire, mais également une solution complète, gérée de façon globale, et comprenant l'offre éventuelle d'un moteur de recherche en complément : Inktomi, Altavista ou Fast. Club Internet et Freesurf utilisent Inktomi et Worldonline n'utilise aucun moteur, Altavista France utilise... Altavista ;-). La cible principale de l'outil est bien entendu les grands portails et les FAI. Le modèle économique est variable selon les clients et les prospects, sur la base d'un coût de mise en oeuvre initial (50 KF) couplé à un pourcentage sur la publicité générée. Sur ce dernier point, tout dépend, bien entendu, du potentiel publicitaire du site portail. Rappelons qu'aux Etats-Unis, Looksmart paye Altavista pour que celui-ci utilise son annuaire...

De la même façon, lorsqu'un site portail (Altavista, Club Internet) propose sur ses pages la procédure de soumission payante de Looksmart, le revenu généré peut être unilatéralement reversé à l'annuaire ou partagé entre les deux parties.

Conclusion

Looksmart France a frappé très fort dès son entrée sur le marché français et a justifié le temps passé à préparer sa base données, peaufinée depuis l'été dernier. Il y a fort à parier que de nouveaux accords avec d'autres FAI ou portails francophones importants soient annoncés dans les jours qui viennent (à mon avis, Excite France devrait être l'un de ceux-là, de par son accord aux Etats-Unis avec Looksmart).

Sa base de données de sites semble effectivement très bien fournie. Looksmart France se positionne d'entrée parmi les trois ou quatre annuaires majeurs du monde francophone. Il est indispensable, donc, d'y être présent si vous êtes gestionnaire d'un site. Deux regrets, cependant, concernant cet outil : un manque certain de possibilités au niveau des opérateurs de recherche disponibles et aucune solution pour soumettre un site de façon gratuite. Certains l'appelleront donc "un annuaire de riches", même si cette appellation est grandement erronée, l'outil proposant de très nombreuses sources d'information de type "personnel". Mais une possibilité de référencement gratuite, voire "contrôlée", aurait été de bon aloi, à mon avis. Si le principe de la soumission payante est très "tendance" et n'est plus remis en cause aujourd'hui, il aurait quand même été sympathique de donner leur chance à des sites qui ne peuvent pas payer plus de 1 000 F pour cette procédure, au risque de se voir refusé...

■ Open Directory Francophone : interview de Michelle Carrupt

Q : Michelle Carrupt, Pouvez-vous nous présenter votre cursus en quelques mots ?

R : Bonjour, je suis formatrice pour adultes, indépendante à Genève, spécialisée dans les formations informatiques pour les professionnels de la santé et les médecins. Depuis 1995, le web est devenu mon outil de travail quotidien.

Q : Pouvez-vous nous expliquer en quelques mots ce qu'est l'Open Directory Project ? (ODP)

R : C'est un projet gigantesque. Démarré sous le nom de NewHoo en juillet 1998 par deux programmeurs californiens, repris par Netscape à la fin de la même année, (Directory Mozilla, dmoz) il vise à développer l'annuaire le plus complet sur Internet, en faisant appel à des éditeurs volontaires du monde entier. Netscape met à leur disposition l'infrastructure, les outils d'édition et est propriétaire des droits d'auteur. Avec plus de 2 370 000 sites, 34 260 éditeurs, 342 000 catégories et 62 catégories linguistiques, l'Open Directory fait partie des "annuaires qui comptent" :

<http://searchenginewatch.com/awards/index.html#hall>

Un site référencé dans dmoz sera présent dans plus de 220 annuaires, dans plusieurs langues. La version francophone est disponible à l'adresse <http://www.dmoz.ch/>

Q : Certains disent que le nombre de 34 260 éditeurs inclut ceux qui ont travaillé un temps à l'ODP et sont partis depuis. Est-ce vrai ?

R : Ce sont les chiffres officiels qui figurent sur la page d'accueil <http://dmoz.org> . Ils représentent le nombre total des éditeurs ayant contribué à la création de l'annuaire depuis 1998. Le nombre d'éditeurs actifs n'est pas publié, mais de source bien informée il serait proche de 5 500 actuellement.

