

La guerre de positionnement des moteurs de recherche

Mots clefs / Keywords

Google, Yahoo, MSN, moteurs, recherche, outils, services, Internet, positionnement, index, pertinence, algorithme, crawl, engine, search, classement, stratégies, groups, images, vidéos, SWOT, fiche technique, toolbar, desktop, Groupwares, blog, diversification, développement, communication, innovation

Présenté par : Julien Lombard-Donnet

Structure de l'étude

INTRODUCTION (page 4)

PREMIERE PARTIE (page 5)

Pour chaque moteur de recherche, cette partie comprendra :

- une fiche technique ainsi que des sites clefs permettant d'approfondir l'information
- une fiche produit avec un bref commentaire sur les produits et diversifications des sociétés
- une fiche analyse (SWOT), brève et succincte permettant de mieux cerner les enjeux, ainsi qu'une Webographie indicative pour aller plus loin dans la connaissance et la veille du moteur concerné.

DEUXIEME PARTIE (page 15)

1 – Historique et présentation des moteurs de recherche

1.1 – Google

1.1.1 – Présentation

1.1.2 – Historique

1.2 – MSN

1.2.1 – Présentation

1.2.2 – Historique

1.3 – Yahoo !

1.3.1 – Présentation

1.3.2 – Historique

2 – Caractéristiques comparées de ces moteurs de recherche

2.1 – Le nombre de pages crawlées

2.2 – Evaluation de la pertinence

2.2.1 – Aspects techniques du classement des sites

2.2.2 – Comparatif de la pertinence des résultats

TROISIEME PARTIE (page 25)

1 – La guerre de diversification

1.1 – La bataille des services

1.1.1 – La recherche web

1.1.1.1 – La recherche image

1.1.1.2 – La recherche vidéo

1.1.1.3 – La recherche « locale »

- 1.1.2 – L'enjeu des RSS
- 1.1.3 – Concurrence sur d'autres types de services
- 1.1.4 – Les services innovants
- 1.2 – La bataille des outils
 - 1.2.1 – Les Desktop Search
 - 1.2.2 – Providers de Blog Spaces
 - 1.2.3 – Guerre des Bar(res)
 - 1.2.4 – Emails et groupwares : guerre de la communication P2P
 - 1.2.5 – Autres innovations

- 2 – Tableau comparatif
 - 2.1 – Méthodologie et objectifs
 - 2.2 – Tableaux comparatifs
 - 2.2.1 – Les services
 - 2.2.2 – Les outils
 - 2.3 – Analyse

QUATRIEME PARTIE (page 34)

- 1 – Les stratégies de conquête
 - 1.1 – Bilan des stratégies de développement et de diversification
 - 1.1.1 – Stratégies de diversification
 - 1.1.2 – Stratégies de développement
 - 1.2 – La communication : outil prépondérant sur le Web
- 2 – Forces et faiblesses par rapport au marché
 - 2.1 – Enjeux et principaux concurrents
 - 2.1.1 – La gestion de recherche
 - 2.1.2 – Cartographie et Clusterisation
- 3 – Perspectives et prospectives

CONCLUSION (page 42)

Tableau synthétique et conclusion générale

WEBOGRAPHIE (page 44)

Introduction

Les moteurs de recherche sur le web fleurissent de plus en plus. Parmi eux, trois « géants » se distinguent clairement : Google, MSN et Yahoo !. Ceux-ci se livrent une lutte sans merci pour devenir le seul et l'unique moteur de recherche privilégié par les internautes.

Les trois géants de l'Internet ont, chacun à leur manière, avancé leurs pions pour mieux conquérir en 2004 le marché en pleine expansion des liens sponsorisés. Google a racheté à tout va, Yahoo a frappé un grand coup et MSN a brouillé les cartes. Un pactole estimé à 1,4 milliard de dollars rien qu'en 2003: le montant des revenus générés par les liens sponsorisés explique aisément pourquoi Google, Yahoo et bientôt MSN se livrent à une concurrence acharnée dans l'exploitation économique de la recherche d'informations sur le web. Le marché des liens promotionnels ou sponsorisés a réussi à redonner, presque à lui seul, de la vigueur au secteur publicitaire, en plein marasme.

Quel est le meilleur moteur de recherche ? Qui fournit les meilleurs services ? Quel moteur est (et sera) le plus utilisé ? Quels sont les concurrents potentiels ? Comment s'opère la concurrence ? Quel sera le futur de la recherche ?

Pour répondre à ces questions, il conviendra d'étudier les produits des trois principaux moteurs de recherche. Nous nous cantonnerons en effet à cette solution, car étudier la plupart des moteurs de recherche nous amènerait d'une part à être moins qualitatif et d'autre part n'aurait que peu de rapport avec le sujet, tant les Big 3 dominant le marché. Cependant, il faudra néanmoins comparer les produits des majors à ceux qui se sont spécialisés dans tel type de marché.

Nous devons aussi étudier les caractéristiques des produits ainsi que les développements poursuivis. Enfin, nous tenterons de dégager les perspectives et prospectives et nous comparerons les caractéristiques de ces moteurs de recherche

Première partie – Fiches Techniques des Big 3

Il s'agira de présenter ici les caractéristiques principales nécessaires à la compréhension de l'analyse concurrentielle de ces moteurs de recherche. Ne seront dégagées ici uniquement les key issues.

Pour chaque moteur de recherche, cette partie comprendra :

- une fiche technique ainsi que des sites clefs permettant d'approfondir l'information
- une fiche produit avec un bref commentaire sur les produits et diversifications des sociétés
- une fiche analyse (SWOT), brève et succincte permettant de mieux cerner les enjeux, ainsi qu'une Webographie indicative pour aller plus loin dans la connaissance et la veille du moteur concerné.

Les moteurs étudiés ici seront :



Général

Site Web	www.google.com
Date de mise en service	07/09/1998
Robot Indexeur	Googlebot
Nombre de pages indexées	plus de 8 milliards

Caractéristiques

Nombre de langues acceptées	35
Nombre de mots acceptés par requête	30
Fréquence des mises à jour	28 jours
Type de fichiers indexés	plus de 220
Opérateurs Booléens	oui
Recherche Web	oui
Recherche Images	oui
Recherche Groupes	oui
Recherche Annuaire	oui
Recherche Actualités	oui

Particularités

Classement	PageRank
Dédoublonnage	oui
Préférences	oui
Personnalisation de la recherche	oui
Cache	oui
Traduction	oui

- Pour aller plus loin -



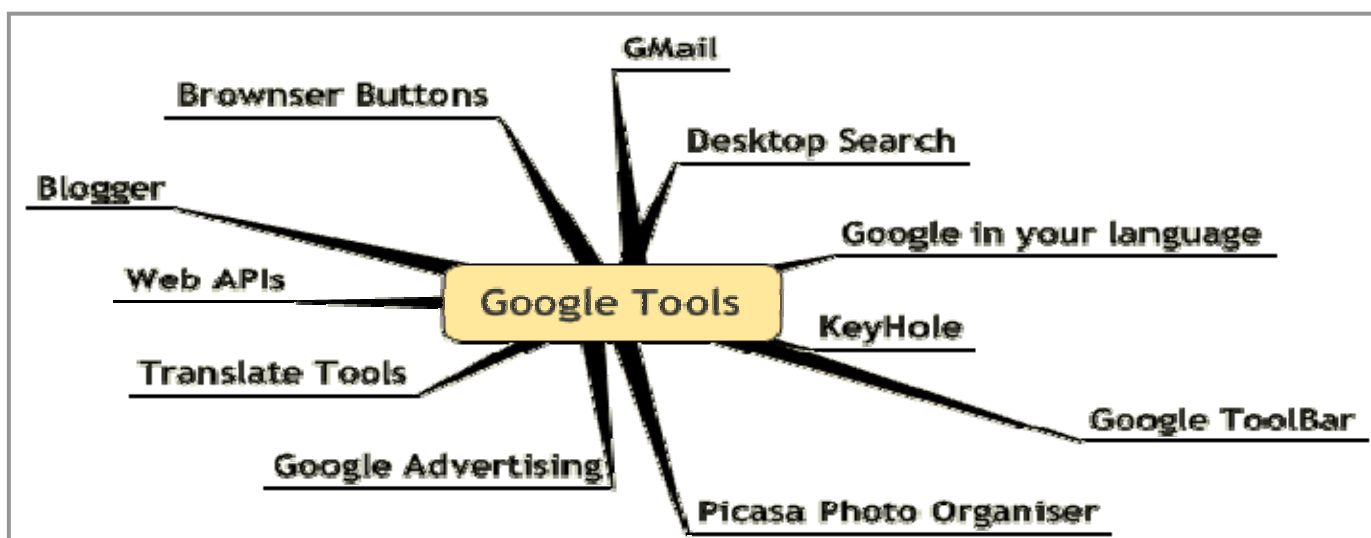
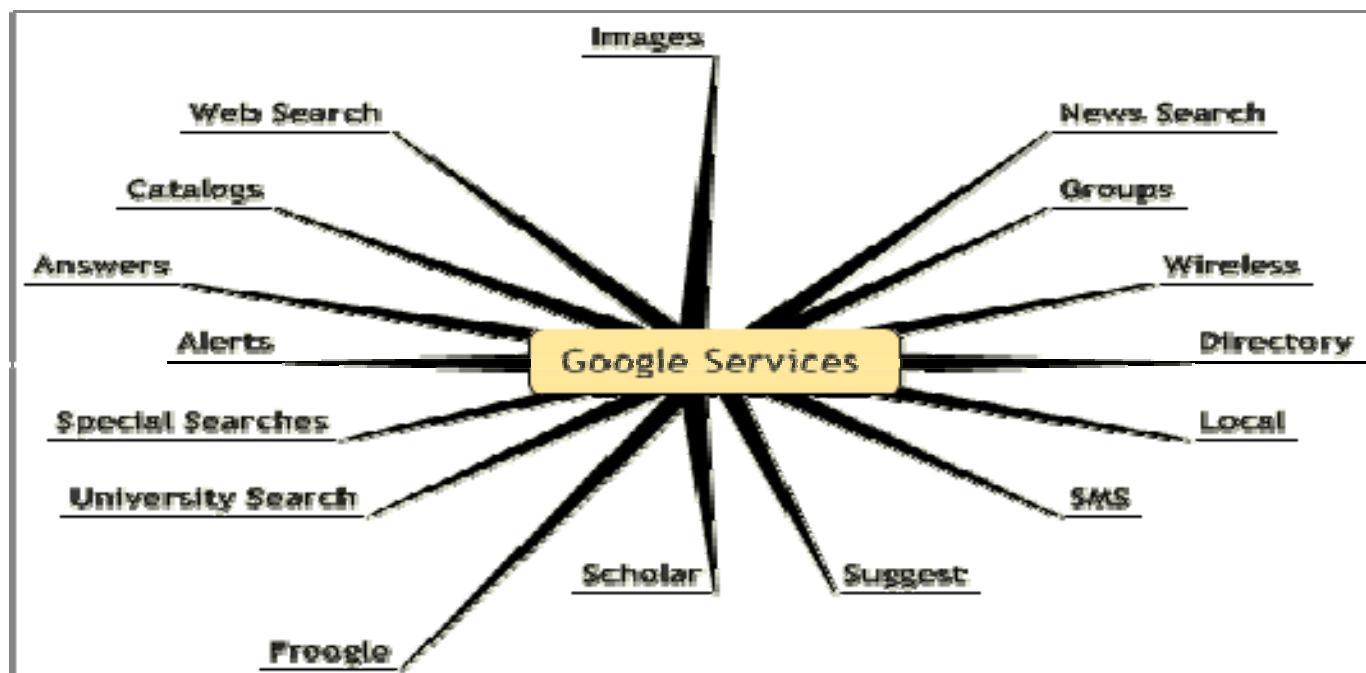
DSI (<http://www.dsi-info.ca/>)

Ce site est dédié aux sites de recherche (automates de recherche ou moteurs de recherche ainsi que les répertoires de recherche) sur le web. De nombreuses informations sur les moteurs de recherche ainsi que des méthodologies appliquées y sont présentes.



Google Indicator (<http://google.indicateur.com/>)

Annuaire spécialisé pour tout apprendre et tout savoir sur Google. Ce site propose de nombreuses catégories comme l'actualité, la recherche, la méthodologie, les services proposés, les partenariats...



Google propose deux types d'offre : les services et les outils. Ce moteur de recherche englobe les images, les NewsGroups, les actualités ou encore les annuaires. Mais la recherche de Google s'étend aussi sur des documents universitaires (Scholar), sur le shopping (Froogle). De plus, Google offre des services tout à fait innovants comme la suggestion de mots (Suggest), la veille automatisée (Alerts) ou encore la recherche d'universités US, de localisation US ou services SMS.

L'offre de Google s'agrémente d'outils qui ont plus ou moins à voir avec le domaine de la recherche. Parmi eux, on trouve un service de mail performant (Gmail), un logiciel de photo, un logiciel de cartographie par satellite, un service de Blog ou encore des outils de traduction. A cela s'ajoute la Google Toolbar, qui permet de naviguer plus facilement sur le web. Enfin, Google Advertising permet aux différents sites de se faire sponsoriser et par la même financer leurs sites.

