



LA GESTION DE TRAFIC DES SITES EN LIGNE

Thème : Les moteurs de recherche et annuaires.

1/ Les moteurs de recherche et annuaires

Le terme outils de recherche englobe les annuaires et les moteurs de recherche disponibles sur le web. La différence majeure entre l'annuaire et le moteur de recherche est la présence de l'homme pour classer les sites dans le premier cas et l'utilisation d'un algorithme dans le second cas. Cette différence est parfois difficile à observer du fait qu'un site puisse coupler la présence d'un annuaire et d'un moteur de recherche sur une unique page web. Par exemple, le portail Wanadoo est un outil de recherche qui possède un annuaire et un moteur¹ de recherche.

Sachant que les Internautes préfèrent utiliser le moteur de recherche de leur choix ; une société doit choisir les outils de recherche sur lesquels son site va être référencé. Habituellement pour qu'une campagne de communication soit rentable, la société effectue un ciblage des personnes susceptibles d'être intéressées par le service ou le produit vendu. Elle utilise les supports de communication en adéquation avec sa cible. Pour le référencement sur les moteurs de recherche, la méthode n'est pas la même. L'une des particularités des outils de recherche est qu'ils touchent indifféremment toute catégorie socioprofessionnelle, quelque soit le sexe et l'âge de l'utilisateur. Comme les moteurs sont utilisés par tous, on ne peut pas cibler l'utilisateur en choisissant un moteur de recherche plutôt qu'un autre. C'est l'utilisateur du moteur de recherche qui ciblera sa recherche ; et **c'est à la société de choisir les bons termes de référencement pour être en adéquation avec sa cible.**

Pour être visible, un site en ligne doit être référencé sur les principaux outils de recherche.

1.1 Sur quels outils de recherches faut-il se référencer ?

La pertinence du référencement d'un site **dépend de la technologie utilisée** par le moteur de recherche. Pour une même requête, les critères de classement des sites diffèrent selon la technologie utilisée. Si le travail de référencement est correctement effectué pour l'une des technologies, la société doit référencer son site sur les outils de recherche utilisant cette technologie. C'est pourquoi **il est important de connaître les différentes technologies et de savoir quels sont les portails qui les utilisent.** Après avoir déterminé les technologies utilisées par les outils de recherche les plus utilisés, je répondrai à la question : **sur quels outils de recherche faut-il se référencer ?**

¹ Le portail Wanadoo utilise la technologie de Voila pour son moteur de recherche

1.1.1 Les fournisseurs de technologies de recherche :

Présentation de Google Inc² :

La société Google Inc. développe un moteur de recherche d'informations sur le Web. Le moteur nommé Google répond à plus de 200 millions de demandes de recherche par jour et son index est constitué de plus de Quatre milliards de pages Web. La technologie de recherche proposée par Google peut être exploitée à partir de PC de bureau et de multiples plates-formes sans fil telles que les téléphones WAP ou i-mode.

Technologie de Google :

La technologie de recherche développée par Google³ est totalement automatisée. Son utilisation ne nécessite aucune intervention humaine. Google intègre plusieurs procédés de recherche. La plus innovante est l'analyse de correspondance hyper textuelle développée par Larry Page et Sergey Brin et baptisée « *PageRank* ».

L'analyse de correspondance hyper textuelle examine la totalité du contenu de chaque page Web en tenant compte des attributs de la police utilisée et de la position des termes figurant sur la page. Le contenu des pages Web voisines est également analysé par ce procédé. L'ensemble de ces données permet de renvoyer des résultats correspondant aux requêtes des utilisateurs. La technologie du « *PageRank* » classe les pages web par ordre d'importance. L'ordre ainsi défini est le résultat d'un algorithme comprenant 500 millions de variables et 3 milliards de termes. Une autre spécificité du « *PageRank* » est sa faculté à classer les pages web en fonction du nombre d'affiliations à d'autres pages. La qualité des pages affiliées est également prise en compte dans le calcul du classement.

Présentation de Voila :

Le moteur nommé Voila est développé par la société Echo, filiale de Wanadoo. Echo Interactive est initialement connue pour son moteur de recherche Web. Mais cette filiale de Wanadoo a évolué dans la plupart des technologies du Net, et elle gère aujourd'hui un ensemble de savoir-faire nécessaires à la production de services en ligne.