Q : Quel est votre rôle au sein de l'ODP?

R : J'ai fait mes premiers pas dans l'ODP en octobre 1999. Je suis une éditrice parmi les 5 500 éditeurs actifs, actuellement responsable de la catégorie World/Français, privilège que je partage avec 9 autres personnes. J'accorde une attention toute particulière à la catégorie World/Français/Santé, j'y ai inscrit la plupart des sites.

Q : Qu'est-ce qui vous a poussé à devenir éditeur de l'ODP?

R : Le hasard d'une navigation nocturne! Sur le site suisse www.search.ch j'ai cliqué sur un lien "become an editor". Je pensais tenir là enfin un outil simple pour le classement de mes signets, utilisables en formation. Ce fut le début de l'engrenage et ma première motivation. Mais j'ai aussi découvert le monde du référencement et ses méthodes, ce qui constitue un excellent apprentissage!

Q : Quels sont les liens entre www.dmoz.org et <http://www.dmoz.ch> ? Pourquoi n'existe-t-il pas de [dmoz.fr](http://www.dmoz.fr)?

R : La base dmoz peut être téléchargée librement et installée sur n'importe quel site web. <http://dmoz.org/license.html> Un éditeur suisse passionné l'a fait, après avoir acquis les noms de domaine dmoz.ch, dmoz.li, dmoz.de, dmoz.at. Son site donne directement accès aux catégories en français, allemand, italien et romanche. A ma connaissance, le domaine dmoz.fr est libre, aucun éditeur français ne l'ayant acquis.

Q : Vous partagez la responsabilité de la partie francophone de l'ODP avec 9 autres personnes. Comment se répartit le travail entre



Michelle Carrupt, Open Directory

vous?

R : Il n'y a pas de règles définies. Simplement, nous essayons de ne pas intervenir dans les catégories de prédilection des co-éditeurs. Les 9 personnes de cette branche peuvent éditer dans toutes les catégories francophones. Nous agissons surtout en fonction du nombre de sites en attente. La plus grande partie de notre activité à ce niveau est la répartition des sites vers les sous-catégories, et les conseils aux nouveaux éditeurs sur le forum. Une bonne entente règne parmi les "seniors editors" du World/Français!

Q : Comment devient-on éditeur sur l'ODP?

R : C'est très facile de devenir éditeur: un clic sur le lien "devenir éditeur" sur la page de la catégorie qui vous intéresse, un coup d'oeil attentif aux conseils donnés, et il ne vous reste plus qu'à trouver 3 sites non encore référencés. Leur donner un titre, les décrire brièvement en français, et respecter les règles de la langue française. (orthographe, accents, ponctuation). Votre demande sera évaluée, acceptée ou refusée par un éditeur appelé "meta editor". En principe, vous recevez une lettre de confirmation ou de refus en français. Vous avez ensuite 30 jours pour effectuer votre première édition d'un site. Les directives en français sont disponibles ici : <http://www.dmoz.org/special/francais/directives.html>

Q : Y a-t-il une sélection ou tout le monde peut-il vraiment le devenir?

R : La sélection est devenue beaucoup plus restrictive, spécialement pour les catégories commerciales. Les critères de sélection sont basés sur la pertinence par rapport à la catégorie, la qualité de la description, l'orthographe. Il faut savoir que chaque "meta editor" reçoit jusqu'à 200 demandes par jour, et qu'il est également bénévole!

Q : Quelle est la charge de travail quotidienne d'un éditeur de l'ODP?

R : La charge de travail est très variable: de 1 à 2 heures mensuelles à quelques heures par jour ou nuit pour l'édition dans la catégorie World/Français! Chacun donne le temps qu'il veut, il n'y a aucune obligation. Une catégorie avec un nombre élevé de sites sera plus exigeante.

Q : N'y a-t-il pas un risque qu'un éditeur prenne en compte une rubrique uniquement pour y refuser par la suite ses concurrents?