- SWOT Analysis -

<p>Forces</p> <ul style="list-style-type: none"> - accès et service rapide - connu et populaire - pertinence et puissance de son algorithme - pertinence des résultats - cache intéressant - nombreux services proposés - Investissements en R&D nombreux 	<p>Faiblesses</p> <ul style="list-style-type: none"> - la plupart de ses services sont en version bêta - Moteur de référence donc souvent critiqué - Pas de technologie de cartographie ou de clustering
<p>Opportunités</p> <ul style="list-style-type: none"> - peut optimiser encore plus ses services - peut optimiser ses interfaces - peut aisément conquérir d'autres marchés grâce à son image de marque - peut innover et facilement supplanter ses concurrents sur des nouveaux marchés 	<p>Menaces</p> <ul style="list-style-type: none"> - MSN et Yahoo lancent leur propre moteur de recherche - Introduction en Bourse controversée - Souvent copié (Video Search) - Ne semble pas se préoccuper de services comme la cartographie ou le clustering qui peuvent devenir des enjeux clefs

- Pour en savoir plus -



Webrankinfo (www.webrankinfo.com)

Site web spécialisé sur le référencement, qui propose ici une étude complète sur Google et ses produits, sur le fonctionnement du Googlebot et sur les techniques de référencement de Google.

Site en Français.



Observatoire Google (<http://google-observatoire.blogspot.com/>)

Blog sur l'actualité du moteur de recherche. Se tient à jour de toutes les news de Google depuis juillet 2004.

Equivalent du Inside Google US

Site en Français



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

Blog officiel du moteur de recherche.

Toutes les nouveautés sont ici données en avant première.

Général

Site Web	http://search.msn.fr
Date de mise en service	02/02/2005
Robot Indexeur	MSNbot
Nombre de pages indexées	plus de 5 milliards

Caractéristiques

Nombre de langues acceptées	10
Nombre de mots acceptés par requête	150 caractères
Fréquence des mises à jour	21 à 30 jours
Recherche RSS	oui
Opérateurs Booléens	oui
Recherche Web	oui
Recherche Images	oui
Recherche Groupes	oui
Recherche Annuaire	non
Recherche Actualités	oui

Particularités

Classement	oui
Dédoublonnage	oui
Préférences	oui
Personnalisation de la recherche	oui
Cache	oui
Traduction	non

- Pour aller plus loin -



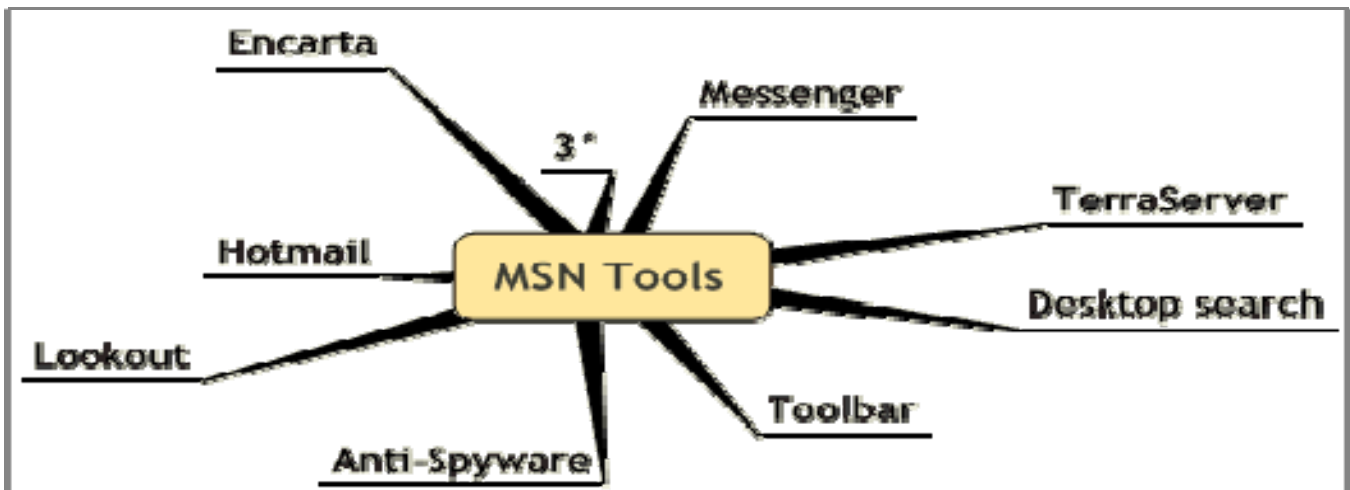
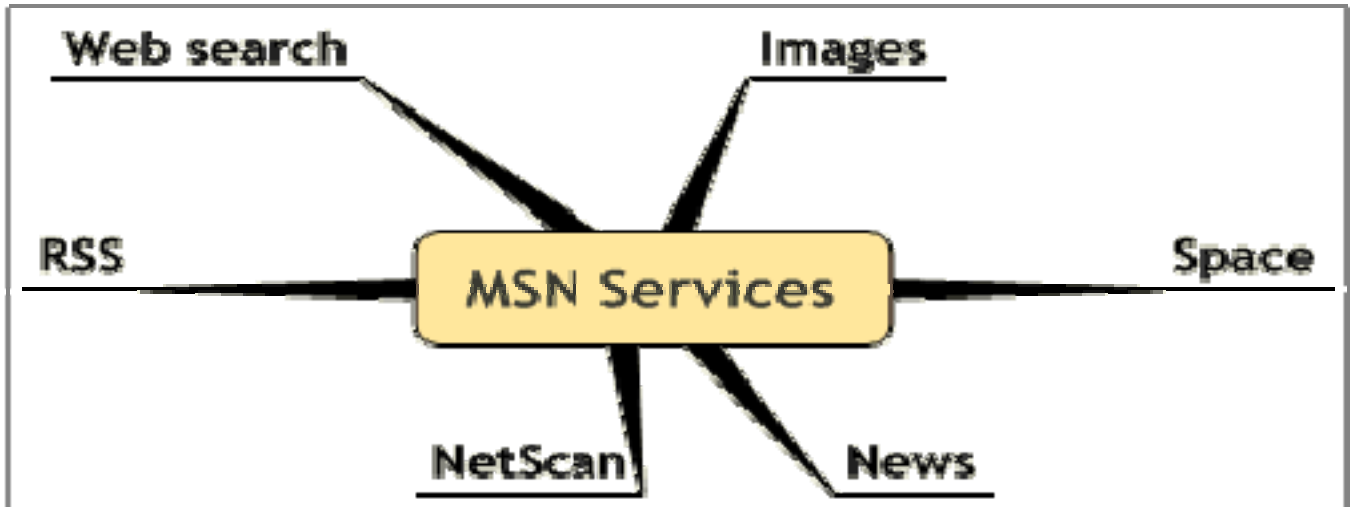
DSI (<http://www.dsi-info.ca/>)

Ce site est dédié aux sites de recherche (automates de recherche ou moteurs de recherche et les répertoires de recherche) sur le web. De nombreuses informations sur les moteurs de recherche ainsi que des méthodologies appliquées y sont présentes.



Sam Mag (<http://www.sam-mag.com/P1,1,5,402,,,default.aspx>)

Ce site propose une description des produits de MSN Search ainsi qu'une lettre de Bill Gates qui permettent d'en savoir un peu plus sur le tout nouveau moteur de Microsoft.



L'offre de MSN est toute nouvelle. La firme de Redmond doit en effet réagir rapidement afin de faire face à Google sur le marché des moteurs de recherche en particulier et sur le marché d'Internet en général.

Pour cela, MSN développe des services sensiblement identiques à ceux de la société de Mountain View comme la recherche de news, d'images et surtout un moteur de recherche optimisé et internalisé. En plus de cela, MSN innove grâce au suivi des recherches en RSS, une gestion de Blogs (MSN Space) plus fine que le Blogger de Google. Net Scan permet quant à lui de rechercher parmi de nombreux newsgroups.

Le point fort de MSN réside surtout dans sa diversification produit. Bien sur, La société de Redmond profite de tout son savoir faire et de sa force de frappe financière pour acquérir des sociétés et innover. Service webmail (Hotmail), desktop search, toolbar, Lookout (recherche dans les mails), TerraServer (qui veut concurrencer Keyhole) sont tout autant d'outils destinés à prendre des parts de marché à Google. En revanche, Microsoft se dégage de Google sur d'autres segments tels la messagerie, l'anti-spyware ou la possibilité de consulter Encarta online (équivalent de Google Answers)

- SWOT Analysis -

<p>Forces</p> <ul style="list-style-type: none"> - puissance financière - bonne connaissance du marché - pertinence des résultats - la plupart des produits sont finis - nombreux services proposés - investissements en R&D nombreux - Possède plusieurs monopoles 	<p>Faiblesses</p> <ul style="list-style-type: none"> - Suit Google et semble avoir du mal à vraiment innover - mauvaise image de Microsoft auprès des internautes et à contrario, celle de Google est excellente - pas de technologie de cartographie ou de clustering
<p>Opportunités</p> <ul style="list-style-type: none"> - peut concurrencer la technologie de Google - peut redorer son image de marque en proposant des logiciels et application de très haut niveau - peut aisément conquérir d'autres marchés grâce à sa puissance - doit conforter ses monopoles 	<p>Menaces</p> <ul style="list-style-type: none"> - Google et Yahoo tentent de le supplanter et il est difficile de suivre leur cadence - mauvaise image de la société et les résultats des autres compagnies font que les internautes préfèrent se fidéliser ailleurs - ne semble pas se préoccuper de services comme la cartographie ou le clustering qui peuvent devenir des enjeux clefs. Manque d'innovations franches

- Pour en savoir plus -



MSN (<http://www.msn.fr>)

Ce site est le portail de la filiale de Microsoft. Tous les produits proposés par MSN y sont représentés.

Ce site est la vitrine de MSN, et on y trouve des informations sur MSN Search.



MSN Search Weblog (<http://blogs.msdn.com/msnsearch>)

Le blog officiel de MSN; On y trouve toutes les informations ainsi que les nouveautés concernant le nouveau moteur de recherche de Microsoft.

Général

Site Web	http://search.yahoo.com
Date de mise en service	17/02/2004
Robot Indexeur	Slurp, Feedseeker, MMCrawler
Nombre de pages indexées	plus de 4 milliards

Caractéristiques

Nombre de langues acceptées	35
Nombre de mots acceptés par requête	Indéfinissable
Fréquence des mises à jour	Environ 1 mois
Recherche contextuelle	oui
Opérateurs Booléens	oui
Recherche Web	oui
Recherche Images	oui
Recherche Groupes	oui
Recherche Annuaire	oui
Recherche Actualités	oui

Particularités

Classement	oui
Dédoublonnage	oui
Préférences	oui
Personnalisation de la recherche	oui
Cache	oui
Traduction	oui

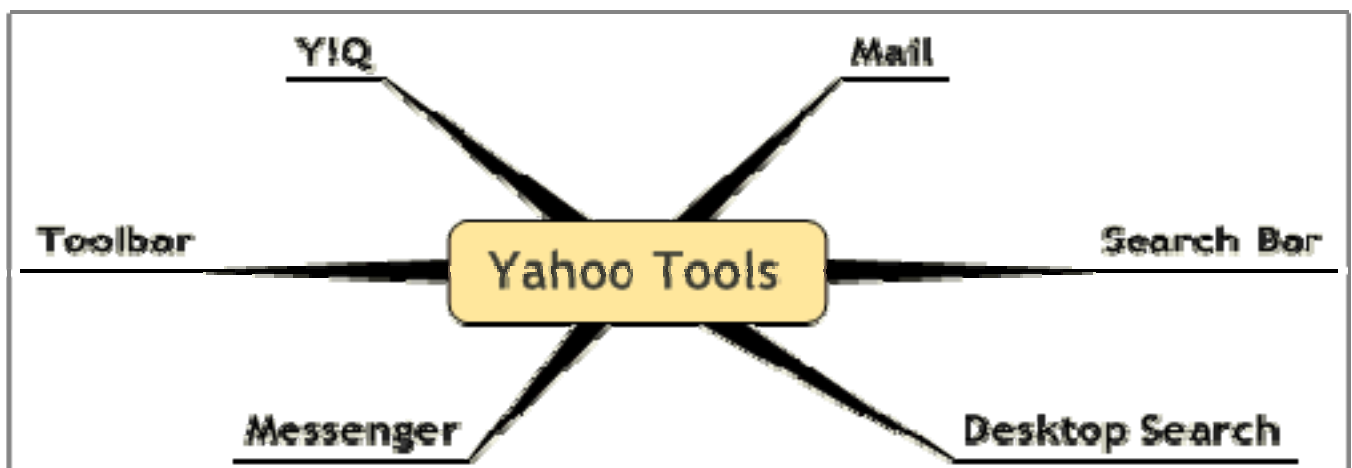
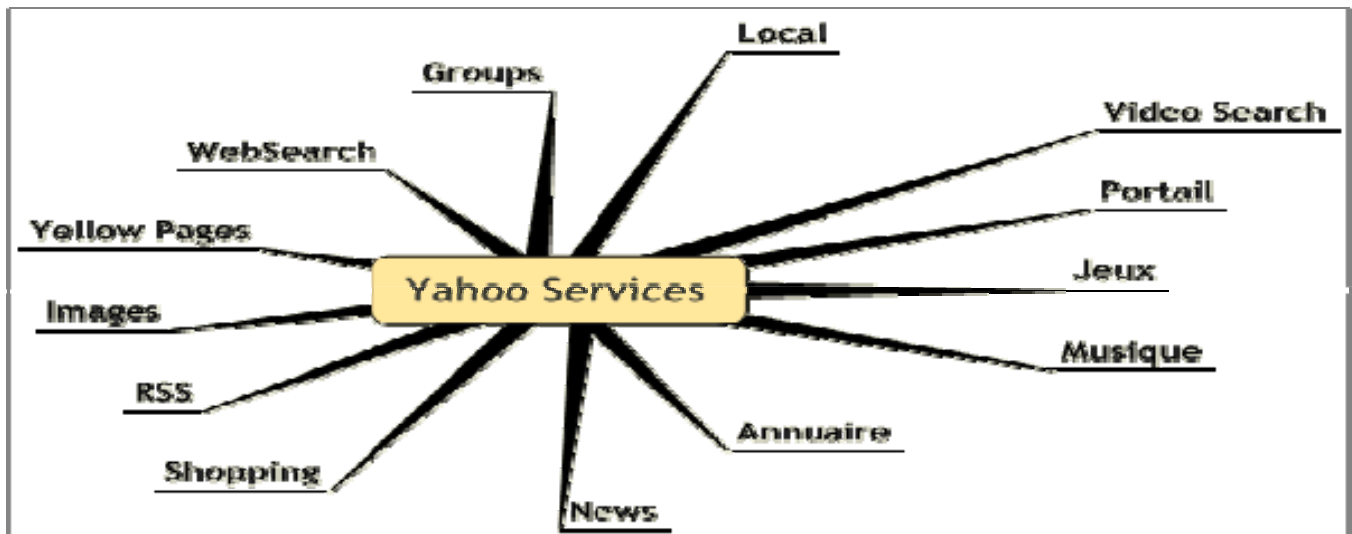
- Pour aller plus loin -



