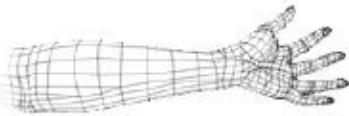
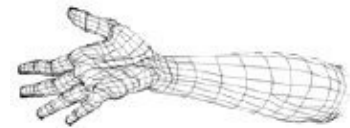


numéro nonante-quatre ( #94 )

le mercredi 12 novembre 2003



Sommaire



- [Sommaire](#)
- [Editorial](#)
- [Les entreprises et les outils de recherche](#)  
Les liens sponsorisés – partie VI
- [Le monde selon Google](#)  
Google Deskbar
- [Le référencement marketing](#)  
Trouver un compromis entre marketing et référencement
- [Humour et Internet](#)  
Excuse bidon
- [Référencement de A à Z](#)  
Référencement sur les annuaires
- [Optimisation et indexation commerciale](#)  
La croissance prévue du e-commerce n'aura pas lieu pour tous
- [Enfin, les informations](#)
- [Les outils de recherche francophones](#)

[Le dernier Moteurzine en html](#)

[Le dernier Moteurzine en PDF](#)

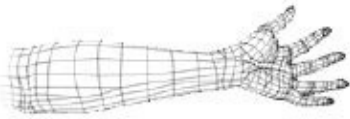
[Les archives de Moteurzine](#)

[Ecrire à Moteurzine](#)

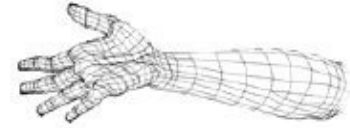


[E-commerce Internet Services](#)

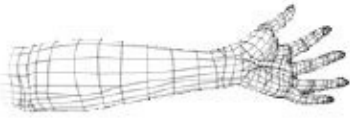
- L'article / l'entretien  
Le "topic sensitive pagerank"
- Conclusion



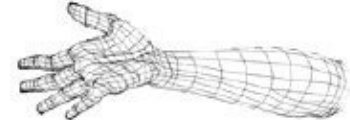
Espace publicitaire



– **vosre publicité lue par 25.000 personnes !** –



Éditorial



### **Un MoteurZine long, long, mais long...**

Bonjour à toutes et à tous,

Aujourd'hui, il y a énormément de chose à partager, ou tout simplement à dire.

Avant tout, je tiens à remercier Philippe YONNET qui s'est proposé de rédiger de manière irrégulière des articles plus techniques sur le monde des outils de recherche. Je suis persuadé que cela va satisfaire la curiosité de plus d'un d'entre vous !

Depuis quelques mois, je m'occupe de faire un peu de ménage dans la base de données d'Enfin. Ces derniers jours, j'avais à vérifier la validité de quelques référenceurs. Je dirais que 9 sur 10 ont été supprimés et pour diverses raisons... l'une d'entre elle, c'est la volatilité du lien en retour qu'ils ont pu mettre vers Enfin. J'entends tout de suite les grincheux : *on ne veut pas mettre de boutons qui va défigurer notre site et rien nous apporter ! on préfère payer !!*. Je préfère prendre les devants et répondre de suite. Mettre le bouton **Enfin.fr** sur votre site ne dévisage rien du tout, c'est un label de qualité. Si Enfin ne vous apporte pas de trafic, pourquoi y soumettre votre site ! Quand au fait de devoir *payer*, je constate régulièrement avec MoteurZine que c'est une belle parole qui ne dépasse pas ce stade.

Sinon, je reçois régulièrement (toujours sur Enfin) des soumissions de référenceurs qui offrent comme service des soumissions automatiques dans plusieurs milliers d'outils de recherche. Il va de soit que je vous conseille d'être prudent, c'est votre site que vous mettez en jeu si vous vous tournez vers ces sociétés. Je préfère vous conseiller mon propre outil de référencement qui n'a rien d'un robot terminatorien automatisé, mais qui fonctionne bien avec des résultats plus que satisfaisants. Le site, c'est référencement payant.

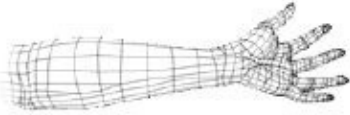
Dans ce numéro de MoteurZine, vous allez découvrir un interview de Alain SANJAUME, le Directeur Général d'Espoting Media en France. Il répond en toute franchise à 7 questions vérité !

Un petit message pour finir, si vous êtes intéressé, il y aura peut être du travail en freelance (avec facture et tout). Le travail correspond à quelques heures par jour à passer devant son

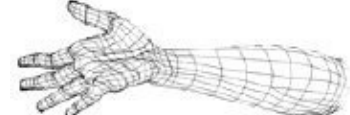
écran. Les contraintes... connaître internet, savoir parler français et être logique. Pour en savoir plus, la première prise de contact se fait via cet email : [freelance@idf.net](mailto:freelance@idf.net)

Bonne lecture...

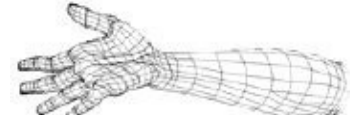
par CHRIS HÉDÉ



Rechercher sur Enfin



Les entreprises et les outils  
de recherche



### **Les liens sponsorisés – partie VI**

Nous vous avons promis d'obtenir une interview des dirigeants des sociétés OVERTURE et ESPOTTING pour clôturer notre série de chroniques consacrées aux grands acteurs du positionnement promotionnel.

Nous tenons à remercier Alain SANJAUME, Directeur Général de ESPOTTING d'avoir accepté de répondre à nos questions en exclusivité pour Moteurzine.

### **7 questions à Alain Sanjaume, Directeur Général d'ESPOTTING MEDIA France**

Votre entreprise est actuellement leader européen des liens sponsorisés. Pouvez vous nous parler un peu de votre propre parcours ?

A 23 ans j'ai co-fondé ma première entreprise, Eurosystems, qui officiait dans le secteur des services informatiques. J'ai développé cette entreprise dans plusieurs pays européens avant de la céder en 1998 à un groupe anglo-saxon, 4Front Technologies Inc.

J'ai ensuite assuré la direction et le développement de cette entreprise en France jusqu'en 2000, date à laquelle j'ai géré la fusion de ce groupe avec un grand acteur américain du marché : NCR Corp.

Ayant géré des sociétés de 2 à 900 salariés dont certaines cotées en bourse, j'avais envie de repartir dans la création d'entreprise. Je souhaitais aussi investir dans le secteur Internet, à commencer par un segment porteur de ce marché.

C'est fin 2001 que j'ai croisé le chemin d'ESPOTTING qui correspondait au-delà de mes espérances à mon projet personnel. Depuis deux ans, et grâce à l'expérience acquise, j'ai bâti et développé une entreprise sur des bases saines.

C'est une joie quotidienne que de gérer une équipe jeune et motivée qui porte très haut des

valeurs humaines de service clients, de flexibilité et de rapidité d'exécution. C'est par le professionnalisme de mon équipe que, tous les mois, nous affichons une croissance à deux chiffres.

A ce jour combien d'affiliés français donnent les résultats sponsorisés de ESPOTTING ?

Nous comptons plus de 150 affiliés en France aujourd'hui. Les plus connus sont Lycos, Copernic, Tiscali, Kartoo, La Poste, 9Telecom, Télé2, France 2 et France 3, Le Monde, La Tribune, Top Santé, Libération, FHM, NRJ, MCM,...

Vous noterez ainsi qu'ESPOTTING se démarque fortement de ses concurrents. En effet, ESPOTTING représente véritablement un **RESEAU** de sites affiliés, et non pas une « régie externalisée » pour quelques gros moteurs américains.

En effet au-delà d'une présence dans les moteurs nous sommes présent dans des sites communautaires (Seniorplanet, Gay.com, Autojournal, ?) où des carrefours d'audience permettent de toucher des micro-cibles très qualifiées pour nos annonceurs.

Nos intégrations sont variées, aussi bien dans les pages réponses des moteurs que dans des annuaires ou des chaînes thématiques.

Par exemple, un article de « Libération » traitant du tourisme en Grèce contiendra en bas de page des liens de nos annonceurs proposant des offres de voyage en Grèce.

Nos sites affiliés accueillent des internautes fidèles, plus « matures », possédant une expérience et une culture Web. Ce sont ces internautes qui n'hésitent pas à compléter des questionnaires pour obtenir des informations et acheter en ligne.

Sur les outils plus généralistes, le nombre de « butineurs » (autrement dit les nouveaux entrants sur le Web) et de « surfeurs » est beaucoup plus élevé. Leur attitude habituelle consiste à cliquer partout uniquement parce que c'est nouveau. Ces internautes correspondent plutôt à un public volatile et bien souvent à la recherche de produits ou services gratuits.

Ayant privilégié depuis notre création la qualité de notre réseau à la quantité, nous avons des taux de transformation très élevés. Le nombre d'études de cas que nous avons réalisées dans des secteurs aussi différents que le voyage, la rencontre ou l'e-learning, montrent bien notre avantage concurrentiel.

Nous avons donc des mois d'avance sur nos concurrents dans ce domaine. Et de nouveaux contrats vont encore être annoncés sous peu.