R : Bien entendu, et c'est peut-être ce qui motive l'adhésion de certains! Mais c'est compter sans la vigilance des éditeurs des catégories supérieures. Ce type de comportement est assez vite repéré, et la surveillance par les pairs assez efficace.

Q : Pouvez-vous nous décrire de la façon la plus précise possible en quoi consiste le travail d'un éditeur de l'ODP au quotidien?

R : Avant toute chose, l'éditeur doit prendre connaissance des directives d'édition, qu'il doit respecter. Celles-ci ont été traduites récemment en français, comme la plupart des interfaces, au grand soulagement des éditeurs francophones ! L'éditeur accède à sa page d'édition, protégée par mot de passe. Le nombre de sites en attente d'acceptation dans chacune de ses catégories y est indiqué. Pour chaque site soumis, il :

1. Vérifie si le site est déjà inscrit dans l'annuaire. Un double référencement n'est possible que selon des critères bien définis dans les directives, et nous disposons d'outils qui permettent de repérer les sites déjà inscrits.
2. Visite le site
3. Contrôle l'URL, le titre et la description, les corrige si nécessaire
4. Vérifie la pertinence de la catégorie
5. Accepte, supprime ou transfère le site dans une autre catégorie en commentant ses décisions.

Toutes les activités d'édition sont enregistrées dans un journal d'édition, consultable par les autres éditeurs de la même catégorie ou des catégories parentes. S'il n'y a aucun site à réviser, il est nécessaire de faire des recherches pour trouver des sites à insérer. Les nombreux outils mis à notre disposition rendent ces tâches assez aisées.

Nous disposons également d'un forum destiné à recueillir nos échanges, nos questions, nos demandes de changement, nos trucs et astuces, et ... nos coups de gueule! Certains sujets déclenchent les passions, (la tauromachie, les partis politiques...) mais notre sens de l'humour prend toujours le dessus, qu'il soit suisse, belge, français ou québécois! Une lecture des forums anglophones est parfois nécessaire pour recueillir des informations, et c'est par-là que je termine ma session d'édition, souvent très tard dans la nuit.

Q : Est-il possible de soumettre un site dans plusieurs catégories ? Un site bilingue peut-il être soumis à la fois dans la partie anglophone (www.dmoz.org) et dans la partie francophone (www.dmoz.ch) ?

R : Dmoz.ch présente simplement la partie francophone de dmoz.org. Les liens vers les catégories de dmoz.ch aboutissent en fait à dmoz.org. Il n'est donc pas nécessaire de faire un double référencement sur les deux versions de l'Open Directory. Une double soumission est cependant possible en cas de :

- site multilingue: surtout si le choix est présent sur la page d'accueil.
- caractère régional (au sens dmoz du terme, concerne un pays ou une région) Exemple: Le CHU de Rouen est référencé en catégorie Santé, et dans la catégorie Régional/France/Santé/Hôpitaux et cliniques.

En principe, un référencement dans deux catégories thématiques n'est pas possible. Le CHU de Rouen, ne pourrait être référencé en santé et en médecine.

Q : Comment jugez-vous la qualité de la partie francophone de l'ODP? Et de la partie mondiale (environ 2 millions de sites recensés) ?

R : La qualité de la partie francophone s'est beaucoup améliorée ces derniers mois, surtout depuis la traduction des interfaces. Avec 33 700 sites inscrits, dmoz est mieux connu des webmasters et le nombre de sites soumis en constante augmentation. Les différents articles parus sur Abondance nous ont bien aidés, comme la reprise de la base dmoz dans le directory Google.

Cependant, les catégories sont assez inégales, en fonction de l'engagement et de la qualité des éditeurs. Un expert d'un domaine aura à coeur de créer et maintenir une catégorie faisant référence. Par exemple, la catégorie Gynécologie et obstétrique est gérée par un médecin spécialiste, qui a visité et commenté tous les sites référencés. Un spécialiste du droit français sur Internet édite une catégorie en rapport avec son activité, tandis qu'un passionné de littérature a créé une riche arborescence de 1167 sites.

L'ensemble du répertoire a un taux très bas de "file not found", ce dont nous sommes très fiers! Un nettoyage régulier des sites en non-réponse permet de proposer à l'internaute des réponses valables.