Fiche de synthèse :

Nom : Google

Date de lancement : 1998

Propriétaire : Google.Inc

Taille de l'index : 4.3
Milliards de pages

Délai de rafraîchissement de l'index : 4 semaines (période appelée la « Google Dance »)

Technologie utilisée par les portails : Free, Club-Internet (France), AOL, Earthlink, Netscape.

Fiche de synthèse :

Nom : Voila

Date de lancement : Juillet
1998

Propriétaire : Wanadoo

Taille de l'index : 60
Millions de pages en Français

Délai de rafraîchissement de l'index : 4 semaines

Technologie utilisée par les portails : Wanadoo, Voila.fr.

² Une présentation générale de Google est consultable à l'adresse <http://www.google.fr/press/overview.html>

³ Une présentation générale de la technologie Google.Inc est consultable à l'adresse http://www.google.com/press/overview_tech.html

Présentation de Exalead⁴ :

Exalead Corporate fournit des « *outils d'accès à l'information centrés sur l'utilisateur* ». Les outils fournis par Exalead s'adressent particulièrement aux sociétés présentes sur le web et qui souhaitent mettre en place un moteur capable d'effectuer des recherches sur **leur** site. Le portail AOL.fr utilise la technologie Exalead comme système d'indexation.

Technologie Exalead :

Exalead est l'équivalent d'un moteur de recherche « *personnalisé* ». Le moteur Exalead est configuré pour effectuer ses recherches dans de nombreux formats de documents tels que Word, Adobe Acrobat, jpg, png, mp3, et Macromedia Flash...

Présentation de Inktomi :

Inktomi Corp. est spécialisée dans la conception, le développement et la commercialisation de logiciels destinés aux environnements Internet et Intranet. La gamme des produits Inktomi comprend un moteur de recherche sur le Web, une solution logicielle de shopping comparatif en ligne, un logiciel de cache mémoire destiné à améliorer les performances d'Internet et des réseaux et un moteur d'annuaire automatisé. Ses produits sont distribués au travers d'un réseau de revendeurs présents dans sept pays : France, Royaume-Uni, Allemagne, Autriche, Suisse, Suède et les pays du Benelux.

Technologie de Inktomi Search Engine :

Inktomi est un moteur de recherche indirect, c'est à dire qu'il ne propose pas lui-même une interface commerciale de recherche sur le web. Pour ses clients moteurs de recherche, les requêtes lui sont transmises par ses clients et Inktomi renvoie les résultats. Pour ses clients annuaires, Inktomi intervient lorsque l'annuaire n'a trouvé aucune réponse correspondant à la requête de l'utilisateur. Les annuaires effectuent une première recherche sur leur base d'information propre puis font appel à Inktomi pour approfondir la recherche.

Fiche de synthèse :

Nom : Exalead

Date de lancement : 2002

Propriétaire : Exalead

Taille de l'index : 50

Millions de pages en Français

Délai de rafraîchissement de l'index : 4 semaines

Technologie utilisée par les portails : Exalead, AOL.fr.

Fiche de synthèse :

Nom : Inktomi Search Engine

Date de lancement : 1996

Propriétaire : Inktomi

Taille de l'index :

Délai de rafraîchissement de l'index :

Technologie utilisée par les portails : Snap, HotBot, GoTo.com, Canada.com, NTT, @Home Network, MSN France, AOL France, et Nomade France, Yahoo.

⁴ Une présentation générale d'Exalead est consultable à l'adresse : <http://www.exalead.com/cgi/exalead>

Présentation de Fast⁵ :

Overture Services Inc. propose une gamme complète de produits et de services de recherche sur Internet tel que le moteur de recherche Fast / All The Web.

La société Overture est une filiale de Yahoo depuis 7 Octobre 2003. Elle est dirigée par Ted Meisel et son siège social est basé à Pasadena, en Californie.

Technologie de AllTheWeb :

AlltheWeb index plus de 3 Milliards de pages. La technologie utilisée permet de trouver différents formats de documents : PDF, Word et de nombreux types de formats multimédias. Cette technologie gère 36 langues différentes dans lesquelles l'utilisateur peut effectuer une recherche.

Présentation de Altavista⁶ :

Altavista est la propriété de Yahoo par le biais de sa filiale Overture Services Inc. Cette filiale, fournit des services de recherche et de technologie. Elle est installée à Palo Alto, en Californie et possède 61 brevets déposés et liés à la recherche.