Quels sont les atouts majeurs d'ESPOTTING pour ses clients ?

Je vais développer ma réponse en trois points : **Meilleure qualité du service, Meilleur rapport qualité prix et meilleur réseau d'affiliation.**

Tout d'abord la qualité du service que nous apportons à nos clients. Vous parlez mieux à des gens qu'à des emails ;-)) Nous sommes les seuls à faire des formations dédiées à nos clients tous les vendredis.

Nous avons une équipe dédiée très importante, en relation permanente avec nos annonceurs, quelle que soit leur taille.

En effet, si les liens sponsorisés représentent aujourd'hui une des meilleures solutions du Web pour recruter des membres, des prospects, des nouveaux clients, il n'en demeure pas moins vrai qu'une campagne doit être suivie et optimisée pour obtenir des résultats conséquents.

Ce service est réalisé en interne avec notre personnel français. Nos équipes ont une parfaite connaissance de tous nos clients et de leurs problématiques. Dans la majorité des cas nous anticipons les demandes des clients en les contactant ou en leur envoyant des informations permettant de résoudre un problème ou de mieux optimiser leurs campagnes.

Avec une enchère minimale de 10 centimes d'€ et un prix moyen au clic des plus raisonnables, ESPOTTING vous permet d'obtenir un excellent retour sur investissement.

Une campagne ESPOTTING représente un des meilleurs rapports qualité/prix actuels du marché. N'oublions pas que contrairement à d'autres modes de promotion, un annonceur ESPOTTING conserve à tout moment la maîtrise de ses dépenses !!!

Et bien entendu, la **qualité de notre réseau d'affiliation** qui comme je le précisais à la question précédente, apporte à nos annonceurs un trafic important en volume mais surtout des plus qualifiés.

Existe-t-il une taille minimale d'entreprise pour aborder les liens sponsorisés chez ESPOTTING ?

Certainement pas. J'ai même envie de dire qu'avec un système comme le notre, tous les annonceurs sont sur un pied d'égalité et ne sont pas forcément dépendant de leur budget marketing pour se battre sur des marchés parfois fortement concurrentiels. Plus de 80% de nos annonceurs sont aujourd'hui des PME.

Nous avons constaté que des sociétés de taille plus modeste tenaient tête dans les enchères à des sociétés plus importantes, tout simplement parce qu'elles avaient des taux de transformation meilleurs et donc des coûts d'acquisition client plus faibles. Une société qui a des bons produits avec un positionnement tarifaire intéressant pour le client final a toute sa place sur notre réseau.

Avez-vous un programme spécifique pour les professionnels qui distribuent vos produits ?

Nous assurons, bien sur, depuis notre site [la promotion de nos distributeurs](#). Mais, au-delà, nous mettons à leur disposition des outils concrets tels que notre interface API qui permet de gérer automatiquement les enchères ou d'afficher notre interface client aux couleurs de notre distributeur.

Nous donnons à nos distributeurs une véritable valeur ajoutée face à leur client. [Jeremy Garamond](#) est à leur disposition pour étudier leurs demandes.

A l'heure où nos sources de trafic sont de plus en plus complémentaires aux prestations de référencement nous voyons un afflux de demandes de SEO's . Les distributeurs les plus

actifs dans ce domaine sont entre autre : Synodiance, @position, 1èreposition, Brioude Internet/Takezo, ?

Vous avez annoncé officiellement votre rapprochement avec la société américaine FINDWHAT. Que va concrètement apporter ce regroupement pour les clients d'ESPOTTING ?

Tout d'abord, FINDWHAT et ESPOTTING n'ont pas consommé le mariage car certains points restent à discuter. Nous devrions en savoir plus d'ici la fin de l'année.

Ceci dit la logique industrielle de rapprochement est évidente : les deux sociétés ont un positionnement et un fonctionnement très similaire par rapport à tous les points évoqués précédemment (qualité du réseau de distribution, service client, politique éditoriale). Il était donc assez naturel pour ces deux entités de se rapprocher pour former un acteur international incontournable de ce marché.

Bien entendu, **ce rapprochement va permettre à nos annonceurs internationaux à court/moyen terme de pouvoir annoncer outre atlantique à partir d'un seul point d'entrée**. Cette facilité d'action et d'utilisation de nos outils et de nos services est très appréciée par les annonceurs. En un coup de téléphone, un email, avoir la possibilité de toucher 100% de sa cible potentielle est un avantage considérable.

Pour terminer, que souhaiteriez vous dire à nos lecteurs qui hésitent encore à devenir annonceurs ESPOTTING ?

Si en moins de 3 mois nous sommes passés de 16000 à 20000 clients nous n'avons plus à prouver l'intérêt du marché pour notre solution. De plus, 92 % de nos clients nous sont restés fidèles depuis notre création démontrant ainsi leur attachement à notre produit,

ESPOTTING est une société rentable et pérenne qui affiche un chiffre d'affaire dépassant les 100 millions d' €

Avec un milliard de requêtes par mois, ESPOTTING vous offre un maximum de chance de toucher un prospect qualifié, et ceci dans dix pays européens.

Le coût d'entrée pour une prise en charge **totale** de votre compte est de 259 €. Ce qui vous permettra de bénéficier de nos années d'expérience et d'expertise dans ce secteur !

### **ESPOTTING**

28 Bld Haussmann

75009 Paris

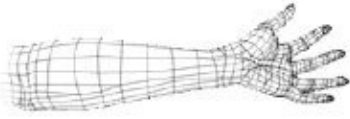
Téléphone : 0 825 377 681

[Informations générales](#)

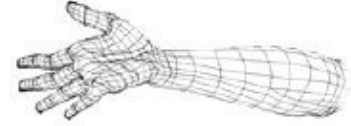
[Service commercial](#)

[Affiliation](#)

par Jean-Claude BENARD de Quadramultimédia



Le monde selon Google



## Google Deskbar

Google vient de lancer un nouveau petit logiciel à installer sur un PC sous Windows. Il permet de lancer des recherches sur Google à tout instant, sans quitter l'application en cours, et sans ouvrir de navigateur web !

La Google Deskbar crée une petite zone de saisie en bas de l'écran, près de l'horloge, dans laquelle on peut taper quelques mots depuis n'importe quelle application. Une recherche sur Google est alors automatiquement lancée et les résultats s'affichent dans une fenêtre qui s'ouvre en bas à droite. Cette petite fenêtre est en fait une version réduite d'Internet Explorer. Elle peut être agrandie, et d'un clic on peut se retrouver dans son navigateur par défaut.

Des raccourcis clavier permettent d'accélérer encore les recherches. Il suffit en effet de sélectionner un ou plusieurs mots n'importe où dans un logiciel sous Windows, et de taper CTRL-ALT-G puis Entrée, pour voir apparaître immédiatement les 10 premiers résultats.

Quelques options permettent de paramétrer la Google Deskbar. Ainsi, on peut choisir le raccourci clavier pour remplacer CTRL-ALT-G s'il est trop compliqué. On peut désactiver les statistiques qui sont envoyées à Google (nombre d'utilisations de chacune des fonctionnalités de la barre) : on aurait préféré que par défaut aucune information ne soit envoyée à Google...

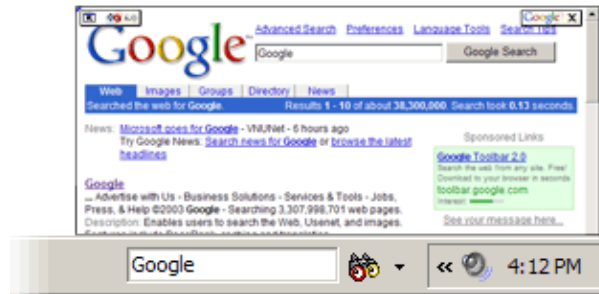
Plus étonnant encore, il est possible de paramétrer la barre pour qu'elle lance des recherches sur d'autres sites que Google. J'ai testé cette option et cela fonctionne bien : par exemple, en indiquant l'URL des recherches sur WebRankInfo avec le format imposé par la Google Deskbar, et en choisissant un raccourci-clavier, je peux lancer des recherches sur mon site préféré (!). C'est vraiment pratique et très rapide : en consultant une page web ou un document sur mon ordinateur, je sélectionne un ou plusieurs mots, je tape F8 (mon raccourci à la place de CTRL-ALT-G que je trouve trop long) puis CTRL-K (mon raccourci pour les recherches sur WebRankInfo) et me voilà avec les résultats qui apparaissent !

Pour l'instant la Google Deskbar n'est disponible qu'en anglais, en version beta, malheureusement uniquement pour les utilisateurs de Windows. Il est néanmoins possible de choisir d'effectuer les requêtes sur le site google.fr par exemple.

La Google Deskbar présente des avantages indéniables pour l'internaute. Et si un grand nombre d'internautes se mettaient à l'utiliser, c'est Microsoft, un grand concurrent attendu de Google, qui pourrait en pâtir. Celui-ci cherche en effet à maîtriser le "bureau de Windows" (desktop), et à intégrer la recherche sur le web au coeur même de Windows (c'est un des grands objectifs de Longhorn, la prochaine version du système d'exploitation de Microsoft).