Les signets de la BNF (http://signets.bnf.fr/html/notices/n_3060.html)
 Etude sur le moteur de recherche Yahoo et sur sa façon de fonctionner.
 Collection de bookmarks pour une aide en ligne sur le moteur.
 Raccourcis pour accéder directement aux services Yahoo par cette web page.



Abondance (<http://outils.abondance.com/yahoo-moteur.html>)
 Abondance.com est un site de référencement qui propose sur cette page certaines caractéristiques techniques de Yahoo Search, ainsi que des tips de référencement sur l'un des pionniers du web.



Yahoo ! se concentre essentiellement sur les services. Comme ses concurrents directs, il propose de rechercher dans les groupes, dans les pages jaunes, dans les annuaires ou encore dans les news. Yahoo ! propose en plus d'autres services dont il est le seul à offrir l'accès comme les jeux ou la musique. De plus, il met à disposition un fil RSS sur ses recherches et, comme MSN, il possède son propre portail. Fort de son propre moteur de recherche depuis 2004, Yahoo ! se positionne comme Le concurrent de Google.

Cela se confirme avec les outils qu'il propose. Tout comme ses deux concurrents directs, il offre un service mail ainsi qu'une toolbar et un Desktop Search en version Bêta. Contrairement à Google et parallèlement à MSN, Yahoo ! possède aussi un client de messagerie. Les innovations sont aussi au cœur de Yahoo !. Premièrement, une taskbar est téléchargeable en version bêta qui permet de faire des recherches depuis son bureau et d'avoir les news directement sur la barre des tâches. Mais là où Yahoo ! se distingue réellement, c'est dans son nouveau concept Y !Q en version bêta qui est de rechercher contextuellement des informations en copiant/collant une page web. Ceci peut devenir la clé de la réussite pour Yahoo !

- SWOT Analysis -

Forces <ul style="list-style-type: none">- bien connu par le public du fait de sa longue présence sur le net- moteur de recherche performant- propose à travers son annuaire une multitude de services qui font que le grand public l'adopte en page de démarrage	Faiblesses <ul style="list-style-type: none">- pas de technologie de cartographie ou de clustering- ne suit pas Google sur tous les terrains.- s'est trop fait grignoter ses parts de marché par Google ce qui implique une remontée difficile
Opportunités <ul style="list-style-type: none">- Y!Q à développer- développer la qualité de ses services- indexer de plus en plus de pages- rejoindre Google en terme de qualité et de pertinence car les deux moteurs sont très proches.	Menaces <ul style="list-style-type: none">- arrivée de MSN- innovations technologiques de Google- pas aussi « cool » que Google ce qui peut lui porter préjudice au niveau des internautes

- Pour en savoir plus -



SearchEngineShowdown

(<http://searchengineshowdown.com/features/yahoo/review.html>)

Cette page offre une étude assez complète et technique sur Yahoo ! avec par exemple les opérateurs acceptés, et autres caractéristiques



Les moteurs de recherche

(<http://www.lesmoteursderecherche.com/yahoo.htm>)

Page web retrace l'historique de Yahoo ! depuis sa création jusqu'à l'achat d'Overture



Yahoo Official Weblog ([http:// www.ysearchblog.com/](http://www.ysearchblog.com/))

Weblog officiel pour se tenir au courant des nouveautés sur le moteur de recherche de Yahoo.

Deuxième partie – Présentation et Evaluation

Il s'agira ici de présenter uniquement les moteurs de recherche en eux-mêmes sans prendre en compte leurs « produits dérivés » comme les services ou les outils.

Après les avoir présenté et avoir résumé brièvement leur historique, l'étude s'attardera à comparer les caractéristiques de recherche de ces moteurs.

Ensuite, une tentative de comparaison sur les pages indexées poussera à évaluer la pertinence de ces moteurs, en distinguant à la fois les aspects techniques et en évaluant le degré de relevance des résultats proposés.

STRUCTURE :

1 – Historique et présentation des moteurs de recherche

1.1 – Google

1.1.1 – Présentation

1.1.2 – Historique

1.2 – MSN

1.2.1 – Présentation

1.2.2 – Historique

1.3 – Yahoo !

1.3.1 – Présentation

1.3.2 – Historique

2 – Caractéristiques comparées de ces moteurs de recherche

2.1 – Le nombre de pages crawlées

2.2 – Evaluation de la pertinence

2.2.1 – Aspects techniques du classement des sites

2.2.2 – Comparatif de la pertinence des résultats

1 – Historique et présentation des moteurs de recherche

1.1 - Google

1.1.1 - Présentation

Avec plus de 70% du trafic généré par l'ensemble des moteurs de recherche, Google est le moteur de recherche incontournable du moment. Développé à l'Université de Stanford, Google fait son apparition on line fin septembre 1999 après près d'un an de développement et tests.

Son succès s'est établi sur 2 axes principaux : son système de classement, le [PageRank](#) ainsi que la [rapidité](#) et la [simplicité](#) de son interface. En l'occurrence, Google a opté pour une interface très dépouillée, avec uniquement de la publicité sous forme de liens texte, en concentrant ses services sur la recherche de pages web et rien d'autre. Le moteur jouit également d'une très grande [rapidité](#) dans l'interrogation de sa base de données. Par rapport aux moteurs déjà existants, Google a proposé quelque chose de simple et de rapide. Il est nécessaire de rappeler que le haut débit à cette époque était loin d'être démocratisé, et que le temps de chargement des pages était long pour une grande partie des internautes. En allant sur des pages annuaire type Yahoo !, et ce, pour une simple recherche, l'internaute perdait beaucoup de temps. C'est pourquoi Google a marqué un point face à ses concurrents du moment. De plus, fort d'un [index](#) déjà très imposant, d'une [pertinence](#) prouvée en peu de temps et d'un temps de recherche optimisé (un petit script montre la durée de la recherche), Google a su se positionner dès son entrée sur le marché comme l'un des meilleurs.

1.1.2 - Historique

- Le 9 mai 2000, Google a officiellement présenté une interface configurable en 10 [langues](#), parmi lesquelles le français. Cette décision a été motivée par le fait que 35% des visiteurs sont des utilisateurs internationaux.
- C'est en fin juin 2000, tout de suite après avoir annoncé que plus d'un milliard de pages avaient été recensées (desquelles 560 millions enregistrées entièrement), que [Google a été choisi par Yahoo!](#) comme base de recherche pour les sites qui ne font pas partie de son archive
- Août 2002: [AOL](#), important provider américain et fameux portail [choisit Google](#) comme moteur de recherche pour le web en substituant ainsi la base de données d'Inktomi utilisée durant plusieurs années.
- Septembre 2002: Selon Google au moins la moitié des recherches effectuées chaque jour en début juillet provenaient de l'extérieur des Etats-Unis
- Octobre 2002: Google ajoute à ses services une nouvelle section dédiée à la recherche de nouvelles et les présente comme dans un vrai site qui, automatiquement, extrapole les nouvelles plus récentes et importantes prises de [4000 sources](#) officielles et différentes se trouvant sur le web.
- Décembre 2002: Google affirme avoir introduit certaines modifications à la méthode de crawling de sa base de données; modifications qui ont influencé

comme d'habitude, le positionnement ainsi que le trafic de nombreux sites importants

- Avril 2003: Google et le Content Target Adversting. Après l'acquisition de Blogger.com, Google propose à ses annonceurs une nouvelle façon de publier les AdWords: dans des pages Web liées aux mots clés choisis.

- Mai 2003 - Google a acheté Applied Semantics, société de services d'achat de noms de domaine, publicité contextuelle et solutions d'information retrieval pour la société

- Septembre 2003 - T-online, un des Internet Provider européen plus connus, a signé un accord avec Google pour la gestion de la recherche sur le net, en rompant son précédent rapport avec Overture et donc avec la base de données de FastSearch.

- Octobre 2003 - Fin de cet été Google a racheté la société Kaltix (www.kaltix.com), une jeune start-up née en juin de cette année et créée par 3 chercheurs de l'université de Stanford, qui développe des instruments pour la [personnalisation](#) des recherches.

Source : <http://www.lesmoteursderecherche.com/google.htm>

1.2 -

1.2.1 – Présentation

MSN a un parcours très différent de celui de Google. Issu de Microsoft, MS-Network était initialement développé afin de devenir une alternative aux leaders de service on-line type AOL ou CompuServe. L'effet escompté n'a pas eu lieu et MSN a progressivement migré vers un [portail on-line](#).

MSN Search était un peu différent d'un vrai moteur de recherche, et ressemblait plus à un catalogue de liens vers les produits et services de Microsoft. Même si une barre de requête était disponible, les liens tendaient essentiellement vers les informations de Microsoft.

Le problème est que les internautes demandent des requêtes pertinentes et non des liens « commerciaux ». Microsoft a donc tenté une première fois de développer son propre moteur de recherche mais sans succès. Après des partenariats avec Inktomi Altavista et plus récemment Overture, MSN a voulu redorer son blason sur le marché des moteurs de recherche, voyant Google prendre largement toutes les parts de marché.

Malgré tous ces efforts, MSN reste encore très loin de Yahoo et n'est pas en mesure de concurrencer Google. L'équipe de Redmond a alors décidé de développer [son propre moteur](#), avec son propre algorithme afin de se positionner au mieux sur le marché. La version bêta de ce moteur a largement bien fonctionné et MSN espère pouvoir rentrer efficacement dans la conquête des moteurs de recherche. Encore très méconnu au niveau des internautes MSN va lancer un buzz pour promouvoir son nouvel outil de recherche.

1.2.2 – Historique

- En septembre 1999, MSN Search laisse Inktomi pour un accord quinquennal avec LookSmart (pour les catégories thématiques des sites) et surtout avec AltaVista avec lequel les recherches standard sont effectuées.
- En avril 1999, sur MSN Search 3,5 millions de recherches ont été effectuées.
- En mars 2002 il a inséré parmi ses résultats ceux d'Overture, grâce à un accord renouvelé jusqu'en 2003, c'est pourquoi les 3 premiers résultats sont repris directement des plus de 53'000 annonceurs d'Overture.com.

Source : <http://www.lesmoteursderecherche.com/msn.htm>

1.3 – **YAHOO!**

1.3.1 – Présentation

Depuis le 17 février 2004, Yahoo ! utilise son [propre moteur de recherche](#) sur yahoo.com, et se libère ainsi de la présence de Google dans son univers. Il se pose en grand rival de Google et on devrait entendre parler de lui de plus en plus fréquemment. En effet, Yahoo! utilisait précédemment la technologie de Google et suite au rachat d'Inktomi et d'Altavista les équipes de recherche Yahoo! ont développé ce moteur qui leur offre l'indépendance par rapport à leur concurrent de toujours.

Sans compter qu'il est une des ressources on-line plus utilisée pour les recherches, Yahoo! est actuellement un des sites web les plus [fréquentés](#) au monde. En 1993, deux étudiants de l'Université de Stanford mirent en ligne leurs favoris. Une demande accrue pour rentrer dans leur catalogue en ligne est arrivée rapidement. C'est en [cataloguant manuellement](#) par rubriques que Yahoo ! s'est forgé sa réputation. Après une capitalisation en Bourse des plus réussie, Yahoo ! a pu se donner les moyens de ses ambitions. Optimisant son site, sa base de donnée (Oracle), et ses serveurs, Yahoo ! a pu gérer les millions d'internautes adeptes de son portail. En développant de nombreux services, Yahoo ! s'est positionné comme l'outil incontournable du Web.