Technologie de Altavista :

Depuis novembre 2002, une nouvelle version du moteur Altavista est disponible. La présentation, la technologie de classement des résultats et d'indexation ont changé. Une technologie nommée « Prisma » permet désormais d'affiner les résultats d'une recherche en proposant des termes connexes. L'algorithme de recherche a été amélioré et le délai de rafraîchissement des résultats a été raccourci.

Fiche de synthèse :

Nom : Fast / All The Web

Date de lancement : Mai 1999

Propriétaire : Yahoo

Taille de l'index : 3.2 Milliards de pages

Délai de rafraîchissement de l'index : 2-3 semaines

Technologie utilisée par les portails : Lycos, Tiscali, Yahoo.

Fiche de synthèse :

Nom : Altavista

Date de lancement : Décembre 1995

Propriétaire : Yahoo

Taille de l'index : 1.1 Milliards de pages

Délai de rafraîchissement de l'index : 4 semaines

Technologie utilisée par les portails : Altavista, Yahoo.

1.1.2 Le classement des principales technologies de recherche :

Les 6 technologies majeures et utilisées par les outils de recherches sont : Google, Voila, Exalead, Inktomi, Fast et Altavista. Le classement utilisé dans cette partie du document date de Mai 2003. Depuis ce classement, il est important de noter que la société Overture propriétaire de Fast / AllTheWeb est devenue une filiale de Yahoo le 7 Octobre 2003.

⁵ Une présentation générale de Fast/AllTheWeb est consultable à l'adresse :

<http://www.alltheweb.com/info/about/index>

⁶ Une présentation générale de Altavista est consultable à l'adresse : <http://fr.altavista.com/about/>

Tableau 1 : Classement des outils de recherche

Classement des moteurs de recherche et annuaires - internautes français		
outil de recherche	%	Technologie "moteur" utilisée
google	52,89 %	google
wanadoo	11,55 %	voila
yahoo	11,16 %	google
aol	8,12 %	exalead
msn	5,64 %	inktomi
voila	3,83 %	voila
lycos	2,16 %	fast
tiscali	1,53 %	fast
free	1,16 %	google
altavista	0,67 %	altavista
club - internet	0,59 %	fast
...		
fast / alltheweb	0,05 %	fast
hotbot	0,02 %	inktomi


Les outils de recherche **Google, Wanadoo, Yahoo et Aol** représentent plus de 80 % des parts de marché.

Source : Etude « Moteurs de recherches: 5 acteurs, 95 % des recherches ! » par e-stat et @position – Mai 2003

Le tableau ci-dessus met en évidence l'importance de Google en tant que moteur de recherche et en tant que portail. En Mai 2003, la technologie Google possède 61.85 % de parts de marché⁷ grâce aux portails Google, Yahoo et Free. Cette technologie est la plus présente sur le marché français. Vient ensuite la technologie Voila qui détient 15.37 % du marché avec les portails Wanadoo et Voila. Les autres technologies utilisées possèdent toutes individuellement moins de 10 % de parts de marché. C'est le cas par exemple d'Exalead, une technologie moteur exclusivement utilisée par Aol. C'est également le cas pour Inktomi, Fast et Altavista.

Tableau 2 : Classement des outils de recherche par technologie

Si la société adapte son site aux technologies moteur de Google, voila et Exalead et qu'elle le fait référencer sur tous les sites utilisant ces technologies, elle touche plus de 85 % des utilisateurs de moteurs de recherche !

 Depuis le rachat d'Overture services inc. par Yahoo, les technologies moteur Inktomi, Fast et Altavista ont vu leur part de marché croître.

Classement des technologies moteurs			
Technologie moteur	%	Portail utilisateur	%
google	61,85 % répartis en :	google	85,5%
		yahoo	12,6 %
		free	1,9 %
voila	15,37 % répartis en :	wanadoo	75,1 %
		voila	24,9 %
exalead	8,12 % répartis en :	aol	100 %
inktomi	5,66 % répartis en :	MSN	99,7 %
		hotbot	0,3%
fast	4,10 % répartis en :	Lycos	53,6 %
		tiscali	30,4 %
		club-internet	14,6 %
		fast / alltheweb	1,3 %
altavista	0,67 % répartis en :	altavista	100 %