Q : Les sites qui utilisent l'annuaire Dmoz proposent-ils tous la même version de l'outil mise à jour au même moment ? Est-ce qu'il y a un système de "réplication", toutes les semaines par exemples, sur les sites partenaires ?

R : Il s'agit certainement de la question la plus fréquemment posée au sujet de dmoz! La base est mise à jour chaque semaine (la nuit du lundi sauf erreur) et téléchargeable à cette adresse <http://dmoz.org/rdf/>) Dmoz n'a aucun contrôle de ce qui est fait avec ces données. Il appartient à chaque annuaire ou moteur de recherche de procéder au téléchargement et à la mise à jour. Certains le font chaque semaine, chaque mois, ou jamais... Sur l'original <http://www.dmoz.org> les données sont mises à jour immédiatement, sauf le moteur de recherche qui est indexé tous les 2-3 jours.

Q : Quels sont les délais entre une soumission et l'inscription?

R : Ils sont très variables d'une catégorie à l'autre, et dépendent vraiment de l'éditeur responsable. Cela peut aller de quelques heures (éditeur ei.

ligne, site de qualité, titre et description complète et correcte, soumission dans la bonne catégorie) à quelques semaines, voire quelques mois si la catégorie n'a pas d'éditeurs. Il y a même des catégories délaissées! Lors de la soumission d'un site, assurez-vous d'avoir choisi la bonne catégorie. L'Open Directory a une arborescence riche, et si vous recherchez la meilleure catégorie, votre site sera passé en revue dans de meilleurs délais. Une catégorie munie d'un éditeur avec de nombreux sites en attente est assez vite remarquée par un éditeur du niveau supérieur. Celui-ci peut éditer à tout moment, et il corrige en général la situation. En cas de délais anormalement longs, il est possible d'envoyer un message à l'éditeur, en cliquant sur le lien "Votre avis" au sommet de la page.

Q : Justement, comment se passe, dans la pratique, un référencement sur une catégorie qui ne dispose pas d'un éditeur attiré ? Qui effectue l'évaluation du site ?

R : Une catégorie sans éditeur ne signifie pas que les sites soumis ne soient pas traités.

Un exemple: Je suis éditrice de la catégorie World/Français. Je peux éditer dans toutes les catégories de cette branche, même si mon nom ne figure pas comme éditeur responsable de toutes les sous-catégories. Peuvent traiter les sites en attente: les 10 éditeurs de World/Français, les 60 éditeurs appelés "editall editors" et qui ont accès à tout le répertoire ainsi que les 40 "meta editors" chargés de l'acceptation des nouveaux éditeurs.

Q : Les éditeurs sont-ils difficile à trouver?

R : Chaque éditeur a développé une stratégies pour convaincre de nouveaux éditeurs, avec plus ou moins de succès. (Amis, famille, liste de discussion...) Le bénévolat est une notion parfois difficile à faire passer! Malheureusement, depuis la traduction des interfaces, le taux d'acceptation est anormalement bas pour les catégories francophones. Est-ce dû à la difficulté qu'éprouvent les "meta editors" non francophones pour juger des soumissions en français? A ce jour, nous n'avons pas reçu d'explication. Mais cette situation devrait changer rapidement, il en va de l'avenir de l'ODP francophone.

Q : Comment jugez-vous la qualité des fiches descriptives de sites que l'on vous propose ?

R : Être éditeur à dmoz permet de juger du sérieux des entreprises de référencement. Certaines fournissent des descriptions ne nécessitant aucune révision, soumettent les sites dans les bonnes catégories. D'autres en revanche... Vous pourriez sans crainte leur confier les sites de vos ennemis! Si la charte du référencier était appliquée, ces soumissions passeraient immédiatement à la trappe!

<https://www.abondance.com/docs/charte.html>

Si je comprends les difficultés des webmasters amateurs, je suis beaucoup moins indulgente avec des sociétés de référencement, dont les services facturés à prix d'or n'ont rien de professionnel. Visiter des milliers de sites à dmoz m'a appris beaucoup sur les techniques de référencement, et j'encourage les étudiants de ce domaine à venir nous rejoindre, c'est une excellente école!