Puis, pour faire face à la recrudescence des requêtes, Yahoo ! s'est servi du moteur de [Google](#) pour permettre aux internautes de rechercher sur le web. Mais Google prenant de plus en plus d'ampleur, Yahoo ! s'est vu contraint pour continuer à exister de développer son propre moteur de recherche, sorti en février 2004.

1.3.2 – Historique

- Fin 1993 – création de Yahoo ! qui à l'époque s'appelle « Jerry's Guide »
- 1995 – Yahoo est introduit en Bourse par [Séquoia](#) qui débourse pour 1 million de Dollars de part de marché
- 1997 – Yahoo ! a racheté la fameuse société Four11 à travers laquelle il peut maintenant offrir gratuitement des adresses e-mail et les services qui leur sont liés.

- 1998 - Yahoo! a conclu aussi un accord avec AT&T (qui a pris le dessus sur MCI avec laquelle il avait un accord précédemment) pour commercialiser des kits de connexion à Internet avec une interface personnalisée de Yahoo!.
- fin janvier 1999, Yahoo! a acheté [GeoCities](#), un des sites les plus visités au monde, et siège d'une communauté de 22 millions de pages Web.
- Octobre 2002, Yahoo surprend et se renouvelle. Après des mois et des mois de spéculations sur le nouveau partenaire technologique de Yahoo!, le colosse du web a choisi de renouveler le contrat, bien que lié à d'importants changements. Google reste ainsi le moteur Back-end de Yahoo! mais ses résultats font désormais partie intégrante de l'annuaire, ainsi le travail de Google géré par sa technologie et celui human-powered de Yahoo! se fondent ensemble.
- Juillet 2003 [Yahoo ! achète Overture](#). Après avoir acheté Inktomi cette année, Yahoo! réalise un nouvel achat stratégique. Yahoo! achètera Overture pour 1.63 milliards de dollars en actions et en comptant, un vrai tremblement de terre dans le monde de la publicité online. De cette façon Yahoo! fait un bond en avant défiant le grand player du secteur, Google, également sur le plan du Keyword Marketing.

Source : <http://www.lesmoteursderecherche.com/yahoo.htm>

2 – Caractéristiques comparées de ces moteurs de recherche

2.1 – Le nombre de pages crawlées

Google se distingue des deux autres moteurs par la taille gigantesque de son index qui comprend 8,1 milliards de pages (et qui avoisinerait officieusement les 10 milliards de pages crawlées). Loin derrière, Yahoo ! et MSN évoluent entre 4 et 5 milliards de pages.

Bien entendu, le nombre de pages recensées *stricto sensu* n'est pas un indicateur relevant. En effet, comme le souligne Génération NT, certaines requêtes sont censurées par MSN. La présence du mot *sex* dans une page déclenche son « anti-indexation » dans MSN Search. Ces pages ne sont pas comptabilisées par MSN Search dans son total de pages indexées. Or une simple recherche sur Google pour le mot *sex* rapporte plus de 300 millions de réponses.

Google a commencé la guerre de la taille de l'index des moteurs de recherche en clamant un index de plus de 8 milliards de pages. Microsoft avait quand à lui planifié de prôner le titre du plus gros moteur de recherche en annonçant 5 milliards de pages.

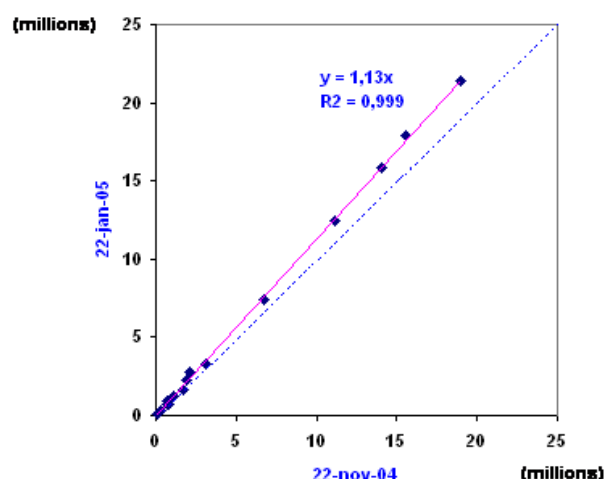
Mais réalité et statistiques sont souvent très différentes. Par exemple, Google inclus dans ses résultats ce qu'il appelle les pages « partiellement indexées » ou plus communément les pages « link-only ». Google connaît ces pages car des liens pointent vers elles mais rien n'en a été indexé. Or, normalement, la taille du moteur de recherche ne devrait pas compter les pages dupliquées, les pages de Spam... pourtant cela semble apparaître dans le compte de Google.

Yahoo quant à lui essaye de rester en dehors de cette guerre de la taille de l'index. Quand il a lancé sa propre technologie de recherche, il a refusé de donner des chiffres sur la taille de son index, et a seulement indiqué qu'il était comparable aux autres.

"As in the past, we are not disclosing the size of our index for competitive reasons. That said, we believe our index is highly competitive. Search quality is comprised of a variety of factors including freshness, relevance etc. and we continue to deliver high quality results for our consumers to ensure that they are able to find the best results for what they are looking for," *Stephanie Lwamasa – Porte Parole Yahoo*

La quantité de pages indexées n'a donc que très peu de valeur et il faut s'attacher beaucoup plus à la pertinence des résultats et à la profondeur des pages.

Une étude intéressante a été réalisée par Jean Veronis, publiée sur son blog [<http://aixtal.blogspot.com/2005/01/web-google-searching-9105590456-pages.html>], où il démontre clairement que Google indexe plus de pages qu'il ne prétend en avoir. Selon lui, Google, qui annonce 8 milliards de pages indexées, en recense plus de 9 milliards, chiffre qui devrait augmenter rapidement pour dépasser les 10 milliards fin Mars.



2.2 – Evaluation de la pertinence

2.2.1 – Aspects techniques du classement des sites

+ GOOGLE ET LE PAGERANK

Le PageRank est une **formule mathématique** utilisée par Google pour déterminer l'importance d'une page web. Un lien émis par une page A vers une page B est assimilé à un « **vote** » de A pour B. Plus une page reçoit de « votes », au plus cette page sera considérée comme **importante** par Google.

Il est à noter que toutes les pages n'ont pas le même pouvoir de « vote » qui, par exemple, émis par la page d'accueil d'un site majeur tel que Microsoft ou Le Monde, pèse beaucoup plus lourd qu'un vote émis par une page personnelle quelconque.

Le fonctionnement du PageRank [Traduction Google]

Nous assumons qu'une page A reçoit des liens (ou "votes") émis par les pages T1...Tn.

Le paramètre d est un facteur d'amortissement pouvant être ajusté entre 0 et 1.

Nous donnons généralement à d la valeur 0.85.

De même, C(A) est défini comme le nombre de liens émis par la page A (liens sortants). Le PageRank de la page A est défini comme suit :

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

Le PageRank peut être calculé en utilisant un simple algorithme itératif, et correspond au vecteur propre principal de la matrice normalisée des liens du Web.

Avec :

PR(A) le PageRank de la page A

PR(Tn) le PageRank de la page Tn

C(Tn) le nombre de liens émis sur la page Tn

d tous les « votes » sont additionnés, mais pour en limiter l'importance, le total est multiplié par ce coefficient d'amortissement (0.85)

1 - d permet de garantir que la moyenne des PageRank de l'ensemble des pages du Web sera de 1.

L'examen de cette formule permet de voir que le PageRank d'une page n'ayant aucun lien entrant sera de 0.15 ,

soit : $(1 - 0.85) + 0.85 \cdot (0) = 0.15$

En conclusion, le PageRank d'une page B ne dépend que **de 3 facteurs** :

- le nombre de pages Ak faisant un lien vers B,
- le PageRank de chaque page Ak,
- le nombre de liens sortants de chaque page Ak

Il ne dépend donc pas des éléments suivants :

- le trafic des sites faisant un lien vers B
- le nombre de clics sur les liens vers B dans les pages Ak
- le nombre de clics sur les liens vers B dans les pages de résultats sur Google.

+ MSN SEARCH

Le nouveau moteur de recherche de Microsoft étant très récent, il existe peu d'études ou d'interpolations sur la façon dont MSN classe les sites indexés. Toutefois, MSN Search utilise un ensemble d'instructions pour classer chaque page dans l'index. De nombreux facteurs sont pris en compte pour déterminer ce classement, les plus importants étant les suivants :

- **Langue** dans laquelle le site a été écrit
- Combien d'autres sites contiennent un **lien** vers ce site [popularité]
- **Qualité** du contenu et son volume

En règle générale, les sites qui sont mis à jour régulièrement, qui contiennent beaucoup d'informations et qui sont associés à de nombreux autres sites sont mieux classés que les autres. Le classement des sites change continuellement selon la fréquence de leur mise à jour, les nouveaux sites découverts par MsnBot et le type de contenu figurant sur chacun d'eux.

MSN utilise-t'il l'analyse au niveau des blocs ?

(Via WebMasterHub - <http://www.webmaster-hub.com/publication/article128.html>)

La première tâche importante consiste à identifier les "blocs sémantiques" différents sur une page. Pour cela, les chercheurs de Microsoft utilisent l'algorithme VIPS (VIsion-based Page Segmentation) : Segmentation de la page sur critères visuels. Il s'agit d'extraire la structure d'une page en se basant sur des informations données par sa présentation visuelle (en fait le code HTML de la page). Le résultat est une structure arborescente, dans laquelle chaque noeud constitue un bloc. Les blocs se voient attribuer un "degré de cohérence", qui permet de déterminer à quel point le bloc en question est clairement séparé des autres blocs.

A ce stade, il est possible d'éliminer le "bruit de fond" sémantique, en identifiant les blocs de navigation, les publicités, et les éléments de décoration, qui sont en général facilement reconnaissables.

Une fois les blocs identifiés, il reste à analyser les liens qu'ils contiennent, et la manière dont les blocs relient les pages entre elles...

Cela signifie que l'on ne va plus utiliser un graphe de relations "page à page", mais "bloc à page" et "page à bloc".

La relation "bloc à page" identifie le fait qu'un bloc sémantique contient un lien vers une page donnée (c'est une page entière, en l'absence d'ancre, un lien émanant d'un bloc pointe vers une page entière, et ne donne pas d'indication sur la partie de la page qui est "visée" par le lien).

La relation "page à bloc" identifie l'appartenance des blocs à une page donnée.

A partir de ces graphes de relations, les algorithmes HITS et Pagerank restent utilisables sans modification majeures : on obtient deux algorithmes dérivés, le BLPR (Block Level Page Rank) d'une part, et le BLHITS (Block Level HITS d'autre part).

+ YAHOO ! ET LE WEBRANK

Les résultats tirés de Yahoo ! sont triés par pertinence décroissante par rapport aux termes de la requête. Les critères utilisés par Yahoo ! sont nombreux parmi lesquels on peut citer le texte situé sur la page, le titre de la page, la description de la page, les liens associés à cette page, ainsi que d'autres critères.

Le WebRank de Yahoo est calculé différemment du PageRank de Google. En plus de calculer la relevance d'une page en fonction des **liens** la désignant, le WebRank prend en compte **le nombre de**

visites de cette page. Donc, lorsqu'un site très populaire et donc très visité pointe vers cette page, alors celle-ci s'en trouve intuitivement mieux référencé (car désigné par un « major », dont les visiteurs iront peut-être sur la page).

De plus, il est certain que Yahoo ! accorde beaucoup d'importance aux sites étant référencés dans des [métarépertoires](#) comme Yahoo Directory ou encore DMOZ.

2.2.2 – Comparatif de la pertinence des résultats

Un comparatif intéressant de la pertinence des résultats entre Google et Yahoo ! a été mis au point par le Journal du Net (<http://www.journaldunet.com/0411/041116googleyahoo.shtml>).

Plusieurs indicateurs ont été mis en place.

Google semble se montrer globalement le [plus performant](#) et rend de meilleurs services à l'internaute que Yahoo ! Cependant, l'écart entre les deux moteurs reste très minime, et le taux de satisfaction reste élevé pour les deux.

Sur la seule pertinence des résultats, l'écart entre les deux moteurs se révèle plus grand que l'écart général. Cela s'explique notamment par la plus grande pertinence des résultats pour Google lors des [recherches complexes](#). Toutefois, tous les deux obtiennent la moyenne sur ce critère. Il est à noter que le PageRank de Google jouit d'une excellente réputation et que la qualité de son algorithme n'est plus à démontrer.

Concernant le niveau de liens brisés, l'enquête permet de constater l'effort des deux moteurs pour tenir à jour leur index et éviter ainsi de pointer vers des pages supprimées ou déplacées. Sur ce critère Google devance Yahoo. Pour les résultats hors sujets, c'est-à-dire ne répondant pas du tout à la requête lancée, les deux moteurs sont sensiblement pareils.