Source : Etude « Moteurs de recherches: 5 acteurs, 95 % des recherches ! » par e-stat et @position – Mai 2003

Il est aisé de référencer un site sur tous les moteurs de recherche existant. Cependant, un site peut être référencé sur de nombreux outils de recherche sans que cela n'accroisse sa visibilité aux yeux des Internaute. Pour qu'un site soit correctement placé dans les moteurs de recherches, il est nécessaire de travailler sur l'infrastructure et sur le contenu du site. Je

⁷ Le marché étudié est Français

préconise d'adapter le site aux technologies de Google, voila et Exalead pour que celui-ci soit correctement placé lors d'une recherche effectuée à partir des pages Google, Yahoo, Free, Wanadoo, Voila et AOL.

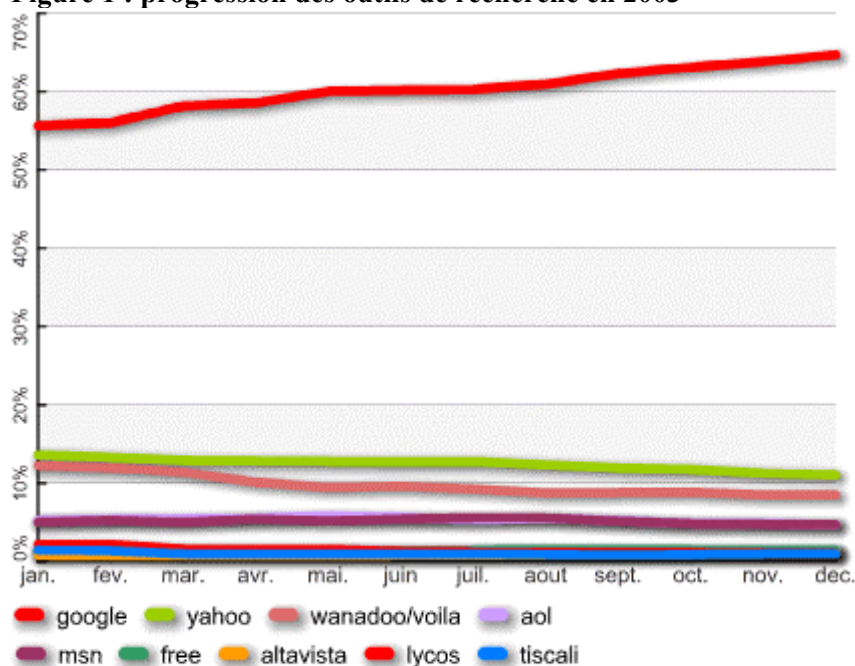
Depuis le début de l'année 2004 Yahoo n'utilise plus la technologie moteur Google. Yahoo en faisant acquisition⁸ des moteurs Inktomi, Altavista et Fast privilégie désormais ses propres outils de recherche. Ce changement doit être pris en considération dans les préconisations faites précédemment. Si la société adapte son site aux seules technologies Google, Voila et Exalead, celui-ci risque d'être mal placé sur le portail Yahoo qui utilise désormais d'autres technologies moteur.

1.2 Quel est l'avenir des outils de recherche ?

Le rachat d'Overture services inc. par Yahoo a provoqué un changement dans la répartition des parts de marché entre les différentes technologies moteur existantes. Selon la stratégie des acteurs propriétaires de technologies moteurs, il est probable que ce type de fusions ou acquisitions se renouvellent dans un futur proche. Lors de ces changements, les propriétaires de sites web doivent adapter leur stratégie de référencement au nouvel environnement. Dans cette partie de l'étude je réalise un prévisionnel qui doit me permettre de répondre à la question : **Quel est l'avenir des différents outils de recherche ?**

1.2.1 La progression d'ensemble :

Figure 1 : progression des outils de recherche en 2003



Source : Etude « L'utilisation des moteurs de recherche par les internautes en 2003 » par @position Janvier 2004

En 2003, Google accroît ses parts de marché au détriment de la concurrence. Vous pouvez observer sur le graphique ci-contre que Google progresse positivement tandis que Yahoo, Voilà, AOL, MSN, Lycos et Tiscali perdent des parts de marché. Parmi les « outsider », nous observons une évolution positive de Free et Altavista. Nous essaierons d'en comprendre le pourquoi dans la partie « progression des outsider ».