Après le grand succès de sa Google Toolbar, Google gagnera-t-il le coeur du système d'exploitation de Microsoft avec sa Google Deskbar ? En tout cas, l'actualité nous montre comment Google répond à la rumeur au sujet du rachat de Google par Microsoft...

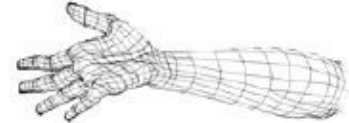
Pour plus d'infos : le site officiel et le groupe de discussions.



par Olivier DUFFEZ de WebRankInfo



Le référencement marketing



### **Trouver un compromis entre marketing et référencement**

Dans l'introduction au référencement marketing le mois dernier nous étudions l'intérêt de l'optimisation marketing du référencement. Et notamment, comment permet-il de drainer du trafic supplémentaire simplement en améliorant l'affichage du site dans les pages de résultats.

Si l'intérêt est prouvé, il reste toutefois une question en suspend : comment améliorer la qualité de son affichage dans les résultats sans porter préjudice au positionnement du site ?

Ce risque existe et si le positionnement du site décrochait de manière importante dans les pages de résultats, l'approche n'aurait plus d'intérêt.

Aujourd'hui, cette fiche pratique de méthodologie vous explique en 4 points comment améliorer votre image dans les résultats tout en préservant le travail de référencement déjà accompli.

### **Se projeter à la place de l'internaute**

Se projeter à la place de l'internaute, c'est aussi réfléchir à la manière dont vous réagissez lorsqu'une page de résultats apparaît.

Prenons un exemple.

Le contexte : je souhaite acheter du foie gras par internet. Je n'y connais rien dans ce domaine, je sais simplement que d'expérience, le sud ouest est LA région du foie gras en France.

Je tape foie gras dans Google et je tombe sur :



[Le site officiel des professionnels du foie gras](#)

Le site officiel des professionnels du **Foie Gras** : les adresses des producteurs et de nombreuses recettes mais aussi ses vertus et des conseils pour le ...

Description: Le site officiel des professionnels du **foie gras** : les adresses des producteurs en France et des recettes...

Catégorie: [World](#) > [Français](#) > [Maison](#) > [Cuisine](#)

[www.lefoiegras.fr/](http://www.lefoiegras.fr/) - 4k - [En cache](#) - [Pages similaires](#)

[Le site officiel des professionnels du foie gras](#)

Test. Test. Chargement en cours, veuillez patienter ... Mailing list gourmande, Les Producteurs en régions, Espace Journalistes, Espace ...

[www.lefoiegras.fr/default.asp](http://www.lefoiegras.fr/default.asp) - 16k - [En cache](#) - [Pages similaires](#)

[Foie Gras et Gastronomie du Sud Ouest chez vous en 48/72h !](#)

**Foie Gras**, Jambon basque et autres saveurs du Sud Ouest livrés chez vous en 48/72h ! foie\_gras foie\_gras **Foie Gras** du Sud ouest : ...

[www.thefrenchluxury.com/](http://www.thefrenchluxury.com/) - 3k - 9 nov 2003 - [En cache](#) - [Pages similaires](#)

[Lucien Doriath - Foies Gras d'Alsace](#)

Entreprise artisanale de production de **Foie Gras** et produits dérivés du canard et de l'oie.

Description: Producteur de **foie gras** artisanal en Alsace. Visite, repas à la ferme, tarifs et vente de produits...

Catégorie: [World](#) > [Français](#) > ... > [Achats](#) > [Gastronomie et alimentation](#)

[www.lucien-doriath.fr/](http://www.lucien-doriath.fr/) - 8k - [En cache](#) - [Pages similaires](#)

[Sonoma Foie Gras Home Page](#) - [ [Traduire cette page](#) ]

**Foie Gras** This site is temporarily unavailable. More Information.

[www.sonomafoiegras.com/](http://www.sonomafoiegras.com/) - 5k - [En cache](#) - [Pages similaires](#)

[Sonoma Foie Gras - Products](#) - [ [Traduire cette page](#) ]

What Is **Foie Gras**? ... Grades of **Foie Gras**. **Foie Gras** A is the top grade of **foie gras**. It is silky smooth and buttery and will melt in your mouth when sautéed. ...

[www.sonomafoiegras.com/product.html](http://www.sonomafoiegras.com/product.html) - 6k - [En cache](#) - [Pages similaires](#)

[Foie gras de Strasbourg - Edouard Artzner - Strasbourg. Alsace. ...](#)

---

A votre avis, sur quel résultat vais-je cliquer ?

Voici le cheminement de ma recherche :

1– Les deux premiers résultats sont un site officiel des professionnels du foie gras. Cela ne m'intéresse pas puisque je cherche pour acheter du foie gras.

Je passe mécaniquement mon chemin.

2– "Foie gras, gastronomie du Sud Ouest en 48/72h !". Voilà qui tombe bien ! Cela répond à mon critère de recherche, et en plus, je sais que je serai livré rapidement. Pour terminer, ce n'est pas mon besoin principal mais comme le site semble vendre des produits du sud ouest, je vais peut-être trouver un petit blanc sympathique gorgé de soleil pour l'accompagner.

Mon choix est déjà fait. Je prends tout de même la peine de regarder les autres résultats.

3– "Lucien Doriath – Foies Gras d'Alsace". Je suis tellement inculte dans ce domaine que je me demande si Lucien Doriath est un producteur de renommée en France. Au risque de faire un affront aux connaisseurs, je vous avoue que je passe mon chemin : je suis déjà convaincu par le résultat précédent et ce résultat n'accroche pas particulièrement.

Je jette un rapide coup d'oeil sur les suivants puis je clique sur The French Luxury.

Lorsque je clique sur ce site, je suis déjà un prospect sans même l'avoir regardé. Il répond au besoin et peut-être même plus (il vend peut-être du bon vin blanc sucré), je sais que je suis livré rapidement, et il est largement au dessus du lot par rapport aux concurrents.

### **Enseignements**

Je tiens à vous rassurer, je n'ai aucun lien avec ce site et pour assurer une impartialité de rédaction, je n'ai effectivement cliqué sur le lien qu'une fois cet article terminé.

Vraisemblablement, ce site a compris l'intérêt du référencement marketing : le site est bien référencé et en plus, sa présentation surpasse ses concurrents dans les pages de résultats.

A partir de cet exemple, chacun est à même de vérifier en quoi ses pages sont bien ou mal présentées dans les résultats de recherche. Et nous le verrons, comment améliorer la présentation sans porter préjudice au positionnement du site.

### **Trouver le compromis en 4 étapes**

#### *1– Trouver une cible et adapter les mots-clés pour chaque page*

Il est inutile d'envisager un référencement sans déterminer à qui il s'adresse.

Le choix des mots-clés à optimiser doit donc être le plus précis possible. Et différent sur chaque page. Cela permet de déterminer précisément comment doit-on vous trouver. Et aussi, de faciliter la présentation en page de résultats.

Cela doit même s'appliquer aux pages "annexes", comme les pages "A propos de notre société" , ou encore "Nous contacter".

Par exemple, sur une page "A propos", pourquoi ne pas choisir des combinaisons de mots-clés secondaires ?

Si la société est dans le chauffage électrique, il est possible de positionner sur des mots comme "fabricant chauffage électrique", "conseil chauffage électrique", "installation chauffage électrique" ou d'autres combinaisons relatives aux activités du site et tapées dans les moteurs.

#### *2– Optimiser le référencement de chaque page*

Ces mots-clés et combinaisons de mots clés en main pour chaque page, vous devez effectuer le travail nécessaire pour optimiser le référencement de chacune. Pour cela, les Chroniqueurs de MoteurZine vous aide par le biais de leurs articles.

#### *3– Améliorer l'affichage de chaque page*

Nous revoici avec notre entreprise de chauffage électrique et sa page "A propos". A ce stade, la page est positionnée sur des mots-clés et combinaisons de mots-clés précis. Le prévisionnel de trafic est restreint comparé aux pages principales. Mais très ciblé.

L'approche est de dire : au lieu de mettre en balise titre ressemblant à quelque chose comme : "A propos de Ma société de chauffage électrique" – qui n'attirera aucun visiteur –, autant exploiter la place avec les mots-clés et optimiser l'affichage des résultats.

Il est ainsi possible de transformer cela en :  
"Chauffages électriques : fabrication, conseils et installations".

Tout va tenir dans la page de résultats et tous les mots-clés s'y trouvent. On peut dire que vous avez exploité le potentiel de la page sur des mots ciblés, qu'en page de résultats la page va apparaître correctement, et pour finir, qu'elle vous fournira des prospects qualifiés.

D'une manière générale, plus la présentation est ciblée, meilleurs sont les résultats. Allez à l'essentiel, créez quelque chose de lisible par les humains, et ciblé aux mots-clés se trouvant dans la page.