En termes de résultats redondants, là aussi Google et Yahoo s'en sortent. Ergonomiquement, Google aurait pu d'ailleurs obtenir un score supérieur grâce à une présentation plus pertinente : lorsqu'il affiche sur une page de résultats deux liens pointant vers un même site (mais des pages différentes), il prend le soin de coller les deux résultats et d'afficher le second avec un léger décalage vers la droite. Visuellement, l'internaute peut donc constater que les deux résultats sont liés. En revanche, chez Yahoo, aucun effort particulier n'est mis en oeuvre pour signaler ces résultats d'un même site.

Sur le critère des résultats "banals", c'est-à-dire étant jugés comme liés à la requête mais n'y répondant pas directement ou seulement de manière partielle, Google démontre qu'il est capable d'afficher des résultats légèrement plus pointus que Yahoo.

Ce sont sur les pages que l'on constate le plus grand écart entre les deux. Google se révèle plus efficace que Yahoo pour chasser ce genre de pages inutiles dans l'avancée de la recherche de l'internaute (car purement publicitaires et souvent mal ciblées) ou qui faussent les résultats des moteurs.

Du côté de [MSN](#), les premiers essais ne sont [pas très concluants](#). En lançant la requête « qui est François Mitterrand » dans MSN Search et dans Google ce dernier renvoie dès le premier lien vers une biographie officielle du site de l'Élysée alors que MSN Search ne donne un résultat pertinent qu'au dixième lien. Une faiblesse corrigée par la recherche sur [Encarta](#), mais celle-ci n'est pas affichée directement. D'autres améliorations sont prévues, pour un moteur si jeune, et MSN va rapidement gagner en pertinence.

Quelle que soit la requête, Google, Yahoo et MSN donnent des résultats différents : une bonne chose, permettant de multiplier les chances de visibilité d'un site. MSN devra en revanche compter sur une pertinence accrue de ses résultats pour pouvoir gagner un public plus large. Il est certes déjà troisième du marché de la recherche Internet, et profite de la [force de son portail](#).

+ CONCLUSION

La question à se poser est « qui se place au mieux sur le marché de la pertinence et pourquoi les internautes choisissent tel type de moteur ? ».

[Google](#), malgré son très performant PageRank, semble quand même [limité](#) par rapport aux différentes offres de ses concurrents. En effet, il n'intègre pas son Google Directory, alors que [Yahoo !](#) a l'avantage de mettre en avant la catégorie de son [répertoire](#) correspondant à la requête. C'est aussi pour cela que Google est en train de se rapprocher de Wikipedia afin de compléter la pertinence de ses résultats, à l'instar de MSN et Encarta.

De plus, au niveau de la recherche [francophone](#), [Google](#) possède un [quasi-monopole](#), du fait qu'il n'a pas de concurrents sérieux. En effet, AskJeeves n'existe pas en Français, Yahoo ! FR est largement inférieur à Yahoo ! US, et Exalead est encore méconnu du grand public. Enfin, MSN, tout nouveau, est en phase d'entrer sur le marché, et n'est encore pas reconnu.

Google reste donc facilement le moteur de recherche préféré des Français.



Troisième partie – Différentiation par diversification

Il s'agira ici de comprendre quels sont les objectifs des moteurs de recherche. En effet, ils proposent tous de multiples services et de multiples outils pour l'internaute. Le but à demi avoué est de fidéliser les internautes en leur proposant tout ce dont ils ont besoin.

On remarquera que souvent, lorsqu'il y a une innovation ou tentative, les réactions des concurrents sont immédiates. Il est très difficile de se démarquer. De plus, nous mettrons en évidence la structure de la stratégie de diversification que ces firmes opèrent.

STRUCTURE :

1 – La guerre de diversification

1.1 – La bataille des services

1.1.1 – La recherche web

1.1.1.1 – La recherche image

1.1.1.2 – La recherche vidéo

1.1.1.3 – La recherche « locale »

1.1.2 – L'enjeu des RSS

1.1.3 – Concurrence sur d'autres type de services

1.1.4 – Les services innovants

1.2 – La bataille des outils

1.2.1 – Les Desktop Search

1.2.2 – Providers de Blog Spaces

1.2.3 – Guerre des Bar(res)

1.2.4 – Emails et groupwares : guerre de la communication P2P

1.2.5 – Autres innovations

2 – Tableau comparatif

2.1 – Méthodologie et objectifs

2.2 – Tableaux comparatifs

2.2.1 – Les services

2.2.2 – Les outils

2.3 – Analyse

1 – La guerre de la diversification

1.1 – La bataille des services

1.1.1 - La recherche web

La recherche web ne s'étend pas uniquement au domaine des pages html ou autre. En effet, les trois moteurs de recherche se livrent une guerre sans merci au niveau de la recherche web et cherchent à améliorer leur positionnement vis-à-vis des internautes en offrant des possibilités de recherches plus accrues et plus originales.

1.1.1.1 – La recherche image

La possibilité de rechercher des images a été l'une des premières dépenses en recherche et développement des moteurs de recherche. En effet, les internautes sont friands de photos, dessins, et ont pour habitude d'en rechercher beaucoup. Il est donc presque obligatoire pour ces méta-moteurs de proposer un tel service. Google propose [Google Image](#), qui permet de crawler le web et trouver près de 1.200.000.000 images. Dans le même esprit, Yahoo! a suivi Google et propose lui aussi plus d'1 milliard d'images. En revanche, le moteur de MSN ne permet pour le moment que de trouver 400.000.000 d'images.

Le concurrent : Cydral

Cydral est un moteur de recherche français d'images, qui crawl les images du monde entier, et qui a pour atout de proposer un filtre parental. A la grande différence de Google, par exemple, ce moteur de recherche ne censure donc pas des photos automatiquement (type les photos d'Abu Grahیب, etc.).

<http://fr.cydral.com/>

1.1.1.2 – La recherche vidéo

La nouvelle tendance, le nouveau bataillon de ces sociétés semble être la recherche de vidéos sur le Net.

Selon Softpedia, "*Video search is hitting prime time for a several reasons. Many people now have high-speed Internet access at home and work, and are warming to online video now that it's not excruciatingly choppy and slow. The costs of creating, hosting and delivering video have also dropped, making more multimedia available. Finally, the Internet is maturing into an entertainment platform for television, via convergence devices that combine PC and TV features, and search will be essential for people to find and watch media, whether its available over broadband, pay-per-view cable or broadcast.*

Le concurrent : Blinkx TV

Bien sûr, Blinkx n'est pas en reste, et, lui qui avait déjà séduit nombre d'utilisateurs sur le secteur des moteurs de recherche internes compte bien concurrencer les deux majors du club sur les news vidéos.

Blinkx TV propose les vidéos smart folders qui mettent à jour automatiquement leur contenu quand des nouveautés sont disponibles. Il permet aussi de télécharger des vidéos de bonne qualité, de comprendre le contenu de vidéo clips et d'indiquer à l'internaute quels sont les points clefs de la vidéo.

<http://www.blinkx.tv/>

*For search providers, offering searchable video is an extremely attractive new market because it not only keeps them **relevant to consumers hungry for multimedia**, but it also helps them appeal to **brand advertisers**, which spend about \$60 billion annually on commercials. Major TV advertisers are comfortable with the effects of commercials, and they're likely to wake up to Internet opportunities once on-demand video is ubiquitous.*"

On comprend donc pourquoi ces entreprises s'intéressent à ce pan du marché des moteurs de recherche. Les autres moteurs de recherche ont eux aussi voulu s'octroyer la plus grosse part du gâteau et se sont lancés sur ce marché. **Yahoo!** ou encore **Blinkx** veulent se positionner comme à la fois des précurseurs, mais aussi en même temps réaffirmer leur statut dans ce marché très concentré où l'innovation et la créativité en sont les enjeux.

C'est pourquoi **Yahoo!** s'est penché sur la "mode" de la recherche vidéo. Après une phase de développement rapide compte tenu du fait que Google n'allait pas tarder à rendre sa copie, Yahoo! a voulu se concentrer sur la technologie RSS, afin de pouvoir s'imposer sur le pan de la **vidéo news en temps réel**.

1.1.1.3 – La recherche « locale »

L'un des nouveaux axes de recherche dans la bataille du renseignement des moteurs de recherche est la recherche locale. Google vient juste d'annoncer sa version beta de **Google Map** (<http://maps.google.com/>) visant à proposer à l'internaute un service de cartes et de localisation de business. Plutôt bien réalisé, Google semble proposer le meilleur service de cartographie. De plus, il semble presque certain que la firme de Mountain View intégrera son outil de cartographie par image satellite (Keyhole). Ce service a été lancé par Google afin de contrer l'offre de Yahoo!, qui avait déjà proposé son service de cartes (<http://maps.yahoo.com/>). De plus, MSN Search, proposant aussi le même type de service de recherche vidéo [http://maps.msn.com/\(cxcmyxzibl3v2h55ftodetm1\)/Home.aspx](http://maps.msn.com/(cxcmyxzibl3v2h55ftodetm1)/Home.aspx), Google se devait d'être présent sur le marché. Chose que Google fait plutôt bien et qui lui donne un avantage et donc atout précieux dans la guerre de positionnement face à l'internaute. Il est à noter que la version 3 de la **Google Toolbar** propose étonnamment les services maps de ses concurrents.

Le concurrent : MapQuest

Mapquest est présent sur le marché depuis longtemps et jouit d'une assez bonne réputation. Leader incontesté du domaine, Mapquest pourrait voir ses parts de marché grappillées par les Big 3, et notamment par l'offre de Google.

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

1.1.2 – L'enjeu des RSS

La technologie RSS est en train d'exploser dans le monde d'Internet. C'est pourquoi il y a un fort enjeu à se positionner sur ce segment. Après quelques tests, **Google hésite** toujours entre le RSS et le format Atom, ce qui fait l'affaire de ses concurrents comme Yahoo ! ou MSN qui propose déjà des solutions de suivi de recherche de leur résultats. En version « très beta », MSN propose un service qui fonctionne très bien, mais pas très intuitif. En effet, pour les recherches, il suffit d'aller sur le moteur de recherche, de taper sa requête, d'ajouter "&format=rss" à la fin de l'URL et de copier ce lien vers un agrégateur. Toutes les nouveautés parue sur la requête s'agrègeront facilement et permettront d'être à jour sur les nouveautés parues sur le web.

Dans le même esprit, Yahoo ! propose un accès RSS pour ses internautes et d'une manière plus intuitive que MSN (http://e.my.yahoo.com/config/promo_content?.module=ycontent&.url=). Enfin, Google n'est pas en reste en proposant, et ce, plus tôt que ses deux concurrents, un service d'alerte web, qui propose le même genre de services, mais qui n'est pas basé sur la même technologie (<http://www.google.com/alerts>).

1.1.3 – Concurrence sur d'autres types de services

La concurrence sur les types de service se fait rude. Même si on remarque que MSN, contrairement à Yahoo !, n'axe pas sa stratégie sur le développement de services mais plutôt sur les outils, les Big 3 tentent d'être les meilleurs sur tous les segments. Petit tour d'horizon des principaux :

- La recherche sur les groupes de discussion montre que Yahoo !, MSN et Google ne veulent pas laisser passer un enjeu aussi important. En effet, proposer un tel service leur permet de fidéliser plus facilement les internautes. Sur ce point, Yahoo !, qui développe ce service depuis longtemps, semble avoir un temps d'avance sur ses concurrents.
- La recherche de news a été l'un des enjeux clefs de la fin de l'année 2004. En effet, ce service propose à l'internaute de se fidéliser au mieux aux moteurs, sans avoir à passer par un service tiers type *DayPop* ou autres. Cela a deux objectifs : le premier étant de se positionner comme un provider omnipotent face à l'internaute, le deuxième étant de barrer l'entrée à de nouveaux concurrents potentiels. Ici, Google semble dominer le marché même si quelques questions subsistent essentiellement sur le fait de la sélection des sources.
- La recherche Business, où Google semble avoir un certain monopole, le service de ses concurrents n'étant que très peu développé et accusant un certain retard.
- La recherche shopping, où MSN a commencé en premier, grâce à son portail. Le problème de l'exhaustivité s'est fait ressentir, étant donné que MSN ne proposait que des produits Microsoft ou partenaires de Microsoft. Yahoo !, égal à lui-même, propose une offre de qualité. Enfin, Froogle, développé récemment par Google, montre un certain potentiel et est en passe de remporter le marché, grâce à la comparaison de prix et le choix sur différents sites.

Dans le domaine de la recherche pure, les moteurs tentent de porter la compétition sur la pertinence. Pour cela, MSN faisait office de leader grâce à son service encarta, censé donner les meilleurs résultats pour les requêtes des internautes. Encarta possède une très bonne réputation, et permet à MSN d'obtenir une assez bonne satisfaction des internautes grâce à sa capacité de répondre à des questions posées.

Du côté de Yahoo !, il faut tout d'abord accepter le contenu Beatsham ainsi que la page qui se charge très lentement. L'Encyclopédie Yahoo ne prétend pas éclairer ses lecteurs sur toutes les facettes des questions posées mais une première est traitée. C'est un avantage non négligeable pour certains problèmes.

Voyant ses deux principaux concurrents suivre ces pistes, Google propose d'héberger gratuitement Wikipedia afin de peaufiner sa pertinence.

1.1.4 – Les services innovants

Google est incontestablement le leader de l'innovation dans les services, où ses deux concurrents n'osent pas s'aventurer ou on encore du retard. Google Scholar a été l'innovation majeure de la fin de l'année 2004 par Google. En effet, la recherche (grâce à de multiples partenariats) dans les documents universitaires a permis de franchir un cap dans le domaine du WebSearch.