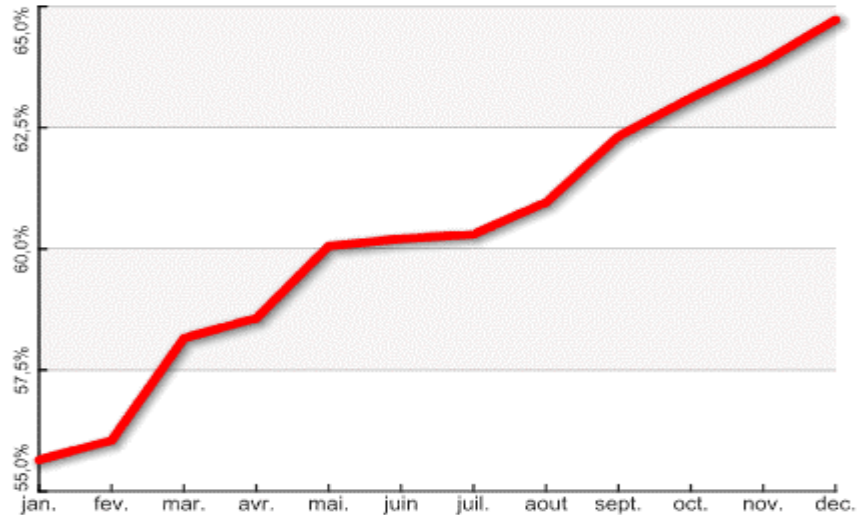
L'important à retenir est la progression de **Google qui continue sur sa lancée et affirme sa position de leader !**

⁸ Yahoo a acquis le moteur de recherche Inktomi en Décembre 2002 pour une valeur de 235 Millions de dollars

1.2.2 La progression de Google :

Google, le leader sur le marché des moteurs de recherche est l'un des seuls à poursuivre son ascension en 2003. Avec plus de la moitié des parts du marché français, **Google est le portail de recherche à privilégier lors du référencement d'un site !**

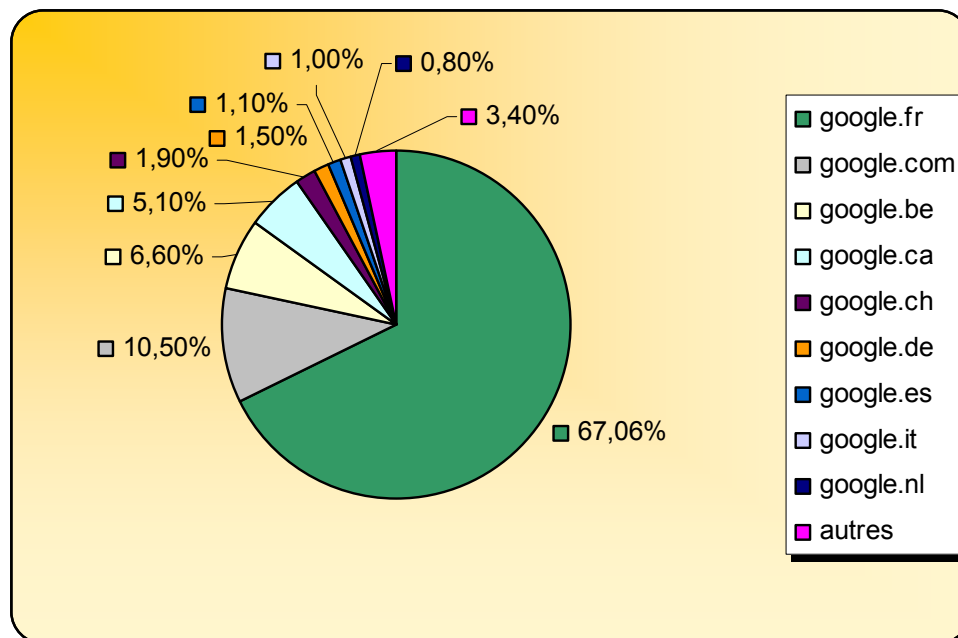
Figure 2 : Progression de Google en 2003



Source : Etude « Google en détail sur 2003 » par @position - Janvier 2004

L'utilisation de Google par les français se fait majoritairement par le biais de Google.fr (67.06%) même si l'extension de Google en .com est encore considérablement utilisée (10.50%). Il faut noter que le pourcentage d'utilisation de l'extension .fr est en nette progression de l'année 2003 à 2004.