#### *4– Vérifier les résultats et modifier en conséquence*

Certains tâtonnements sont nécessaires, notamment par rapport au choix des mots-clés et dans la façon de présenter les pages dans les résultats. La meilleure approche est d'essayer 2–3 versions par page, et de regarder les résultats du trafic en fonction des rafraîchissements dans les moteurs de recherche.

#### **Il existe aussi certains écueils à éviter.**

– Ne mettez pas de noms dans les endroits visibles en pages de résultats. Par exemple, un titre de page "Durand Immobilier, annonces immobilières dans le lot" n'est pas optimisé. Personne ne connaît Monsieur Durand sur Internet.

Le secteur étant ultra concurrentiel, un espace non négligeable n'est pas exploité. A cela, il faudrait préférer par exemple : "Immobilier : Nos annonces gratuites dans le Lot".

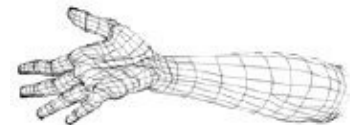
– Evitez les listes de mot-clés et les styles télégraphiques. Les présentations du genre "MonSite : télévision, programmes, programmes télé, tv, chaînes télévision", ou encore "Télévision / Programmes de télévision – Grille des programmes" n'obtiennent pas de bons résultats.

Dans le prochain numéro nous étudierons avec plus de précision l'optimisation de chaque balise META.

par Gautier GIRARD de **Cerise Rouge**



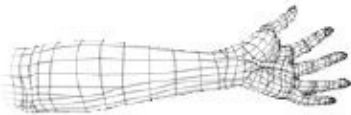
Humour et Internet



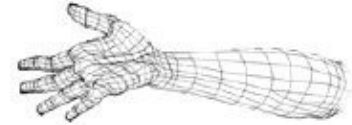
## Excuse bidon



par Frédéric LEPAGE de **Les BD qui bougent**



Référencement de A à Z



### Référencement sur les annuaires

Après avoir choisi les bons mots clefs pour la bonne cible, optimisé notre site, il faut le référencer.

Parmi les outils de recherche, nous distinguerons les annuaires des moteurs de recherche.

### Nous allons commencer par les annuaires.

Les annuaires sont tenus à jour par des documentalistes qui sont donc des humains et non pas des robots.

Les documentalistes indexent les sites en les classant par thématique, rubrique et sous rubriques.

Il existe des annuaires généralistes et des annuaires thématiques, ils sont extrêmement nombreux sur le web.

Lorsque l'on soumet un site à un annuaire, il faut au préalable bien avoir identifié la rubrique où celui-ci a sa place. On peut ainsi naviguer dans les catégories ou faire des recherches sur des mots clefs spécifiques à notre activité pour voir quelles sont les rubriques qui apparaissent dans les résultats. La localisation géographique de l'activité du site peut avoir son importance dans le choix de la catégorie. Si vous hésitez entre plusieurs catégories, la présence de vos concurrents dans l'une d'elle est un indice. Si en plus la catégorie contient peu d'inscrits et n'est pas trop loin de la racine de l'annuaire, le choix en est facilité.

Dans la catégorie choisie on aura souvent un lien pour soumettre un site.

La soumission doit être préparée et soignée, le travail du documentaliste en sera facilité et les chances de voir son site examiné rapidement augmentées.

La plupart du temps, lors d'une soumission à un annuaire il faut soumettre :

- un titre pour votre site
- une description
- l'url
- email du responsable

Attention, le titre et la description sont limités en taille, environ 200 caractères pour une description, il ne s'agit pas d'écrire un roman mais d'être concis et précis. L'utilisation des mots clefs dans la description du site est judicieuse dans la mesure où l'on peut rechercher dans un annuaire avec des mots clefs mais ne faites pas une énumération!

Ce descriptif peut être modifié par le documentaliste, les annuaires s'en réservent le droit. Evitez le langage trop commercial et d'écrire tout en majuscule.

D'autres renseignements annexes peuvent être demandés pour l'inscription.

Une fois le site soumis, son intégration dans l'annuaire peut prendre quelques jours à plusieurs mois dans le cadre d'offre gratuite.

Le référencement payant dans les annuaires s'est largement imposé, permettant l'indexation rapide (7 jours ouvrés le plus souvent) dans l'annuaire.

Nous allons soumettre notre site [Logos-sonneries.biz](http://Logos-sonneries.biz) dans l'[annuaire AOL](#).

Une recherche sur *téléphone mobile* nous fait sortir plusieurs catégories dont une *Sonneries et logos* qui convient tout à fait à notre site.

Dans cette catégorie, en bas de page, nous avons un lien pour *suggérer un site*.

Il faut ensuite accepter les conditions qui sont tout à fait cohérentes (proposer un site en français, ne faisant pas la promotion d'activités illicites, etc.) et soumettre notre site :

**Titre** : Logos et sonneries .biz

**Description** (170 caractères maximum) : Téléchargement de logos et sonneries pour téléphone portable parmi un choix varié de sonneries polyphoniques, logos animés et messages pour répondeurs.

Sur la page suivante votre nom et votre mail vous sont demandés afin de vous envoyer un mail de confirmation. Il faut répondre à ce mail sous 48h pour valider l'inscription.

Les plus gros annuaires francophones généralistes sont **Yahoo, Dmoz, Aol, Voilà, Lycos**. Certains font payer la soumission aux sites à vocation commerciale. Cette soumission payante est souvent la garantie de voir son site examiné rapidement et nullement une garantie d'indexation.

Il existe une multitude d'annuaires thématiques dans lesquels vous devez inscrire votre site, pour les trouver utilisez [Enfin.com](http://Enfin.com) ou [Lesannuaires](http://Lesannuaires.com).

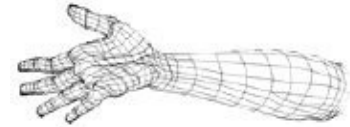
Le trafic généré par ces outils thématiques est certes moindre mais plus ciblé et c'est là le plus important.

Ne négligez donc pas les annuaires secondaires !

par Marie POURREYRON d'**Altitude Référencement**



Optimisation et indexation  
commerciale



### **La croissance prévue du e-commerce n'aura pas lieu pour tous**

Dans un article récent, le département du commerce américain annonçait que les ventes Internet passeront de 80 milliards à 240 milliards, dans les prochains mois. Les entreprises les plus éveillées – et compétentes, ne manqueront pas d'harnacher cette prochaine vague de croissance économique.

Mais voilà qu'une entreprise qui retrouve son site Internet dans le Web Invisible pourrait manquer le bateau.

Le Web Invisible décrit, d'une part, les milliards de sites Internet que les engins de recherche ne peuvent retrouver. Malgré l'avalanche des sites ramenés à chaque requête, la frontière pratique de cette invisibilité commence aussi, d'autre part, souvent dès la deuxième page, et certainement après la troisième. Ces deux formes d'invisibilités se valent.

### **LE CHAOS MALLÉABLE DU WEB**

En effet, qui donc s'aventure au-delà des trois premières pages avant d'abandonner, de changer d'engin de recherche ou tout aussi simplement, de changer de mot clé. Il est tellement facile de rephraser une recherche pour tenter d'obtenir de meilleurs résultats et le chaos du Web est tellement malléable. On a constaté que les internautes utilisent de plus en plus de mots dans leurs expressions clés, 2 dans le passé, jusqu'à 5 maintenant. Certains recommandent même des phrases entières! L'Internaute veut un maximum de vingt propositions de sites, plutôt que cent mille.

Suite à une enquête, le *Georgia Institute of Technology* a démontré que seulement 16 pourcent des sites Web soumis aux engins et répertoires de recherche sont effectivement indexés. Le 74 pourcent restant est introuvable suite aux erreurs de codage de leurs pages Web! En 2002, le World Wide Web consortium (W3C) révélait que seulement 3% des sites de ses membres étaient conformes aux normes de l'Internet – les siennes...! Que peut-il en être des autres?

### **TOUT SE JOUE AVANT... LA 2ième PAGE**

Bien que 55 pourcent des achats en ligne débutent par une recherche satisfaite auprès d'un engin de recherche, selon *Jupiter Media Metrix*, environ 93 pourcent des Internautes ne

dépassent pas la deuxième page des résultats.

Ces statistiques laissent comprendre que pour tout site transactionnel, ou sollicitant d'autres actions concrètes, le positionnement dans les deux premières pages représente l'avantage unique d'effectuer de nombreuses ventes, même par défaut!

Alors si, dans le meilleur des mondes, un engin de recherche retrouve le site Web d'une entreprise, son positionnement hors des deux premières pages semble donner le même résultat que de ne pas l'avoir trouvé!

### **OPTIMISA... QUOI?**

Dans un article récent paru dans une infolettre québécoise d'affaires importantes, un chroniqueur faisait la synthèse des méthodes publicitaires courantes disponibles aux entreprises, sans faire mention de l'optimisation des contenus Web. Bien que cette méthode soit la plus efficace, la plus économique et la plus rentable, cette omission décrit très bien l'addiction des entreprises aux modèles dispendieux des médias traditionnels qui vendent encore leurs lecteurs à l'unité, selon le célèbre CPM des années 1950 ou selon le modèle plus récent du PPC. Ainsi, une entreprise qui paierait jusqu'à \$4.62 pour son visiteur (PPC), sera affichée à côté d'une entreprise qui n'en débourse que \$0.02 pour le sien, avec l'optimisation.