Q : Pensez-vous que le schéma "Open" et bénévole pourra "tenir le coup" si l'Open Directory commence à gagner beaucoup d'argent?

R : L'Open Directory ne gagne pas d'argent : la base est fournie gratuitement à ceux qui le souhaitent.

<http://dmoz.org/World/Fran%e7ais/about.html>

Q : Quelles perspectives pour l'ODP?

R : Tant que l'Open Source (GNU, Gnutella...) existera, l'avenir de l'ODP est assuré, il fait partie intégrante du web.

Q : Le mot de la fin?

R : Vous avez une expertise, un hobby, un sport, un auteur ou chanteur préféré? Inscrivez-vous comme éditeur, une heure ou deux par mois peuvent suffire. Vous contribuerez ainsi à la construction du plus grand annuaire de l'Internet fait par des humains.

http://dmoz.org/Test/Help_Wanted/World/Fran%e7ais/Suggestions/

Merci, Michelle Carrupt (<http://www.mic.ch>) :-)

■ **Etude Hotbot**

Le moteur de recherche Hotbot, dans le cadre de son implantation en Europe, a demandé à une société indépendante de mener une étude sur la façon dont les internautes "voient" les outils de recherche à l'heure actuelle. Cette étude est basée sur le comportement d'un certain nombre d'internautes (deux groupes ont été formés : des utilisateurs expérimentés et des débutants) qui ont été questionnés directement sur leur utilisation des outils de recherche et dont les comportements face à ces outils a été "espionné en live". Les responsables de HotBot France nous ont fait parvenir les résultats de cette enquête (mille mercis à eux ;-)) dont voici les principales conclusions :

- Les utilisateurs font une différence entre le "surf", sans réel but, et la recherche d'information effectuée avec un objectif précis. La tendance est d'aller vers le deuxième comportement.
- L'outil de recherche est considéré comme un point de départ. Une majorité d'internautes positionnent un outil de recherche en page de démarrage de leur navigateur.
- Yahoo! et Lycos sont les outils les plus connus. Atavista, Excite et AOL sont des noms qui apparaissent également. Les utilisateurs confirmés utilisent plus de 5 outils de recherche, contre 1 ou 2 pour les débutants.
- Les raisons pour lesquelles un internaute choisit tel ou tel outil de recherche ne sont pas très claires. Il ne s'agit pas obligatoirement d'une raison de meilleure pertinence la plupart du temps, mais plutôt de renommée. D'ailleurs, de nombreux participants à cette enquête étaient incapables d'énoncer les points forts et les points faibles des outils qu'ils utilisent.
- Par exemple, voici quelques raisons qui ont poussé les personnes sondées à utiliser Yahoo! : "ils font de la publicité off-line", "je l'ai découvert dans la presse", "des gens me l'ont conseillé". Pour Altavista : "je l'ai découvert dans la presse", "des amis m'en ont parlé"... Lycos : "des amis m'en ont parlé", "parce qu'on peut envoyer des messages SMS avec". Etc. On voit que le marketing viral et le bouche à oreille électronique ont de beaux jours devant eux...

- Forces et faiblesses détectées. Lycos : forces = bonne ergonomie, pertinence. Faiblesses = aucune a priori. Yahoo! : forces = rien de particulier, Faiblesses = trop de pub. AOL: forces = à jour, rapide, bon pour le shopping, Faiblesses = mauvaise ergonomie. En règle générale, les forces demandées aux outils actuels sont les suivantes : proposer une ergonomie simple et efficace, être un point d'entrée sur le Web, être rapide, être un facteur de gain de temps, aider l'internaute à trouver une info qu'il est incapable de trouver tout seul, identifier des liens intéressants, nouveaux et inconnus jusqu'alors, obtenir des infos comme la météo et l'actualité (portail), des possibilités d'affiner les recherches, proposer des possibilité de liens vers d'autres outils de recherche. Les faiblesses, en règle générale : proposer trop de résultats, être trop compliqué d'utilisation, avoir une mauvaise ergonomie, afficher trop de pub.