Figure 3 : Parts de marché des principaux sites Google sur le territoire Français



Source : Etude « l'utilisation des outils de recherche : utilisation de Google en France » par Médiamétrie estat - @position - Février 2004

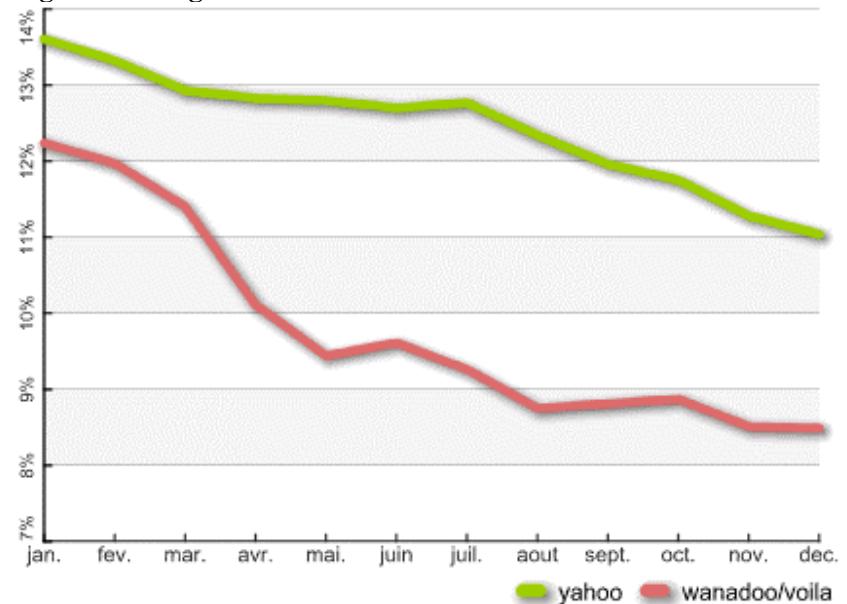
La connaissance de ces résultats doit entrer en compte dans la stratégie de référencement d'un site. En effet, selon l'extension de Google, le même site n'apparaît pas à la même position sur les pages de résultats d'une recherche. **Dans le cas d'un site uniquement adressé aux**

français, il est plus important que le site soit bien placé sur Google.fr que sur Google.com !

1.2.3 La progression de Yahoo et Voila :

Avec plus de 20 % des parts de marché au niveau mondial, Yahoo est le concurrent le plus proche de Google. Cependant sur le marché français de fin d'année 2003, Yahoo possède à peine 11 % des parts de marché. Malgré une courbe de progression dégressive sur l'année 2003, il est possible que Yahoo regagne des parts si la stratégie adoptée par le groupe porte ses fruits : Yahoo utilise depuis le début de l'année 2004, ses propres moteurs de recherches.

Figure 4 : Progression de Yahoo et Voila en 2003

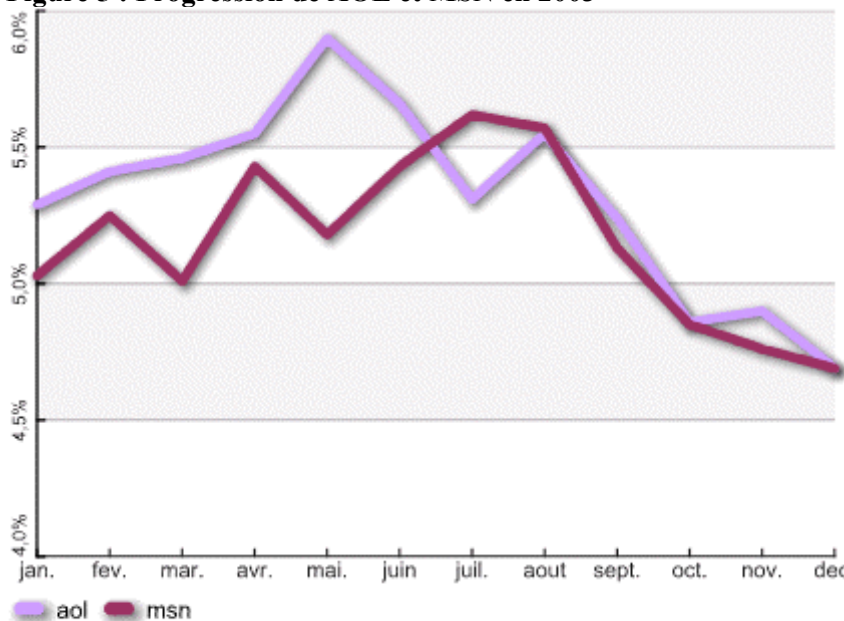


Source : Etude « Yahoo et Wanadoo en détail sur 2003 » par @position - Janvier 2004