Cette optimisation, qui gagne à être connue, est justement ce que la revue américaine *Brand Week Online* appelait "(...) *non seulement une partie des moyens promotionnel (d'un site Web), mais la partie la plus fondamentale. C'est le point de démarrage d'un site. Si vous ne faites rien d'autre, l'optimisation (...) par mots clés peut générer de 80 à 90 pourcent du trafic d'un site, ce qui se traduit par plus de ventes.*" Je connais des entreprises qui ne jurent, maintenant, que par cette "sainte optimisation", après avoir dépensé des fortunes en achats de mots clés (PPC).

### **CONNAÎTRE QUI NOUS RECHERCHE... ET LEUR NOMBRE**

Une autre étude récente témoignait que 72 pourcent des propriétaires de sites Web ne connaissent pas leur taux de conversion de leurs visiteurs Web en clients, ni leur coût d'acquisition de ces visiteurs. En fait, peu se soucient même de savoir par quels mots clés leurs visiteurs les trouvent. Dans pareil cas, toute interrogation sur le coût par vente est superflue...

Quand l'on sait que même avant la mise en ligne d'un site, la simple consultation d'un groupe d'utilisateur du Web (*Web Focus Group*) qui s'intéresse à la future niche de marché d'une entreprise permet non seulement d'obtenir très exactement le nombre total d'Internautes à la recherche d'un tel site, mais aussi le nombre et le libellé exact de tous les mots clés demandés pour y arriver. Ceux qui l'ignorent croient être recherchés par 2, 3 ou même 5 termes. Ceux qui consultent ces *Web Focus Groups* constatent que c'est souvent plus de 100 mots clés distincts que leurs futurs clients potentiels utiliseraient pour les trouver – si ces sites s'en servaient, si ces sites le savaient...

### **L'ISOLEMENT DES RESPONSABLES**

Ces analyses, centrées sur les canaux de revenus du Web – la rentabilité tant recherchée et décriée, nous éloigne des considérations visuelles et techniques de l'univers de la presque

totalité des webmestre, créateurs isolés de la plupart des sites Internet de la planète.

Cet isolement vient du fait d'une centration unique sur les dimensions internes d'un site Web – son codage primaire et son imagerie, et sur l'absence de responsabilité des technologues quant à la rentabilité obtenue par de tels sites en terme de visibilité adéquate, disponible lors d'ajouts de contenus textuels et de codage spécifique au marketing d'un site, de ses produits et service, bref, de l'entreprise. Tant qu'aucune équipe multidisciplinaire ne mettra l'épaule à la roue – proprios en tête, c'est ultimement ces donneurs d'ordres de tels sites qui sont ou seront, en dernière analyse, responsabilisés.

## LA MEILLEURE VOITURE AU MONDE

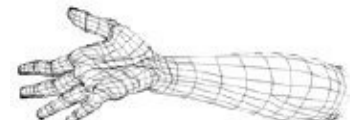
Un jour, Henri Ford demanda à un de ses ingénieurs de lui construire la meilleure voiture au monde. Ce dernier lui revint avec une voiture sans portes! Quand Ford lui demanda le pourquoi de cette anomalie incroyable, l'ingénieur lui répondit que la voiture était plus solide ainsi. Ford lui rétorqua ces deux mots célèbres : Pour qui?

*Merci et à bientôt!*

par Pierre FRIGON de l'**Agence Hyperclics Marketing**



Espace publicitaire



E-commerce Internet Services



Référencement et promotion de sites

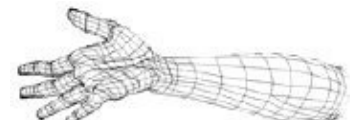


Webmarketing & Référencement

Vous ICI ???



Enfin, les informations



Vous êtes le responsable d'un moteur ou d'un annuaire ? Vous travaillez dans le référencement ? Communiquez et insérez gratuitement votre actualité ici.

## AUJOURD'HUI

ou le monde impitoyable du référencement et des outils de recherche

**08/11/2003 : "Google lance le nouveau logiciel Deskbar"**

« Google vient de lancer un nouveau logiciel bêta qui permet de faire des recherches sur Internet de la barre de tâches du bureau de Windows.

Avec la Deskbar, vous pouvez rapidement faire des recherches dans Google sans avoir besoin pour cela d'une fenêtre de votre fureteur Web préféré. Pour lancer une recherche, il suffit de saisir des mots-clés dans le champ installé à cet effet dans la barre de tâches de Windows, généralement placée au bas de l'écran, puis d'appuyer sur la touche Enter. Les résultats apparaissent alors dans une petite fenêtre qui



semble jaillir de la barre de tâches. »

**08/11/2003 : "Le marché du blog français attise les convoitises"**

« Comment rentabiliser le succès du Weblog, le nouvel espace d'expression à la mode sur Internet? C'est la question que se posent éditeurs, portails, fournisseurs d'accès et autres spécialistes de la publicité ciblée. »

**07/11/2003 : "Arisem s'adosse à un fonds canadien"**

« L'éditeur français de logiciels pour Internet et intranets spécialisé dans le knowledge management (30 salariés pour un chiffre d'affaires de 2,5 à 3 M€ selon la direction) créé en 1996, vient d'intégrer le fonds d'investissement canadien Emergo. Ce dernier détient dorénavant une participation majoritaire dans l'éditeur, opération dont le montant n'a pas été communiqué. » (...)

**06/11/2003 : "Selection de liens sur le SVG"**

RCI Informatique a mis en ligne une série de liens sur le langage SVG, le langage graphique vectoriel standard pour Internet. Ce langage graphique, basé sur XML, convient aux domaines suivants : cartographie, graphiques de gestion, architecture interactive, affichage typographique, graphiques scientifiques, illustration, etc.

**06/11/2003 : "L'alliance Overture–Yahoo profite à Google"**

« Depuis son rachat par Yahoo, la société Overture perd des clients. Plusieurs grands portails internet ont choisi de mettre un terme à leur contrat de fourniture de liens sponsorisés. Pour se tourner vers le concurrent Google. »

**05/11/2003 : "Espotting et Emap signent un contrat pluriannuel"**

« Le spécialiste du lien promotionnel étoffe son réseau de partenaires et signe avec Emap, le numéro deux de la presse en France. Les liens sponsorisés Espotting apparaîtront ainsi dorénavant sur les différents sites du groupe Emap (Top Santé, Téléstar, Télépoche, FHM, l'AutoJournal, France Golf, Bateaux, Addx, Auto Plus, Canal Stars...). »

**05/11/2003 : "Moto net one référence plus de 1500 sites"**

Moto Net One est l'annuaire francophone dédié aux sites moto. Créé il y a un an, Moto Net One vient de passer la barre des 1500 sites référencés. Un "relookage" est prévu dans les mois à venir.

Moto Net One propose l'affichage de son annuaire, en marque blanche, aux sites motos de qualité qui désirent diversifier leur contenu.

**05/11/2003 : "Lycos est mécontent du rachat de Overture par Yahoo!"**

Lycos attaque Overture car celui-ci n'a pas respecté les accords du contrat les liant. En effet, Overture devait consulter Lycos dans le cas du rachat d'Overture par un concurrent de Lycos, ce qui n'a pas été fait. Lycos déclare donc le contrat le liant à Overture est rompu et affiche désormais les liens sponsorisés de Google.

**05/11/2003 : "Terra Lycos prévoit un bénéfice opérationnel au 4e trimestre"**

« Le groupe espagnol de services internet Terra Lycos a réduit sa perte nette de 59% sur les neuf premiers mois de l'année et anticipe de dégager pour la première fois un bénéfice au niveau opérationnel au quatrième trimestre. » (...)

**04/11/2003 : "Ask Jeeves, un nouveau PDG vient d'être nommé"**

Steve Berkowitz vient d'être nommé PDG de la société Ask Jeeves. Steve Berkowitz a été interviewé par Enfin en 2002.

**04/11/2003 : "Annuaire international des Professionnels"**

« L'Annuaire International des Professionnels, outil puissant et rapide de recherche, permet aux entreprises du monde entier de référencer et présenter leurs offres, aux visiteurs de » l'annuaire.

**04/11/2003 : "Yahoo ferme sa division entreprise"**

« Yahoo a décidé de fermer sa division solutions d'entreprise et de stopper la commercialisation de ses différents logiciels dont les offres de portails d'entreprise My Yahoo Enterprise Edition et Yahoo Portal Solutions, et l'offre de diffusion de vidéo Yahoo Broadcast Solutions. (...) »

**03/11/2003 : "Interview de Christophe PARCOT, DG de Overture France"**

Le Directeur Général d'Overture France , Christophe PARCOT, nous livre tous ses secrets, ou presque, en 7 questions sur MoteurZine !