- L'outil de recherche "idéal" doit donc être facile à utiliser, à "lire" (ergonomie), pertinent, rapide, à jour et exhaustif. Rien que ça ;-). Il doit offrir des résultats précis, être capable de filter l'information inutile, proposer une aide en ligne, offrir une possibilité de recherche avancée et s'adapter aux besoins et aux préférences de l'internaute. Résultats sur lesquels s'est reposé Lycos pour lancer HotBot Europe et répondre au mieux à ces exigences. Selon les sondés, aucun outil de recherche n'est aujourd'hui considéré comme "idéal", mais il n'est cependant pas ressorti de très gros griefs contre les annuaires et moteurs actuels. Les outils d'aujourd'hui semblent faire, selon les internautes "aussi bien qu'ils le peuvent".

Peu-être pourrai-je vous donner d'autres informations tirées de cette étude le mois prochain. Restez connectés... ;-)

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

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

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

Les moteurs à la peine ce mois-ci : Infoseek (dead :-), Voila, Excite.

■ Bruits et chuchotements

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

-> Copernic travaille sur une toute nouvelle version de son logiciel, soit la version 2001 qui devrait être lancée au mois de mars prochain. Cette toute nouvelle version sera basée sur les résultats de plusieurs sondages auprès de ses utilisateurs afin de connaître leurs attentes. Avec ces données en poche, Copernic a amorcé, il y a maintenant plus de 5 semaines, le développement de Copernic 2001 qui comportera des changements et ajouts importants par rapport à la précédente version.

-> De son côté, Francité est en train de préparer une nouvelle page d'accueil. Sa mise en ligne est prévue pour le 1er mars. Elle sera plus simple, moins colorée, le fichier HTML sera au maximum de 30 Ko. Certaines parties vont être supprimées et d'autres déplacées, ainsi que les partenaires mis plus en valeur.

-> Hotbot France sera officiellement lancé dans les heures qui viennent. Certaines fonctionnalités de recherche avancée ont été rajoutées au fur et à mesure de la "débétatisation" ;-) de l'outil. Tout devrait être terminé et fonctionnel au moment du lancement.

-> Sur la page d'accueil de Lycos France, 70% des utilisateurs effectueraient une recherche par mots clés, alors que 30% consulteraient l'annuaire. Ces chiffres sont en cours de confirmation par Lycos.

-> Google serait en train de durcir sa politique de lutte contre le spam. Avis aux amateurs...

-> Selon des informations émanant des principaux annuaires francophones, si vous avez un site disponible à l'adresse www.toto.com, il est tout à fait possible d'effectuer plusieurs soumissions sur les annuaires pour des adresses de type www.toto.com/theme1/, www.toto.com/theme2/, etc., en demandant pour chacune de ces sous-adresses des rubriques spécifiques. Cependant, une "zone" d'un site ne sera acceptée dans un annuaire (quelque soit celui-ci, quasiment) que si elle est pertinente (bien sûr) et uniquement si son contenu est disponible de façon autonome par rapport au site global (www.toto.com). Si l'info présentée dans une partie de site peut répondre à une problématique précise dans l'annuaire (si elle répond de façon précise aux demandes des internautes visitant l'annuaire) et si son contenu est tout à fait autarcique par rapport au site auquel elle est rattachée, il n'y a aucun problème pour qu'elle soit référencée. Ceci dit, c'est toujours l'annuaire qui décide, en dernier lieu, de l'acceptation ou du refus du site. Encore heureux... Ceci dit, si la situation est complexe, peut-être faut-il privilégier la demande par mail à celle faite sur le web, au travers de l'interface de soumission "classique"...