De son côté, Wanadoo continue une chute entamée depuis 2002. Il faut rappeler que cette société utilise la technologie de recherche Voila et que cette technologie n'est quasiment pas utilisée au niveau mondial ! Fin 2003 Wanadoo possède moins de 9 % des parts du marché français.

1.2.4 La progression de AOL et MSN :

Figure 5 : Progression de AOL et MSN en 2003



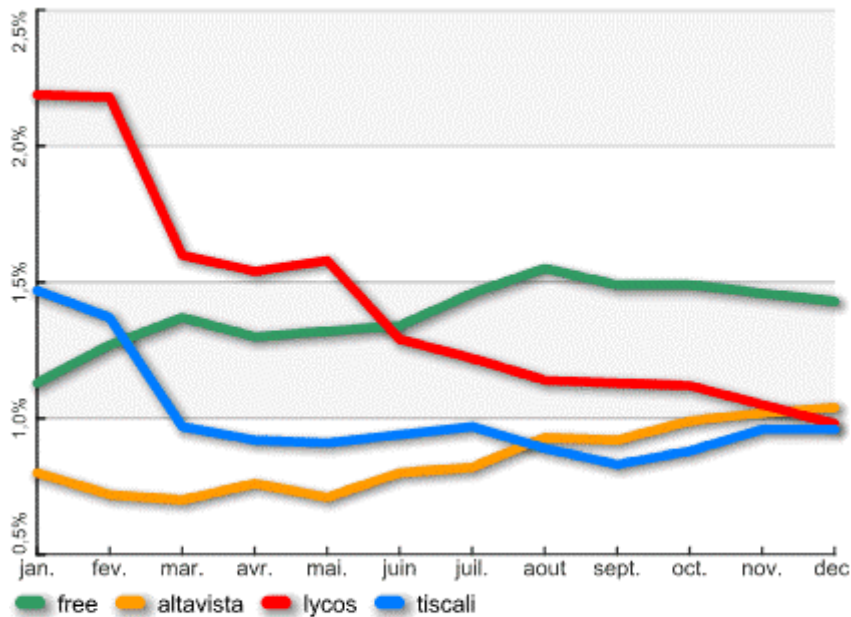
En 2003, l'utilisation des moteurs de recherche via les portails AOL et MSN est en baisse. En effet, fin 2004 les parts de marché de ces outils de recherche sont inférieures à 5 %.

Du côté de Microsoft, cette chute n'est pas en adéquation avec les perspectives d'avenir de la société. En effet, Microsoft souhaite accentuer le développement de la recherche sur Internet.

Source : Etude « AOL et MSN en détail sur 2003 » par @position - Janvier 2004

1.2.5 La progression des « outsider » :

Figure 6 : Progression des "outsider" en 2003



En 2003, deux « outsider » augmentent leur parts de marché : Free et Altavista. Free est une filiale du groupe Iliad. C'est également la marque leader de la société mère. En 2004, Free est le second fournisseur d'accès Internet sur le territoire Français. Son portail de recherche : www.free.fr utilise la technologie de recherche Google.

Source : Etude « free, altavista, lycos et tiscali en détail sur 2003 » par @position - Janvier 2004

Depuis Février 2002, Altavista ne faisait plus partie des 10 portails les plus utilisés sur le marché Français. Il a réintégré ce placement depuis le début de l'année 2003. Cette remontée peut s'expliquer par le rachat de la technologie moteur par Yahoo et par la mise en place d'une version nouvelle du moteur Altavista depuis novembre 2002.

Webographie

Le baromètre ad'oc :

Ad'oc est une agence dédiée au référencement et à la gestion de trafic depuis 1998. Cette agence publie mensuellement un baromètre "référencement et trafic" basé sur un large panel d'outils de recherche. Cette étude restitue l'expérience d'un panel de sites web d'entreprise, exclusivement français et faisant l'objet d'un référencement optimisé. Le baromètre ad'oc est un indicateur d'évaluation qui permet d'apprécier la performance d'un recrutement d'internautes en provenance des outils de recherche.