**01/11/2003 : "Un moteur de recherche pour les femmes"**

Ouverture d'un nouvel outil de recherche dédié aux femmes et à leur passion, le shopping. Cet outil se donne comme objectif d'être juste un guide au niveau de la recherche en proposant aux internautes une sélection de choix.

**01/11/2003 : "Google.ca à la hausse"**

« Le mois de septembre a vu la grimpée de Google.ca, qui a été consulté par plus de 2 millions d'internautes, dans le Top 25 du Web.

Ainsi, Google.ca se rapproche de MSN.ca, seul site canadien qui ces derniers mois attirait plus de deux millions de visiteurs uniques. Et il dépasse Sympatico.ca dans le palmarès des sites Web les plus consultés. »

**31/10/2003 : "Liens sponsorisés: Yahoo capitalise sur ses acquisitions"**

« Overture, racheté au début du mois d'octobre, par Yahoo, vient remplacer Espotting sur les portails européens du géant américain. Reste à franchir un pas autrement plus délicat: quitter Google pour utiliser sa propre technologie de recherche. »

**31/10/2003 : "20.000 annonceurs et 1 milliard de requêtes"**

« Le nombre d'annonceurs pour Q3 2003 est maintenant de 20.000 annonceurs actifs, contre 16.000 annonceurs le trimestre précédent.

Le nombre de requêtes enregistrées au mois de septembre 2003 est désormais de 1 milliard contre 850 millions le mois précédent. Ce volume représente le nombre d'opportunités pour nos 20.000 annonceurs de toucher des prospects potentiels. Ce chiffre exclue les requêtes fournies par Yahoo !

Depuis son lancement, Espotting Media a toujours suivie une stratégie de non dépendance de gros partenaires de distribution. Nous avons un réseau de distribution très large et très varié, mêlant habilement des gros sites portails et des petits sites de niches. L'objectif étant d'apporter à nos annonceurs de la volumétrie mais aussi du trafic qualifié par secteur d'activité. »

**31/10/2003 : "Overture affiche désormais ses résultats sur Yahoo! en Europe"**

« Overture affiche ses liens sponsorisés sur Yahoo! en France, au Royaume-Uni, en Allemagne, en Espagne et en Italie, consolidant ainsi sa position de leader en Europe. (...) »

Overture affichera pour une majorité de mots clés jusqu'à quatre liens sponsorisés en haut de page dans une section intitulée « Liens sponsorisés » et jusqu'à deux liens sponsorisés en bas de page. Pour tous les autres mots clés, deux liens sponsorisés Overture apparaîtront en haut de la page et trois en bas de page. »

**31/10/2003 : "Restructuration annoncée chez MSN"**

« L'information a été révélée par le Wall Street Journal : MSN a débauché le directeur technique d'Overture, Paul Ryan, pour prendre la tête de ses projets en matière de recherche Web et de référencement payant. Un recrutement symbolique qui tombe quelques semaines après que MSN ait renouvelé jusqu'en juin 2005 son contrat avec Overture pour la fourniture de liens sponsorisés aux Etats-Unis et au Royaume-Uni. »

**31/10/2003 : "Microsoft discuterait d'un partenariat avec Google"**

« Microsoft, qui cherche à prendre position sur le marché des moteurs de recherche sur internet, a discuté ces deux derniers mois avec Google d'un éventuel partenariat et même d'une fusion, rapporte le New York Times. »

---

## LES ARCHIVES

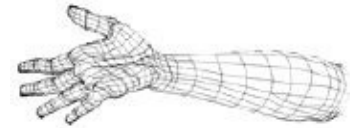
ou il y a un an, l'actualité du moment

### **01/11/2002 : "Est-ce que le pouvoir des outils de recherche menace l'indépendance du web ?"**

Patrick Ahern, Président de Data Recovery Group, a été le témoin de la toute puissance de Google et des difficultés de faire des affaires sans celui-ci. Pendant une année, Data Recovery Group a été le 4ème sur Google avec certaines requêtes "populaires". En janvier, quand Google a enlevé Data Recovery Group de ses pages de résultats sans aucune raison, la société a subi une baisse de CA de 30%.



Les outils de recherche  
francophones



#### **09 novembre 2003**

##### GKannu

Annuaire généraliste et répertoire d'entreprises.

#### **07 novembre 2003**

##### Sexe avenue

Prenez l'avenue qui vous mènent droit aux meilleurs sites de sexe gratuit.

##### Malango

Annuaire des sites des îles de l'Océan Indien.

#### **06 novembre 2003**

##### Guide du référencement

Un annuaire des référenceurs. Plus besoin de rechercher sur le net quel est le référenceur de sa région, en consultant cet annuaire, on peut trouver facilement tous ceux qui se situent près de chez soit.

#### **05 novembre 2003**

##### CD Musique

Le premier annuaire/webzine musical pop proposant des mp3 libres de droit en écoute.

#### **03 novembre 2003**

##### Web-Marseille

L'annuaire de tous les sites du Web marseillais.

##### Undernaute

Moteur de recherche underground européen.

#### **02 novembre 2003**

##### XannuaireX

Annuaire des site gratuits pour adultes.

Annuaire electromenager

Un annuaire pour faire le bon choix et ne pas se tromper dans son achat electromenager!

**01 novembre 2003**

Frenchbar

Si vous voulez tout savoir sur les cocktails. Fiches de + 300 recettes de cocktails, plus de 2.000 pages sur le monde du Bar. Tous les produits et les techniques du barman.

**31 octobre 2003**

Kaaleo

Les meilleurs sites du net notés par les internautes.

Top-manga

Annuaire du monde manga soft et sexy.

Fêtes du Vin

Annuaire alsacien sur le Haut Rhin, Bas Rhin, Les Vosges, Grand Est. L'Alsace et ses caves et son terroir.

**30 octobre 2003**

France annuaire

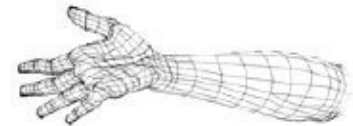
Annuaire de liens sexy.


Sport-access




Tous les sports 24/24 - Avec des milliers de documents concernant tous les sports.



Auto Promotion

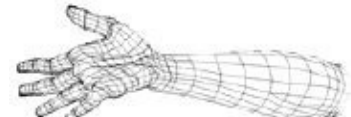


<p>Où trouver la meilleure info sur les outils de recherche ?</p>	
<p>par Caroline de <u>Crea-Interactive</u> (234x60)</p>	
	
<p>par Max, infodesigneur de MZ (170x85)</p>	

	
par Max, infodesigneur de MZ (61x60)	
	
par Alain de <u>Rêve Lémanique</u> (90x38)	
	
par Alain de <u>Rêve Lémanique</u> (90x38)	
par Alain de <u>Rêve Lémanique</u> (468x60)	



L'article / l'entretien



### **Le "topic sensitive pagerank"**

#### **Une piste efficace pour élaborer un moteur de recherche sensible au contexte.**

*Les annonces de concentration et d'alliances se multiplient en ce moment dans le petit monde des moteurs de recherche. Dans la bataille qui s'annonce, la lutte s'effectuera à coup de liasses de billets verts. Mais l'arme stratégique ultime pourrait s'avérer être une avancée technologique. Le " topic sensitive pagerank " (pagerank sensible à la thématique) constitue l'une des pistes les plus intéressantes que le moteur de recherche Google pourra exploiter à l'avenir. Ses inventeurs, les chercheurs Taher H. Haveliwala et Sepandar Kamvar de l'Université de Stanford, ont créé en juin 2003 la société Kaltix pour exploiter les applications de leurs recherches. Cette société a été rachetée par Google fin septembre 2003...*

#### **La pertinence du classement dans les pages de résultats : un enjeu majeur pour les moteurs de recherche.**

Google a montré lors de son lancement à quel point la notion de " pertinence " des résultats pouvait être stratégique. L'algorithme du pagerank a constitué son atout technologique

majeur, en permettant de classer les pages, non plus seulement en fonction de critères liés à leur seul contenu, mais en fonction du nombre de liens pointant vers ces pages. Page et Brin, les fondateurs de Google, sont partis d'une idée intuitive fort simple : si beaucoup de liens pointent vers une page, c'est donc que cette page est importante. Et si une page importante contient un lien qui pointe vers une page, c'est donc que cette page est probablement importante également.

Cette " intuition " a pu être mathématisée, ce qui a produit l'algorithme du Pagerank. Mais le pagerank, s'il a constitué un pas important dans l'amélioration de la pertinence des résultats, ne solutionne pas un problème majeur : les moteurs ne comprennent pas le contenu des pages... Et, pour reprendre un exemple tiré d'un article célèbre, si on tape " jaguar " dans un moteur de recherche, on trouvera aussi bien des pages tirées de sites animaliers que des pages issues de sites de concessionnaires automobiles...

C'est pourquoi, depuis 1998, de nombreux chercheurs ont essayé de développer des moteurs de recherche s'appuyant sur des outils d'évaluation sémantique. Mais ces moteurs sont restés pour l'instant des curiosités de laboratoire... Pourquoi ?