Google développe aussi d'autres services plus ou moins finis tels que Google Suggest, qui permet de suggérer d'autres recherches à partir de mots clefs, qui ne marche que très moyennement, et se fait dépasser par d'autres moteurs permettant la clusterisation. Google Answer permet de poser une question à Google qui se charge d'y répondre au mieux. Des recherches plus approfondies sont toutefois payantes. Les University Search permettent quant à elles de trouver des universités. Enfin, Google Compute permet de réaliser des calculs scientifiques.

De plus, la possibilité de poser des questions à Google par simple SMS et recevoir une réponse rapide pour trouver quelque chose en fonction de l'endroit où est un service très innovant. Google est le champion du service.

MSN ne développe quant à lui que très peu de services, et donc, par voie de fait, innove très peu. Cependant, MSN propose uniquement un portail, largement connu.

Yahoo ! préfère quant à lui se concentrer sur les services. En plus de ce qu'il offre communément, Yahoo ! propose une logithèque assez impressionnante de jeux on-line, de recherche de musique etc.. Son portail est toujours aussi développé, et l'ensemble de ses services destinés à l'internaute lambda suffit largement à remporter une bonne dose de popularité.

Yahoo ! propose aussi des recherches dans les pages jaunes, ainsi que dans son annuaire très développé.

1.2 – La bataille des outils

1.2.1 – Les Desktop Search

Google affiche comme ambition que la facilité de recherche est telle qu'il n'est plus nécessaire de classer les documents. La recherche est aussi aisée que sur le web et se fait sur des informations rafraîchies en continu. Ainsi un courriel est immédiatement indexé dès qu'il est arrivé. L'index est stocké localement. Resterait à avoir des indications sur sa taille.

Google Desktop Search introduit une innovation majeure. Toutes les pages web visitées sont indexées dans l'état et le contenu fait partie des champs de recherche ultérieure. Les contenus de courriels sont présentés sous forme de conversation organisée en fil. Les résultats sont présentés sous forme d'une

Le concurrent : Copernic Desktop Search

Considéré par le plupart des internautes comme le meilleur Desktop search, Copernic se positionne comme le leader de ce segment de marché de la recherche.

Dans une étude comparative sur CNet (http://reviews.cnet.com/4520-3684_7-5536376-1.html?tag=cnetfd.wk), Copernic est de loin devant ses concurrents, notamment grâce à son interface intuitive, à ses performances d'indexation et surtout à sa rapidité qui le distinguent de Google.

De par beaucoup d'utilisateurs, Copernic est le meilleur.

<http://www.copernic.com/en/products/desktop-search/>

page classique de résultats identique à la page habituelle. Autre innovation: dans le cas d'une recherche sur le web les **réponses locales** (éventuelles) sont **également présentées dans la page de résultats**. Cependant, 1 Go sont nécessaires pour le faire fonctionner, aucune indexation en dehors du disque C: n'est prévue et le temps d'indexation est deux fois plus important que celui de Copernic.

L'arrivée relativement tardive de **Yahoo Desktop Search** est compensée par des fonctionnalités absentes des deux produits concurrents. La plus notable est certainement **l'outil de prévisualisation**, qui affiche dans la partie droite de l'écran les fichiers trouvés lors d'une recherche, ce qui dispense l'utilisateur du lancement d'une autre application. Plus fort encore, cet espace permet d'agir directement sur les documents : on pourra ainsi **éditer un fichier ou répondre à un e-mail** directement à partir de l'interface de YDS.

Outre la plupart des fonctionnalités de ses concurrents, **MSN** permet **l'intégration** de son desktop search à **Outlook**, **Internet Explorer** ainsi que **Windows Explorer**. De plus, ce logiciel d'intégrera à sa toolbar, ce qui permettra une recherche plus aisée.

Les trois outils proposés par les trois moteurs ont cependant tous le même problème : la confidentialité. En effet, il est difficile de faire confiance à un moteur de recherche pour ne pas divulguer sur Internet les documents privés de son PC.

1.2.2 – Providers de Blog Spaces

Le récent attrait des internautes pour les blogs et l'explosion du marché laisse entrevoir aux grandes firmes du net une potentialité de développement importante. C'est ce qui a fait que **Google** a acheté **Blogger**, un provider de services blogs. Toujours dans une optique de conquête de l'internaute, cela permet aisément de fidéliser et de s'appropriier des internautes pour la marque. Blogger est un des « gros » services mondiaux de blogs, et le rachat de Google lui a permis d'apporter de nouvelles technologies. Cependant, beaucoup d'internautes pensent que ce service est **trop limité** par rapport à d'autres concurrents, ce qui oblige Google et Blogger à investir en R&D.

De son côté, **Microsoft** propose lui aussi un service de ce genre, avec **MSN Spaces** qui possède beaucoup d'options et qui est tourné vers une clientèle jeune pour des blogs personnels. Beaucoup d'innovations qui font que ce service peut faire de MSN Spaces l'un des **préférés** des internautes qui veulent publier sur des thèmes légers et non professionnels.

A noter que Yahoo ! ne propose pas ce genre de services.

Le concurrent : Bloglines

Récemment acheté par Ask Jeeves, le concurrent le plus sérieux des Big 3, Bloglines, agrégateur de fils RSS qui propose à l'internaute de créer son propre journal en ligne en "piochant" parmi les fils RSS existant sur le Web. Mais Bloglines, créé en 2003, propose également un moteur de recherche de blogs et de fils RSS, un outil de création de blog, un annuaire de fils RSS et une interface utilisateur disponible en six langues dont le français.

www.bloglines.com/

1.2.3 – Guerre des bar(res)

Bien entendu, les trois moteurs de recherche possèdent leur propre toolbar, qui globalement possède les mêmes fonctions. Néanmoins, de petites différences existent. La [Google Toolbar](#) est très réputée et très puissante. Elle permet de faire des recherches web depuis n'importe quel site, d'éliminer les pop-up, de signer on-line intuitivement, de surligner les termes recherchés, Spell Checking, Web Translation, et Autolink pour US only.

Derrière, [MSN Toolbar](#) et [Yahoo ! Toolbar](#) remplissent les fonctions primaires, même si MSN propose plus de services. Il est à noter que Yahoo ! propose dans sa toolbar un anti-spyware, chose étrange vu la récente apparition de Microsoft anti-spyware en version beta qui ne l'a pas intégré dans sa barre.

Enfin, MSN semble posséder la meilleur DeskBar, qui ne nécessite pas un navigateur ouvert pour chercher sur le web, et permet en même temps de faire office de DesktopSearch. Yahoo ! Compagnon n'innove que très peu comme celle de Google.

Le concurrent : Kartoolbar

Avec sa barre d'outils, KartOO, le métamoteur à interface visuelle cartographique innove réellement :

- à la différence de ses "concurrentes", la Kartoolbar est flottante, et il est possible de la glisser n'importe où sur l'écran.

- un simple "copier" d'un texte permet de le placer automatiquement dans la boîte de recherche.

- la Kartoolbar permet de gérer les fenêtres (jusqu'à 20) d'Internet Explorer : navigation et fermeture.

- un menu Favoris aide à mémoriser les documents.

- un menu permet de fermer les fenêtres de pubs "pop-up".

<http://www.kartoo.net/kartoolbar/fr/index.htm>

1.2.4 – Emails et groupwares : guerre de la communication P2P

Après Yahoo ! et MSN, Google se devait de fournir un service webmail. Chose faite avec [Gmail](#), qui possède des caractéristiques intéressantes et innovantes : 1 gigabit d'espace, service rapide et possibilité de recherche dans les mails. Cette puissance de frappe a incité ses concurrents à tenter de s'harmoniser sur Google, avec 250 mégabits pour MSN et 100 mégabits pour Yahoo. De plus, MSN propose un logiciel permettant de chercher dans les mails. Mais l'avantage reste définitivement à Google.

Sur le point de vue [Groupwares](#), c'est [MSN](#) qui domine le marché. Deux offres très intéressantes sont proposées. La première, MSN Messenger, qui est une version de messagerie que la plupart des internautes choisissent. Yahoo ! possède lui aussi un service de messagerie instantanée, mais certainement moins bien aboutit que celui de Microsoft. La seconde, c'est [3°](#), un logiciel qui permet de parler, s'échanger fichier et photo, partager le surf, etc. Sur ce plan, Yahoo ! n'a rien prévu et Google a racheté Picassa, qui permet rapidement de s'échanger des photos et de chatter en direct. Mais [MSN](#) garde le quasi [Monopole](#).

1.2.5 – Autres innovations

La plupart des autres innovations sont encore en [version de test](#). On peut cependant dénoter [Keyhole](#) de Google, qui offre des cartographies du monde à partir d'images satellites, le [translate tools](#) qui permet de traduire des pages en différentes langues.

De son côté, la véritable « innovation de MSN » réside en son [anti-spyware](#), après le rachat du leader du marché Giant Antispyware.

2 – Tableau comparatif

2.1 – Méthodologie et objectifs

L'objectif ici est de comparer simplement les différents services et outils proposés par les moteurs de recherche, en tentant de classer leurs différents services. Cela permettra de dégager une ébauche de stratégie, c'est-à-dire où se diversifient le plus ces moteurs. Qui sont les leaders des services ? Des outils ?

Pour cela, nous allons comparer les différents services et les différents outils. Nous attribuerons des points comme suit : 3 points pour le leader, 2 points pour le deuxième, 1 point pour le troisième et enfin 0 point si le service ou l'outil n'est pas développé.

De plus, pour les services et outils innovants, étant donné qu'il n'y a aucune concurrence, nous attribuerons 2 points pour le service, car la plupart sont en version bêta. Un simple décompte du nombre de services ou outils décrits ci-dessus sera effectué.

2.2 – Tableaux Comparatifs

2.2.1 – Services

Services	Google	MSN	Yahoo !
Recherche Images	3	2	2
Recherche Vidéos	3	2	1
Recherche Locale	3	2	1
Pertinence (encyclopédie)	1	2	3
Gestion RSS	0	2	3
Groupes de discussion	1	2	3
Recherches News	3	2	1
Recherche Business	3	1	2
Recherche Shopping	2	1	3
Services innovants	12	2	8
Total	31	18	27

2.2.2 – Outils

Outils	Google	MSN	Yahoo !
Desktop Search	2	2	3
Weblog	2	3	0
Email	3	2	1
Groupwares	1	3	2
ToolBar	3	2	1
DeskBar	2	3	2
Messenger	0	3	2
Divers	4	2	0
Total	17	20	11

2.3 – Analyse

Sans surprises, [Google](#) est le [leader](#) incontesté des services proposés ainsi que des outils. Sa stratégie de diversification est payante et la firme de Mountain View est présente sur tous les fronts. Avec un total de 48 points, elle est largement devant ses concurrents qui en totalisent 38 chacun.

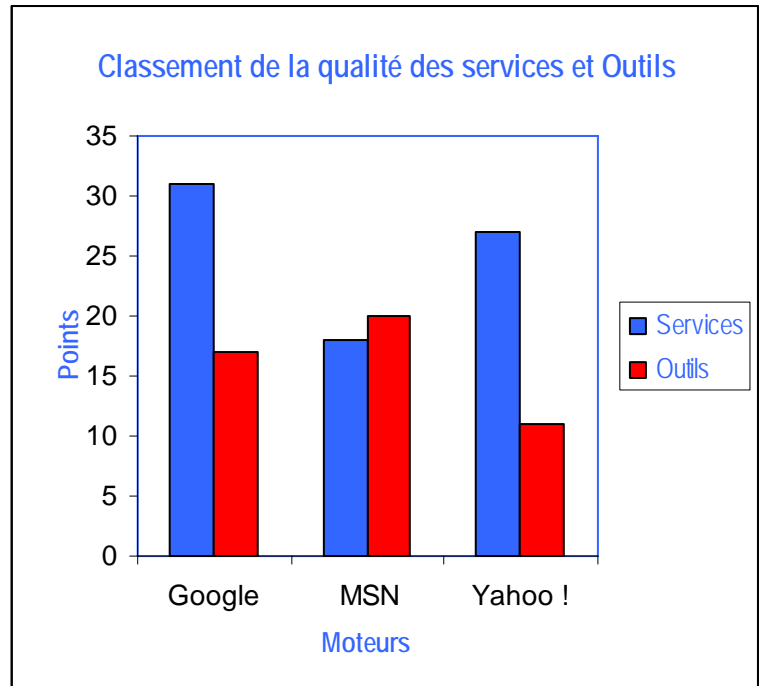
[Yahoo !](#) et [MSN](#) sont très proches. Mais plusieurs éléments se dégagent :

Premièrement, on remarque que [Yahoo !](#) se concentre essentiellement sur les [services](#), et n'offre que très peu d'outils. Sa stratégie est ici de se placer comme un fournisseur de prestations Internet, afin que l'internaute ait accès à tout ce dont il a besoin, en passant par le portail de Yahoo !

Sur le domaine de la prestation de services, Yahoo ! n'est pas loin de rivaliser avec Google. Cela vient essentiellement du fait que Yahoo ! propose des services que Google ne propose pas.