-> Petite info en provenance de Lycos US : "Pages that contain fewer than 100 words in the title and body text are excluded from our search. Our spider ignores the META and COMMENT tags because they cannot be seen on the page. However, we are able to index the ALT attributes of picture files. Perhaps another page from your site can be submitted instead. You may submit more than one URL from your site as long as the URLs represent distinct Web pages. Multiple pages that contain the same content are considered "spam" and will be excluded. "

-> Les pages soumises manuellement sur Inktomi seraient pénalisées dans les positionnements sur les portails utilisant ses index. Seraient privilégiées : les pages crawlées de façon "naturelle" par le spider, en suivant les liens des pages, ainsi que les soumissions payantes ("Paid Inclusion Program"). Inktomi indique (comme Altavista) que 95% des pages qui lui sont soumises au travers des fonctionnalités "Add URL" relèvent de spam... Du coup, Fast, Google, Excite et, donc, Inktomi, prendraient de plus en plus le pli de ne prendre en compte les pages soumises au travers de la fonction "Add URL" que comme des "suggestions", sans plus, contrairement aux pages capturées "naturellement", qui seraient, elles, intégrées beaucoup plus facilement dans les index et de mieux en mieux classées dans les pages de résultat. Altavista serait le prochain à fonctionner de la sorte... Cependant, une page soumise à la main mais qui serait également trouvée "naturellement" par le spider ne serait pas pénalisée. Normal...

■ En Bref

-> Si l'affichage de la page d'accueil de votre site provoque un redimensionnement de la fenêtre du navigateur de l'internaute, il y a de très fortes chances pour que votre site soit refusé par Yahoo! (source : Yahoo!). Un webmaster averti en vaut deux... ;-)

-> Fast annonce un délai de prise en compte, entre la soumission d'une page et son intégration dans l'index, de 3 jours. Wait and see...

-> L'annuaire de Lycos France comprend maintenant 70 000 sites intégrés dans 6 000 catégories.

-> La branche "sociétés" de l'annuaire de Lycos France, qui était en cours de création, se dévoile petit à petit sur l'outil de recherche. Vérifiez si une catégorie correspondant à votre activité a été créée dernièrement.

-> L'intégration entre Spray.fr et Lycos.fr est en cours. Certaines rubriques de Spray (Love@Lycos par exemple) sont venues sur Lycos France, et l'annuaire de Lycos est maintenant intégré sur Spray. A priori, le site www.spray.fr restera en ligne dans les mois qui viennent.

-> Google donne maintenant accès à une recherche sur 13 millions de documents PDF. Les documents sous ce format peuvent être recherchés avec la restriction "inurl:pdf" en complément des (et APRES LES) mots clés demandés. Exemple : [renault scénic inurl:pdf](http://renault.scénic.inurl:pdf). La fonction "inurl:" est proche de "allinurl:" sur Google. Le deuxième impose la présence de TOUS les mots demandés (ET), la première de certains seulement (OU).

-> Excite France a donc mis en ligne un nouvel index et un nouvel algorithme de classement prenant en compte l'indice de popularité. celui-ci est uniquement quantitatif : plus il y a de liens vers votre site, meilleur est son classement, quelle que soit la "popularité" des pages pointant vers vous. Un gros effort de dédoublement a également été fait, notamment dans l'index francophone (même s'il reste encore certaines pages en double qui sont activement chassées) qui comprend 8 millions de pages aux dernières nouvelles. L'objectif à court terme est de 11 millions de pages en langue française. Dans les résultats du moteur, un effort a également été effectué pour enlever les pages émanant directement du site Excite, ce qui était très souvent le cas jusqu'à maintenant, jusqu'à envahir parfois les X premiers résultats. Enfin, nous avons eu la confirmation qu'aucun système d'indice de clic, de type DirectHit, n'est encore utilisé jusqu'à maintenant sur Excite.

Selon Excite, un "rebuild" de l'index est effectué tous les quinze jours, en prenant en compte les soumissions des nouveaux sites soumis. Rappelons que ce délai est celui donné à titre officiel, tout comme le chiffre de 250 millions de pages en toutes langues de l'index ;-)

Excite France se bat et "met la pression" auprès de sa maison mère pour que l'index francophone soit le mieux à jour possible et comprenne le plus rapidement possible les sites soumis sur son portail. Mais tous les index sont traités aux Etats-Unis et seule la structure américaine d'Excite a la main-mise sur les aspects techniques liés au moteur (indexation, mise à jour, "rebuild", crawl, etc.). Pas toujours très facile...