### **L'approche sémantique est restée longtemps bloquée par des limites technologiques**

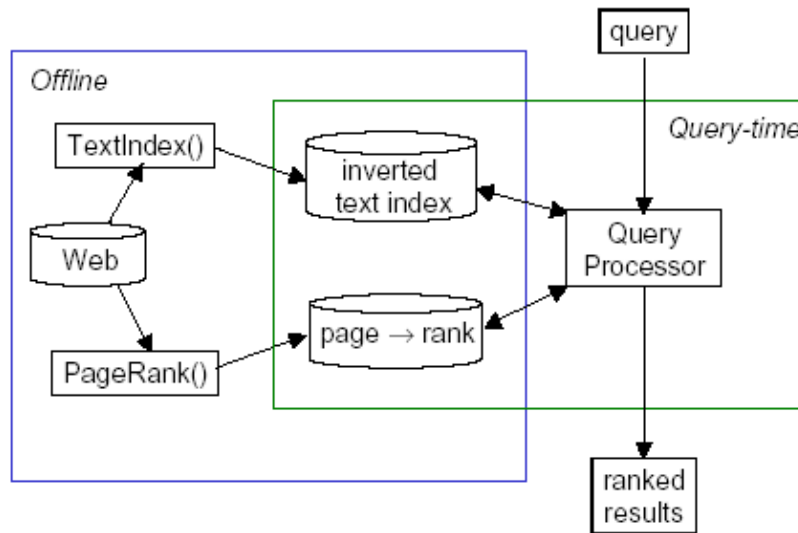
En fait, créer un moteur de recherche sophistiqué peut s'avérer simple sur le papier, mais beaucoup plus compliqué à transformer en moteur " grand public ".

En effet, gérer un index de la taille de la Toile mondiale est déjà en soi quelque chose de compliqué. La masse des données à manipuler est telle que l'on atteint les limites connues des techniques classiques d'interrogation de bases de données... Et pourtant, il est primordial, pour le succès d'un moteur, que le temps de réponse des serveurs soit le plus court possible.

Or, les approches construites autour d'une évaluation sémantique du contenu des pages ont classiquement un inconvénient majeur : elles nécessitent des calculs parfois complexes effectués à la volée... ou utilisent des méthodes qui nécessitent un travail considérable pour élaborer un moteur pertinent sur l'ensemble des requêtes, et non sur une ou deux thématiques choisies...

La méthode du pagerank a de ce point de vue une vertu : elle s'appuie uniquement sur la force brute de calcul. Ces calculs complexes sont effectués " offline ", les résultats (la " valeur " du pagerank de chaque page) sont stockés une bonne fois pour toutes dans l'index. Les calculs qui sont effectués à la volée pour classer les pages dans les résultats du moteur sont simples, rapides, et donc peu gourmands en ressources pour les serveurs...

C'est dans le domaine du savoir faire sur l'architecture de ses datacenters que Google conserve encore aujourd'hui une avance technologique sur ses concurrents. L'algorithme du pagerank est en effet aujourd'hui archi connu, et peut-être utilisé par n'importe quel concurrent. Mais construire une architecture complète capable de répondre à des millions et des millions de requêtes d'internautes chaque jour n'est pas à la portée de tout le monde.



Mais cela n'aura qu'un temps, et de nombreux indices montrent que cette technologie est sur le point de se révéler totalement dépassée. La puissance des machines augmente, les algorithmes se perfectionnent, et Google risque de se voir dépasser rapidement par des concurrents, s'ils ne montrent pas rapidement qu'ils sont capables d'innovations dans le domaine de l'amélioration de la pertinence des résultats.

### Une avancée intéressante : le Pagerank sensible à la thématique

Taher J. Haveliwala et Sepandar Kamvar ont proposé dès février 2002 un algorithme baptisé "topic sensitive Pagerank" (pagerank sensible à la thématique) et qui permet de construire à minima des moteurs de recherche thématiques efficaces.

Ce "nouvel" algorithme, en fait, ne consiste qu'en une réutilisation astucieuse du Pagerank traditionnel. L'idée part d'un constat simple : le pagerank permet de classer par ordre d'importance des pages dont le contenu est identique. Mais il ne permet pas de distinguer une page qui parle du "jaguar" l'animal, ou de la "jaguar" la voiture.

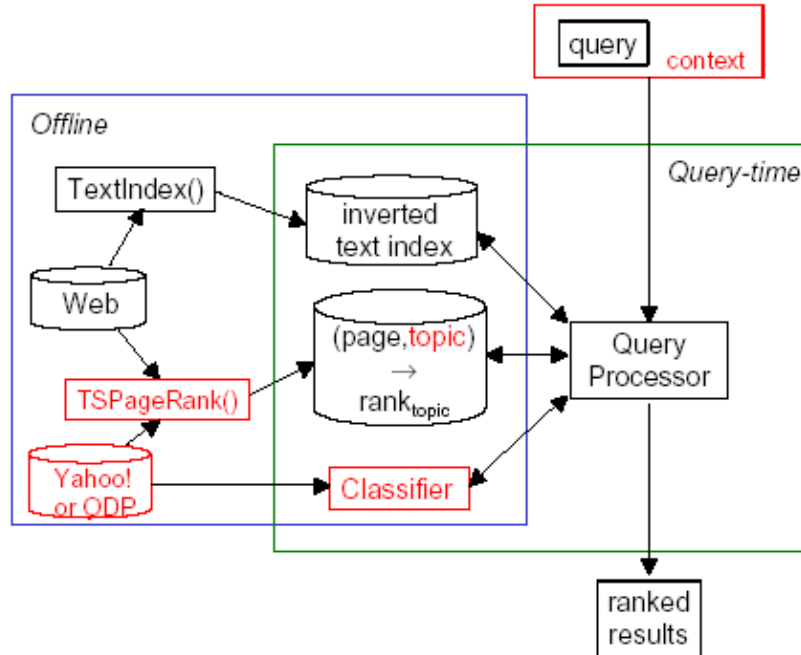
La solution proposée dans le "Topic Sensitive Pagerank" est de biaiser le calcul du Pagerank en donnant plus d'importance aux pages bénéficiant de liens en provenance de sites dont on connaît la thématique de départ. Dans la pratique, on peut par exemple prendre en compte les sites mentionnés dans l'une des 16 rubriques de niveau 1 de l'ODP.

Le résultat donne forcément 17 vecteurs de Pagerank différents pour chaque page : le Pagerank "général", et les 16 pageranks biaisés...

Ces valeurs sont stockées dans l'index avec les pages, comme pour un moteur basé sur le pagerank traditionnel.

Pour utiliser ces valeurs correctement, il suffit ensuite d'être en mesure de déterminer à quelle thématique se rattache une requête donnée. Ceci peut être obtenu grâce à une analyse syntaxique et sémantique sommaire, facile à calculer à la volée. Il suffit en effet de créer un index contenant, pour chaque thématique, le nombre d'occurrences de chaque terme apparaissant dans les pages. Ensuite, la thématique associée à une requête sera évaluée selon une méthode probabiliste, d'autant plus précise que l'on connaît le contexte de la requête...

Comme on le voit dans le schéma ci-dessous, l'architecture d'un tel moteur est assez peu différente d'une architecture traditionnelle avec pagerank. L'essentiel des calculs longs et complexes est toujours effectué offline.



Comme on le voit dans l'exemple ci-dessous, publié dans l'article de Kamvar et Haveliwala, les résultats obtenus sur la requête "bicycle" sont très différents suivant le contexte thématique, et les résultats en général beaucoup plus pertinents.