[MSN](#), quant à lui, se positionne sur le domaine des [outils](#). Il le fait bien car il obtient une meilleure note que Google. Ce qui fait véritablement la différence avec Google, c'est son MSN Messenger qui est très développé, alors que Google n'a quasiment pas investi dans ce créneau, et que très peu dans le groupware P2P. De plus, MSN est présent sur le domaine des services, et se positionne généralement en deuxième position. Le problème de MSN est qu'il n'est qu'un suiveur. Il ne propose en effet que des outils déjà existant, afin de ne pas trop perdre de parts de marché vis-à-vis de ses concurrents. Aucune réelle innovation n'est présente, ce qui peut se révéler comme une menace importante pour MSN.

C'est donc [Google](#) qui rafle la mise, [largement devant ses concurrents](#). Derrière, [MSN et Yahoo !](#) préfèrent opter pour une stratégie de spécialisation, l'un sur les outils, l'autre sur les services.



Quatrième partie – Analyses stratégiques et prospectives

STRUCTURE :

- 1 – Les stratégies de conquête
 - 1.1 – Bilan des stratégies de développement et de diversification
 - 1.1.1 – Stratégies de diversification
 - 1.1.2 – Stratégies de développement
 - 1.2 – La communication : outil prépondérant sur le Web
- 2 – Forces et faiblesses par rapport au marché
 - 2.1 – Enjeux et principaux concurrents
 - 2.1.1 – La gestion de recherche
 - 2.1.2 – Cartographie et Clusterisation
- 3 – Perspectives et prospectives

1 – Les stratégies de conquête

1.1 – Bilan des stratégies de développement et de diversification

1.1.1 – Stratégies de diversification

Les stratégies de diversification et de conquête du tout web font l'objet d'une concurrence féroce. Les trois majors veulent mettre la main sur tous les internautes et se rendre inévitables pour toute recherche sur le web.

Pour cela, les trois méta-moteurs innovent, comme nous l'avons vu précédemment. En proposant une multitude de services et d'outils pour le web, les majors fidélisent leur clientèle. Un exemple flagrant est la rumeur de [Google](#) avec le probable [Gbrowser](#). En effet, avec la popularité de Firefox, Google a engagé le chef de développement de Mozilla, avec pour but de préparer une alternative d'Internet Explorer. Dans ce cas là, la recherche et le web appartiendraient à Google.

Autre exemple, le Google DesktopSearch est censé aussi concurrencer Microsoft lorsque celui-ci sortira son Longhorn en 2006. Aucun avantage n'est laissé à son concurrent.

La bataille des DesktopSearch en est un exemple, celle des VideoSearch en est un autre. Les firmes veulent absolument se positionner comme le leader du marché, prônant l'innovation comme facteur clefs de réussite.

1.1.2 – Stratégies de développement

Le premier objectif de Yahoo ! est de devenir le leader mondial de l'information en ligne, s'entend l'économie, le sport, la culture, etc.

Mais la stratégie de développement de [Yahoo !](#) ne s'arrête pas là. Il propose un vaste choix de services à ses utilisateurs : une adresse et une boîte aux lettres privées; la possibilité de construire son propre site Internet ; la possibilité de se constituer un dossier contenant les adresses des sites préférés et les résultats de ses précédentes recherches ; l'accès à de nombreux chats thématiques ou privés ; l'achat en ligne d'un nombre très important de produits, avec en prime la possibilité de comparer les prix ; et le développement d'accès par mobile (WAP). Tous ces services, accessibles partout dans le monde, sont destinés à [fidéliser les utilisateurs](#).

A cela s'ajoute le [service pour entreprises](#). Yahoo ! offre une vaste palette de services cybermarketing, adaptés à la taille des entreprises. L'ouverture récente d'un bureau à Zurich répond ainsi à cette volonté de Yahoo ! de privilégier simultanément un marketing de proximité et des solutions globales.

[Google](#), quant à lui, mise sur sa [qualité d'innovation](#). En proposant tout (et n'importe quoi) à l'internaute, il veut que le web passe par lui et non par Microsoft. Google possède les moyens financiers et techniques pour mettre un genou à terre à ses concurrents. En quelques mois, Google est devenue une entreprise commerciale à part entière. Depuis son introduction en Bourse, Google ne cesse de grimper en haut du tableau, à un tel point qu'il devient plus gros que les plus grandes multinationales américaines.

La stratégie actuelle de Google est portée vers les [Adwords et Adsense](#), qui lui permettent de se rémunérer à partir du nombre de clics obtenus sur une fenêtre publicitaire. Mais qui dit entrée en Bourse dit comptes à rendre auprès des actionnaires. Or les actionnaires demandent toujours plus d'argent, donc la question est comment les rémunérer lorsque le modèle de Google aura atteint sa maturité voire sa saturation ? Il faudra que la firme de Mountain View se finance autrement. Toujours dans cette optique, il est possible d'extrapoler : qui est dans la capacité de financer une entreprise si gigantesque telle que Google ? La réponse est claire : d'autres entreprises. La seule chose que Google peut leur vendre est bien entendu des espaces publicitaires, et les classer au mieux dans son Page Rank. C'est en tout cas une possibilité. Le revers de la médaille serait une perte de pertinence et donc une migration des internautes vers d'autres moteurs de recherche. C'est pourquoi la solution pour Google reste le domaine de la [recherche et l'innovation](#). En se présentant comme le meilleur, Google encensera les internautes et pourra subvenir tranquillement à ses besoins.

[MSN](#) a une stratégie beaucoup plus [défensive](#). Tout d'abord, sur la diversification. MSN copie plus ou moins bien les logiciels ou les services de ses rivaux mais ne propose pas de nouvelles voies, de nouveaux produits.

De plus, la firme se fait attaquer sur ses domaines de prédilection comme les navigateurs Internet, les messageries, les comptes mails, etc. MSN tente de colmater les brèches tant que faire ce peut, et essaye de produire des outils de qualité. En effet, en tant que « suiveuse », MSN peut repérer les failles et les manques des offres de ses rivaux.

Enfin, MSN peut lancer des campagnes gigantesques à l'échelle de Microsoft pour promouvoir ses outils et services, ce qui est un atout considérable. Grâce à ce pouvoir auprès des mass média, MSN axe son développement sur des campagnes de publicité censée ramener des internautes vers son moteur, et donc générer de la publicité.

1.2 – La communication : outil prépondérant sur le Web

Les stratégies de communication sur Internet sont prépondérantes. En effet c'est le plus souvent le bouche à oreille qui forge la réputation. C'est comme cela que Yahoo ! a forgé sa réputation. De gros coup médiatiques, certes. Mais surtout la bonne réputation auprès des internautes. De plus, la marque, avec ses couleurs peu utilisées à l'époque de son lancement, a joui d'une très bonne popularité.

Mais depuis, [Google](#) a attiré la [sympathie](#) des internautes. En France, Google n'a fait aucune publicité ni de sponsoring. Sa popularité est venue essentiellement du bouche à oreille. D'abord connu des spécialistes (développeurs, "web addicts", fureteurs...) à la fin 1999, il est maintenant découvert et apprécié par le grand public. L'outil plait parce qu'il est bourré de qualités : rapide, assez pertinent, sans fioritures, nombreuses fonctionnalités et « fun ». Le design Google est, avec de la couleur et une communication résolument futée, apprécié des utilisateurs qui ne veulent pas s'ennuyer avec des textes trop long à lire ou de pages trop lentes à charger.

Parmi les principales armes de communication de Google : [les communiqués de presse, un réseau relais de fans, spécialistes ou non du web, et son logo innovant et amusant](#).

Pas une semaine (voire même pas un jour) sans que l'on parle de Google dans la presse spécialisée web, économique ou même grand public. « Quand Google inc. ne produit pas de

communiqués, ce sont d'autres sociétés qui se chargent de parler du moteur. Sur un mois donné, une bonne dizaine de firmes d'audience du net communiquent leurs résultats et baromètres: E-stat, Xiti/1ere position, JupiterMMX, Nielsen/Netratings, NetBooster/Weborama, etc.... Google est souvent en tête des classements et fait donc encore parler de lui, et l'évoquer dans le titre du rapport ou du communiqué, c'est une assurance... d'audience. » (Cristophe Asselin).

MSN, quant à lui, va se baser sur la réputation de Microsoft et bâtir sa popularité sur ses campagnes publicitaires qui seront d'ici peu lancées mondialement.

La communication a donc rapporté la mise à Google, ce qui lui a permis de distancer Yahoo ! et d'écraser MSN. Les internautes sont tous fans de Google, et lui préfèrent sa convivialité et sa pertinence aux autres moteurs de recherche. Stratégie payante pour Google de montrer une image « cool » tout en étant performant. C'est une chose qu'aiment les internautes.

Les autres : La communication classique.

Regardons les moteurs ou portails qui offrent des services de recherche aux internautes français.

Voila.fr bénéficie de la puissance média de son groupe France Telecom. Tous les médias sont utilisés : affichage 4x3, presse ("Faites un vœu et puis Voila"), sponsoring d'émission télé à forte audience (le Loft), événementiel (accompagnement de sortie de films) et bandeaux de pubs on-line. Qui plus est, l'internaute français qui s'abonne via Wanadoo est sensibilisé à Voila.fr dès ses premiers pas sur le net.

Lycos : "Lycos va chercher !" bien sûr. Tout le monde ou presque connaît la truffe du labrador (pardon, du Retriever) noir décliné en télévision, radio, presse, bandeaux on-line et que l'on retrouve habilement sur les sites du groupe.

Aol.fr : le portail vient de s'enrichir d'un très bon outil de recherche mis au point conjointement par Cybion et Exalead. Très difficile d'échapper aux CD-rom, spots télé et bandeaux de la société de Virginie. Une communication rouleau-compresseur assez agressive.

Yahoo! France: les français ont quasiment découvert l'internet avec cette marque mythique. Une communication tout en douceur avec ses chartes graphiques jaunes et violettes que l'on voit un peu en télévision, en radio et lors de quelques "coups" ponctuels : par exemple les portillons du métro parisien décorés aux couleurs de la marque.

Inutile peut-être pour Yahoo! de dépenser des millions lorsque l'on bénéficie d'une telle notoriété auprès des internautes. Pour l'ensemble de la population française, c'est autre chose. L'enquête CSA-Capital du mois de novembre 2001 révèle que 21% des français ne connaissent pas Yahoo! "Ces marques font beaucoup de bruit mais restent confinées à un public restreint, très parisien et plutôt branché" révèle l'étude . De là à parler de la France d'en haut et la France d'en bas, il n'y a qu'un pas.

Voila, Lycos, Aol et Yahoo! ne sont pas des outils de recherche à proprement parler mais des portails

qui proposent parmi leurs services un moteur de recherche (développé par Echo pour Voila, par AlltheWeb pour Lycos..). Les moteurs de recherche purs communiquent rarement en tant que tel. Car un moteur, c'est mécanique et technique, des algorithmes et des milliers de serveurs.

Une exception toutefois : Hotbot, qui a bénéficié au printemps 2001 d'une campagne de publicité presse et radio qui communique sur la précision. Le titre : "Vous ne trouverez que ce que vous cherchez", la signature : "Hotbot, la recherche de précision de Lycos". Le moteur fait en effet partie du groupe Lycos.

Intelligence Center (<http://c.asselin.free.fr/french/stratGoogle.htm>)

2 – Forces et faiblesses par rapport au marché

2.1 – Enjeux et principaux concurrents

2.1.1 – La gestion de recherche

C'est la véritable grande innovation de l'année 2004. La **gestion de recherche** consiste à « **bloquer des sites ou des pages pour améliorer la pertinence d'un résultat de recherche**, à sauvegarder des notices, à les trier, à les classer, à les partager » (DSI-info.ca).

Deux voies sont communément choisies : la gestion de site (Ask Jeeves, Ujiko et Yahoo !) et la barre d'outils (Fooxx et Looksmart). Ni **MSN**, ni **Google n'ont de gestionnaire de recherche** ; ce qui est étonnant pour des sociétés qui tablent sur l'innovation. Ceci peut peut-être représenter une menace pour eux dans leur optique de positionnement comme meilleur moteur de recherche web.

Le **blocage des notices** est la fonction majeure du gestionnaire. Il correspond à l'opérateur ou à la commande d'exclusion. Il permet d'augmenter la pertinence des requêtes en retirant des pages de sites d'achat si la requête ne les requiert pas durant une séance de recherche. Le retrait peut être permanent ou occasionnel.

Pour aller plus loin...

Le site DSI-Info.ca propose une étude complète sur ce sujet. L'analyse de cette partie en est tirée. Il est possible d'avoir un comparatif plus détaillé en allant sur ce site. Enfin, un tableau comparatif résumant les différentes caractéristiques y est présent.

<http://www.dsi-info.ca/chroniques/chroniques-recherche-web2.html>

Le **classement des notices** et spécialement celui des requêtes sont particulièrement utiles pour l'internaute qui fait des recherches fréquentes. Les requêtes peuvent être reprises périodiquement sans que cela soit des alertes et les nouvelles notices peuvent être classées dans leurs dossiers.

Le **partage des notices et des requêtes** fait appel à une communauté d'intérêt, qu'elle soit informelle ou organisée. Il se fait selon quatre canaux : par impression des notices, par courriel, par communautés et par syndication. Les quatre sites de recherche exploitent le partage par courriel. Fooxx ajoute le partage par communautés et My Yahoo ! Search par syndication.