URL : <http://www.barometre.adoc.fr/>

La revue du référencement :

La revue du référencement est la propriété de la société anonyme @position créée en 1999. Cette société développe des prestations de « référencement et positionnement », « d'analyse et stratégie de référencement », de « conseil opérationnel », de « publicité et référencement événementiel », « d'audit sémantique, accessibilité, visibilité », « d'expertise métiers péri-référencement », « d'étude et mesure de visibilité ».

URL : <http://www.revue-referencement.com/>

Etudes :

- « Moteur de recherche ou Annuaire ? » menée par e-stat et @position - 2002
URL : <http://www.estat.com/content/fr/presse/20020410/moteurs.pdf>
- « Moteurs de recherches: 5 acteurs, 95 % des recherches ! » menée par e-stat et @position - Mai 2003
URL : http://www.revue-referencement.com/ZOOM/zoom_techno.htm
- « Panorama du web Français »⁹ par e-stat et @position - Mars 2004
URL : http://www.revue-referencement.com/chiffrescles/panorama_mars04.htm
- « L'utilisation des outils de recherche : Utilisation de Google en France » menée par Médiamétrie / estat & @position - Février 2004
URL : http://www.revue-referencement.com/ETUDES/04_etude_google.htm
- « Yahoo et Wanadoo en détail sur 2003 » menée par @position - Janvier 2004
URL : http://www.revue-referencement.com/ZOOM/yahoo_wanadoo_annee_2003.htm
- « Aol et Msn en détail sur 2003 » menée par @position - Janvier 2004
URL : http://www.revue-referencement.com/ZOOM/aol_msn_annee_2003.htm
- « free, altavista, lycos et tiscali en détail sur 2003 » menée par @position - Janvier 2004
URL : http://www.revue-referencement.com/ZOOM/free_altavista_lycos_tiscali_annee_2003.htm
- « En 2003, le référencement devient la première source de trafic » menée par @position - Janvier 2004
URL : http://www.revue-referencement.com/ZOOM/types_acces_annee_2003.htm

⁹ Ce classement est basé sur 57 052 sites et 45 746 365 visites datées du 01-03-04 au 31-03-04, sans distinction du pays d'origine des visites.

UDA¹⁰ :

L'Union Des Annonceurs est une association Française qui a été fondée en 1916. Cette organisation est l'unique représentant des annonceurs¹¹ en France. Elle regroupe environ 400 adhérent en 2004.

URL : <http://www.uda.fr/>

Etudes :

- « Les chiffres clés des annonceurs » par UDA – 2002
URL : http://www.uda.fr/ressources/pdf/chiffres_cles2003/2_UDAchiffres03.pdf
- « Mix média par taille d'annonceur / Nombre d'annonceurs par média » par UDA – 2002
URL : http://www.uda.fr/ressources/pdf/chiffres_cles2003/4_UDAchiffres03.pdf
- « Evolution de l'Internet en France » par UDA – 2002
URL : http://www.uda.fr/ressources/pdf/chiffres_cles2003/5_UDAchiffres03.pdf

Le Journal du Net :

Le journal du net est un site web propriété de Benchmark Group. Ce site diffuse l'actualité du secteur Internet dans le domaine des entreprises, de la réglementation, de la finance et des technologies.

URL : <http://www.journaldunet.com/>

Dossiers :

- « Dossier : Google » coordonné par Florence Santrot - Mai 2004
URL : <http://www.journaldunet.com/dossiers/google/>

ZDNet.fr :

ZDNet.fr est un site d'information sur l'informatique et les nouvelles technologies. Il est édité par CNET Networks France, la filiale française du groupe américain CNET Networks Inc. Le site ZDNet.fr propose des articles d'actualité, des livres blancs, des études de cas et des services destinés à faciliter les décisions en matière d'achat et d'utilisation de produits high-tech.

URL : <http://www.zdnet.fr/>

Actualités :

- « Publicité en ligne : forte croissance en France et aux Etats-Unis » par Christophe Guillemain - Mai 2004

URL : <http://www.zdnet.fr/actualites/internet/0,39020774,39154469,00.htm>

¹⁰ UDA : Union Des Annonceurs

¹¹ Entreprises, collectivités et tout organisme susceptible d'utiliser des supports de communication