NOBIAS	ARTS
"RailRiders Adventure Clothing" <a href="http://www.RailRiders.com">www.RailRiders.com</a> <a href="http://www.Waypoint.org/default.html">www.Waypoint.org/default.html</a> <a href="http://www.Gorp.com/">www.Gorp.com/</a> <a href="http://www.FloridaCycling.com/">www.FloridaCycling.com/</a>	"Photo Contest & Gallery (Bicycling)" <a href="http://www.bikescape.com/photogallery/">www.bikescape.com/photogallery/</a> <a href="http://www.trygve.com/">www.trygve.com/</a> <a href="http://www.greenway.org/">www.greenway.org/</a> <a href="http://www.jsc.nasa.gov/Bios/htmlbios/young.html">www.jsc.nasa.gov/Bios/htmlbios/young.html</a>
BUSINESS	COMPUTERS
"Recumbent Bikes and Kit Aircraft" <a href="http://www.rans.com">www.rans.com</a> <a href="http://www.BreakawayBooks.com">www.BreakawayBooks.com</a> <a href="http://java.oreilly.com/bite-size/">java.oreilly.com/bite-size/</a> <a href="http://www.carbboom.com">www.carbboom.com</a>	"GPS Pilot" <a href="http://www.gpspilot.com">www.gpspilot.com</a> <a href="http://www.wireless.gr/wireless-links.htm">www.wireless.gr/wireless-links.htm</a> <a href="http://www.linkstosales.com">www.linkstosales.com</a> <a href="http://www.LiftExperts.com/lifts.html">www.LiftExperts.com/lifts.html</a>
GAMES	KIDS AND TEENS
"Definition Through Hobbies" <a href="http://www.flickr.com/gretchen/hobbies.html">www.flickr.com/gretchen/hobbies.html</a> <a href="http://www.BellaOnline.com/sports/">www.BellaOnline.com/sports/</a> <a href="http://www.npr.org/programs/wesun/puzzle/will.html">www.npr.org/programs/wesun/puzzle/will.html</a> <a href="http://www.trygve.com/">www.trygve.com/</a>	"Camp Shohola For Boys" <a href="http://www.shohola.com">www.shohola.com</a> <a href="http://www.EarthForce.org">www.EarthForce.org</a> <a href="http://www.WeissmanTours.com">www.WeissmanTours.com</a> <a href="http://www.GrownupCamps.com/homepage.html">www.GrownupCamps.com/homepage.html</a>
RECREATION	SCIENCE
"Adventure travel" <a href="http://www.gorp.com/">www.gorp.com/</a> <a href="http://www.GrownupCamps.com/homepage.html">www.GrownupCamps.com/homepage.html</a> <a href="http://www.gorp.com/gorp/activity/main.htm">www.gorp.com/gorp/activity/main.htm</a> <a href="http://www.outdoor-pursuits.org/">www.outdoor-pursuits.org/</a>	"Coast to Coast by Recumbent Bicycle" <a href="http://hypertextbook.com/bent/">hypertextbook.com/bent/</a> <a href="http://www.SiestaSoftware.com/">www.SiestaSoftware.com/</a> <a href="http://www.BenWiens.com/benwiens.html">www.BenWiens.com/benwiens.html</a> <a href="http://www.SusanJeffers.com/jeffbio.htm">www.SusanJeffers.com/jeffbio.htm</a>
SHOPPING	SPORTS
"Cycling Clothing & Accessories for Women" <a href="http://www.TeamEstrogen.com/">www.TeamEstrogen.com/</a> <a href="http://www.ShopOutdoors.com/">www.ShopOutdoors.com/</a> <a href="http://www.jub.com.au/books/">www.jub.com.au/books/</a> <a href="http://www.bike.com/">www.bike.com/</a>	"Swim, Bike, Run, & Multisport" <a href="http://www.multisports.com/">www.multisports.com/</a> <a href="http://www.BikeRacing.com/">www.BikeRacing.com/</a> <a href="http://www.CycleCanada.com/">www.CycleCanada.com/</a> <a href="http://www.bikescape.com/photogallery/">www.bikescape.com/photogallery/</a>

### Les difficultés techniques restantes sont en passe d'être résolues

Cette méthode a un inconvénient : elle multiplie le temps de calcul nécessaire du Pagerank par 17 dans l'hypothèse évoquée ici ! Mais depuis, toujours grâce aux travaux des chercheurs de Stanford, des progrès considérables ont été effectués dans les méthodes de calcul du pagerank.

L'algorithme notamment a été perfectionné depuis quelque mois. " L'extrapolation quadratique " permet d'éviter des calculs matriciels systématiques inutiles au cours du calcul du PageRank. Kamvar et Haveliwala ont également mis au point le "PageRank adaptatif" ("adaptive pagerank") un algorithme qui tient compte de la possibilité de ne pas calculer certains termes à différentes étapes de l'itération

Mais l'une des innovations majeures de Taher J. Haveliwala est certainement le " blockrank ", un algorithme permettant de simplifier considérablement le calcul du pagerank en tenant compte de la structure en blocs du web.

L'idée consiste ici à calculer un pagerank "local", sur un bloc de pages (dans la pratique limité à un domaine). L'algorithme converge rapidement dans ce cas particulier. Ensuite, pour le calcul du pagerank définitif, il suffit de faire un calcul sur une matrice de blocs, et non la Toile en entier, pour obtenir une excellente approximation du pagerank réel.

La manière dont les pageranks sont propagés de bloc à bloc dépend toutefois de la fixation de coefficients empiriques. Il est heureusement facile de déterminer ces coefficients de manière précise dès lors que l'on dispose d'un index de référence, avec des pageranks calculés avec un algorithme traditionnel.

Cet algorithme représente une avancée considérable : il diminue le temps de calcul des pageranks de manière considérable (jusqu'à 20 fois moins ?). Il permet aussi de déterminer une bonne approximation du PR effectif pour toute nouvelle page entrant dans l'index.

On soupçonne d'ailleurs fortement Google d'utiliser ce nouvel algorithme depuis mai/juin 2003.

Il est possible de créer aussi, au passage, des vecteurs multiples de pagerank pour chaque bloc... En combinant l'approche du pagerank biaisée et le blockrank, on peut obtenir, avec une économie de moyens remarquable, un moteur de recherche sensible au contexte opérationnel.

### **Cette technique sera t'elle exploitée par Google ?**

Rien ne s'oppose donc à la création d'un moteur de recherche sensible à la thématique de recherche utilisant cette technologie. Et encore moins à la création d'un moteur de recherche thématique, beaucoup plus facile encore à réaliser.

Les chercheurs regroupés au sein de la société Kaltix n'avaient pas encore, en juin 2003, résolu tous les problèmes liés à la mise en place d'une telle solution à base de " block rank ". Mais les écueils qui restent à surmonter sont tellement faibles qu'il ne faudra pas s'étonner de voir apparaître prochainement chez Google des fonctionnalités des pages de résultats sensibles au contexte...

### **Vers l'émergence de technologies concurrente ?**

Le " topic sensitive pagerank " n'est qu'un nouvel avatar de l'algorithme du Pagerank. Il ne permet pas de créer des moteurs réellement intelligents. Il s'agit toujours essentiellement de systèmes utilisant la force brute de calcul combinés avec un zeste d'astuce et de savoir faire.

Pour Google, cet algorithme a l'intérêt d'être facilement exploitable sans remises en cause majeures. Mais l'autre voie, lancée par IBM, vient d'aboutir à un concept d'un genre nouveau : le système WebFountain. Cette technologie, qui fait appel à des techniques de reconnaissance sémantique sophistiquées combinées à des algorithmes originaux (héritiers de HITS et du projet CLEVER), montre que la voie est ouverte pour des moteurs beaucoup plus efficaces et intelligents qu'un Google...

Mais ceci est une autre histoire...

### **Bibliographie :**

*Sur le pagerank :*

[Le PageRank par l'exemple](#), par Dan Hetzel

[L'algorithme du PageRank expliqué](#), par Dan Hetzel

*Principaux articles publiés par les créateurs de la société Kaltix :*

[Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search](#) –

Taher H. Haveliwala / 10 février 2002 revu le 17 mai 2003

Extrapolation Methods for Accelerating PageRank Computations – Sepandar D. Kamvar, Taher H. Haveliwala, Christopher D. Manning, Gene H. Golub / 28 Février 2003

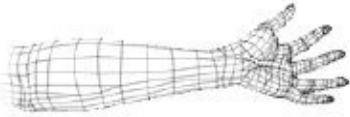
Exploiting the Block Structure of the Web for Computing PageRank – Sepandar D. Kamvar, Taher H. Haveliwala, Christopher D. Manning, Gene H. Golub / 4 Mars 2003

Adaptive Methods for the Computation of PageRank – Kamvar, Sepandar; Haveliwala, Taher; Golub, Gene / 28 avril 2003

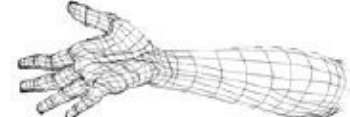
An Analytical Comparison of Approaches to Personalizing PageRank – Taher Haveliwala, Sepandar Kamvar and Glen Jeh / 20 juin 2003

Computing PageRank using Power Extrapolation – Haveliwala, Taher; Kamvar, Sepandar; Klein, Dan; Manning, Chris; Golub, Gene / 16 juillet 2003

par Philippe YONNET de **Webmaster–Hub**



Conclusion



### ATTENTION

La liste des abonnés n'est pas disponible. Nous ne la donnons ni ne la revendons à personne. En revanche, vous pouvez sponsoriser notre lettre d'information : contactez-nous pour discuter des modalités.

Abonnement  
gratuit à  
« Moteurzine »

Désinscription  
de  
« Moteurzine »

les crédits

Chronique sur « Google »

Olivier Duffez de  
**WebRankInfo**

Chronique sur les moteurs

Gilbert Wayenborgh de  
**DeepIndex**

Chronique sur les entreprises

Jean–Claude Benard de  
**Quadramultimédia**

Chronique sur la visibilité

Damien Guigue de  
**Yooda**

Chronique sur le référencement  
marketing

Gautier Girard de **Cerise**  
**Rouge**

---

Humour et Internet

Frédéric Lepage de **Les  
BD qui bougent**

---

Chronique sur le référencement

Marie Pourreyron  
d'**Altitude**  
**Référencement**

---

Chronique sur l'aspect commercial

Pierre Frigon de  
l'**Agence Hyperclics**  
**Marketing**

---

Infodesign

Max, le talentueux !

---

Le reste (mise en page, édito,  
actualités, les nouveautés et ...)

CHRis Hédé de  
**MoteurZine / Enfin**

© 1999 à 2003 par IDF.net