Ask Jeeves se démarque par la gestion de **recherche publique**. **Fooxx**, quant à lui, privilégie **l'approche communautaire** de la recherche et une barre d'outils. **Ujiko** se distingue par ces **icônes** encadrant les notices et la valorisation des notices. Enfin, **Yahoo !** propose un gestionnaire de recherche presque **complet** alliant simplicité, présentation graphique conventionnelle regroupant le tri, le classement des pages et des requêtes, le blocage de sites et un « bouton » d'enregistrement de pages web extra-muros.

2.1.2 – Cartographie et Clusterisation

La deuxième grande tendance de l'année 2004 a été la catégorisation des requêtes et la présentation des résultats sous forme de graphiques. Etrangement, [les 3 majors ne se sont pas tournés vers ce type de solutions](#) qui pourtant, pourraient bien révolutionner le domaine de la recherche sur le web.

De nouveaux outils sont apparus qui permettent de [catégoriser un contenu d'information](#). Par exemple, une requête dans Vivisimo pour « business intelligence » offrira à l'internaute des catégories supplémentaires qu'il pourra interroger. Ici, Vivisimo lui propose: software, data warehousing, business intelligence solutions, overview, insight... Ce nouveau mode de recherche est l'une des clés des moteurs de recherche. Il apporte à l'internaute une nouvelle façon de rechercher, en classant automatiquement sa requête, et en lui proposant d'affiner ses recherches grâce à d'autres catégories qui font aussi office de mots clefs.

Parmi les meilleurs outils de ce type, on trouve Vivisimo, Brainspot, Clusty, Dogpile, Exalead, KWMMap...

Définition : Clusterisation ou Clustering

Faire de la clusterisation (ou « typologie », ou « regroupement », ou « classification automatique »), c'est partitionner une base de données en un petit nombre de sous-bases, appelées « classes », et telles que :

- * Deux individus appartenant à une même classe soient aussi semblables que possible.

- * Deux individus appartenant à deux classes différentes soient aussi dissemblables que possible.

En d'autres termes, la clusterisation tente de décomposer le nuage global de points, représentant l'échantillon, en quelques nuages bien compacts et bien différenciés.

Une autre tendance s'est dégagée aussi : la recherche par [cartographie](#). Utilisant la même technique que le clustering, les moteurs proposant la cartographie proposent en plus une représentation graphique des résultats. Ici, Kartoo, Mapstan ou encore Webrain sont les leaders. Des logiciels clients utilisent la même forme de présentation des résultats comme Grokker.

Cette forme de présentation de résultats [séduit](#) de plus en plus l'internaute, qui se trouve souvent perdu dans le flot de résultats de moteurs traditionnels. Cependant, le défaut principal de ce type de recherche est la lenteur de présentation des résultats, qui, quand l'internaute recherche quelque chose trouvable facilement, s'orientera vers un moteur plus rapide.

Etrangement, Google, MSN ou Yahoo ! ne proposent pas ce type de solutions. Pourtant, ils en ont la capacité technique et financière. Ceci pourrait leur causer des [dommages irréversibles](#) si les internautes commencent à s'habituer à ce type de solutions, d'autant plus que le nombre de pages web ne cesse d'augmenter, et qu'il est de plus en plus difficile de s'y retrouver et de trouver l'information utile et désirée.

3 – Perspectives et prospectives

Il s'agit ici de présenter les futurs développements qui auront lieu sur le marché des moteurs de recherche.

Revue de la 12ème semaine des conférences internationales sur [le World Wide Web](#) à Budapest, présentée par *Christophe Asselin - Intelligence Center*

Une équipe de chercheur japonais de NEC a présenté un document intitulé "Mining the Peanut Gallery" (PDF) décrivant un [outil](#) qui permettrait à des consommateurs [d'obtenir automatiquement des revues de presse sur un produit](#). L'outil Review Seer fonctionnerait sur le recueil de notation de produits effectués par les internautes sur les newsgroups usenet ou sur les forums de sites web. Les internautes pourraient préciser leur recherche en indiquant les fonctionnalités du produit et les tranches de prix souhaitées. Un tel outil doit faire face à certaines difficultés de l'ordre de la subjectivité humaine : comme ces revues de produits qui ne comportent de nombreux commentaires négatifs puis se terminent avec une phrase positive...

D'autres intervenants se sont concentrés sur les nouvelles manières de présenter les résultats aux internautes.

Ben Shneiderman, professeur d'informatique à l'université du Maryland, est convaincu que la plupart des personnes communique mieux avec leurs ordinateurs quand elles peuvent voir des données avec des [interfaces graphiques](#), plutôt que simplement les lire ou les écrire. Il a ainsi présenté plusieurs outils de recherche visuelle, comme le "[TimeSearcher](#)", qui permet d'illustrer par des graphiques les résultats de recherche des données de résultat de recherche. Quant aux outils PhotoMesa et PhotoFinder, ils combinent navigateur et moteur de recherche spécialement conçus pour rechercher des images numériques

Par la suite les moteurs de recherche pourront [consulter les cookies](#) (problème de confidentialité à la clé !) afin de déduire de ses recherches passées que l'utilisateur recherche tel type d'info plutôt qu'un autre (quand vous recherchez avocat vous vous intéressez plutôt au fruit et non au juriste). L'utilisation des cookies pourrait aussi ne renvoyer que les nouvelles informations ou les données adaptées à la localisation géographique de l'internaute.

Mais la personnalisation des résultats de recherche ne sera possible tant qu'un moteur comme Google (déjà connu pour sa vitesse) ne sera pas 10 fois plus rapide !

Google pourrait accroître sa vitesse de recherche en utilisant des techniques développées par les chercheurs de l'Université de Stanford ("génitrice" des fondateurs de Google Page et Brin) qui ont présenté leurs travaux à la conférence.

Sepandar Kamvar, un de ces chercheurs, précise que beaucoup de fausses idées ont été écrites sur l'impact de leurs travaux : "Si nos résultats de recherche sont mis en oeuvre, Google ne pourra pas être 5 fois plus rapide, mais nous espérons toutefois un gain de vitesse de l'ordre de 30%". Selon Kamvar, on pourra bientôt espérer, à l'issue de ces travaux, accélérer la vitesse de recherche, rafraîchir les résultats et les personnaliser.

Le co-fondateur de Google Sergei Brin était parmi le public de la conférence.

A noter que la plupart de ces fonctionnalités évoquées pour des moteurs en ligne existent déjà pour via de gros logiciels de KM et de veille à interface propriétaire et installés localement sur un serveur : [Arisem KM Server](#), [Digimind Evolution](#)...

Comment vont réagir les trois majors ? Apparemment, des développements sont en cours. Mais leur stratégie ne tend pas véritablement à améliorer significativement leur moteur de recherche, mais augmenter leur offre de produits et services.

Pour [Microsoft](#), qui est censé devancer Google, la lutte ne sera pas si simple. La firme devra développer un [algorithme très performant](#) et obtenir une [pertinence](#) égale ou supérieure à son concurrent. Elle se devra aussi de [rafraîchir](#) les documents inclus dans sa base de. Enfin, la [taille](#) de sa base de données doit être égale ou supérieure à celle de Google, et tout comme lui, indexer entièrement chaque site.

[Google](#) quant à lui [maîtrise](#) parfaitement son terrain. La recherche. Grâce à la large couverture médiatique gratuite dont il bénéficie, on parle de Google partout. Il [fidélise](#) les internautes et gagne encore des parts de marché chaque mois. Il [séduit](#) les annonceurs avec ses "AdWords" et en profite pour créer une relation de dépendance avec les webmaîtres en distribuant généreusement ses "AdSense".

[Le positionnement de Yahoo est illisible](#). Son nouveau service payant de soumission doublé du coût par clic qui s'annonce aux Etat-Unis transforme Yahoo en un moteur élitiste. Aussi, il va en opposition chasser sur le terrain propre de ses filiales Overture pour le cpc et Inktomi pour la soumission. La stratégie Yahoo va certainement porter ses fruits dans le [domaine financier](#), mais si l'annuaire historique pousse son raisonnement jusqu'au bout, il y perdra son moteur et ne sera bientôt plus qu'un portail comme les autres.



Conclusion générale

	Google	MSN	Yahoo !
Caractéristiques			
Nombre de pages indexées	3	2	1
Nombre de langues acceptées	3	1	3
Nombre de mots acceptés par requête	3	3	3
Fréquence des mises à jour	2	3	1
Type de fichiers indexés	3	2	2
Recherche Web	3	1	2
Recherche Images	3	1	2
Recherche Groupes	1	2	3
Recherche Annuaire	2	1	3
Recherche Actualités	3	1	2
Particularités			
Classement	3	1	2
Dédoublonnage	3	3	3
Préférences	3	3	3
Personnalisation de la recherche	3	3	2
Cache	3	2	2
Traduction	3	0	3
Evaluation particularités	2	3	3
Services			
Recherche Images	3	2	2
Recherche Vidéos	3	2	1
Recherche Locale	3	2	1
Pertinence (encyclopédie)	1	2	3
Gestion RSS	0	2	3
Groupes de discussion	1	2	3
Recherches News	3	2	1
Recherche Business	3	1	2
Recherche Shopping	2	1	3
Services innovants	12	2	8
Outils			
Desktop Search	2	2	3
Weblog	2	3	0
Email	3	2	1
Groupwares	1	3	2
ToolBar	3	2	1
DeskBar	2	3	2
Messenger	0	3	2
Divers	4	2	0
Total	93	70	78

METHODOLOGIE :

Nous attribuerons des points comme suit : 3 points pour le leader, 2 points pour le deuxième, 1 point pour le troisième et enfin 0 point si le service ou l'outil n'est pas développé.

De plus, pour les services et outils innovants, étant donné qu'il n'y a aucune concurrence, nous attribuerons 2 points pour le service, car la plupart sont en version bêta. Un simple décompte du nombre de services ou outils décrits ci-dessus sera effectué.

Cette réflexion est bien entendue subjective, même si un effort d'objectivité est présent.

CONCLUSION

La suprématie de Google n'est pas à démontrer. Que ce soit d'un point de vue qualitatif ou quantitatif, le moteur de Mountain View est largement devant les autres. Cependant, plusieurs études ont tendance à montrer que ses deux concurrents principaux se rapprochent de lui.

En effet, MSN et Yahoo ! ne sont pas si loin, et même si MSN semble légèrement inférieur à Yahoo !, celui-ci a pour excuse d'être « récent » sur le marché.

Comme la montré cette étude, la guerre de positionnement s'axe sur 3 points principaux :

- + la diversification produit
- + la taille et qualité de l'index
- + les stratégies de communication

Encore là-dessus, Google est le champion toutes catégories. Ses concurrents sont obligés de se spécialiser sur des tranches particulières et suivre les avancées de Google pour le reste. Yahoo ! se spécialise dans les services, MSN sur les outils, et les deux tentent de suivre le rythme effréné imposé par Google.

Toutefois, les majors devront prêter attention aux nouveaux arrivés qui proposent des solutions innovantes et qui pourront leur prendre des parts de trafic importantes.

RECOMMANDATIONS POUR METTRE EN PLACE UN DISPOSITIF DE VEILLE

A l'aide d'un aspirateurs de site et d'un agrégateur de recherche, monitorer les :

- newsgroups
- sites spécialisés
- blogs dédiés
- moteur en eux même (section laboratoire)

Une grande quantité d'information est disponible sur ce sujet, et les sites principaux à monitorer seront présents dans la Webographie.

Webographie

+ Sites généraux

- + ZDNet (www.zdnet.fr)
- + WebmasterHub (<http://www.webmaster-hub.com>)
- + Vunet (<http://www.vnunet.fr>)
- + Génération NT (<http://www.generation-nt.com>)
- + 01.Net (<http://www.01net.com/>)
- + Le journal du net (<http://www.journaldunet.com>)

+ Sites dédiés aux moteurs de recherche

- + Les moteurs de recherche (www.lesmoteursderecherche.com/)
- + Abondance (www.abondance.com)
- + DSI (www.dsi-info.ca)
- + WebRanInfo (www.webrankinfo.com)
- + Indicateur (www.indicateur.com)
- + Intelligence Center (<http://c.asselin.free.fr/>)
- + Search Engine Watch (<http://searchenginewatch.com>)
- + Search Engine Showdown (www.searchengineshowdown.com)
- + Search Engine Guide (www.searchengineguide.com/)

+ Blogs

- + Internet News (<http://www.websearchguide.ca/netblog/>)
- + John Battelle's SearchBlog (<http://battellemedia.com/>)
- + Inside Google (<http://google.blognewschannel.com/>)
- + IT Ligentia (<http://iteligentia.over-blog.com>)
- + Yahoo! Search Blog (<http://www.ysearchblog.com/>)
- + Secret2Moteurs (<http://www.secrets2moteurs.com/>)
- + Google Blogoscoped (<http://blog.outter-court.com/>)
- + Technologies du langage (<http://aixtal.blogspot.com/>)
- + Urfist Info (http://urfistinfo.blogs.com/urfist_info/)

+ Mais aussi les sites des différents moteurs et notamment leur partie laboratoire