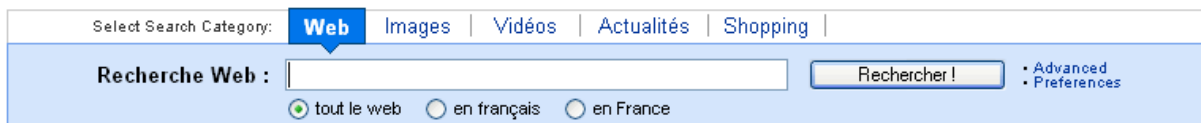


LES ENJEUX DE LA RECHERCHE D'INFORMATIONS SUR INTERNET



Select Search Category: **Web** | Images | Vidéos | Actualités | Shopping |

Recherche Web : [Advanced](#)
[Preferences](#)

tout le web en français en France

Sous la direction de : Catherine Boissonnet

Programme de Management Général
Année 2004/2005

Ecole Euromed Marseille
Domaine de Luminy - BP 921
13 288 Marseille cedex 9

REMERCIEMENTS

Je tiens à remercier ici les personnes qui, par leurs conseils et leurs encouragements ont contribué à l'aboutissement de ce travail.

Mes premiers remerciements iront à Madame Catherine Boissonnet, mon directeur de mémoire qui m'a fait bénéficier de son expérience et de sa connaissance d'Internet.

J'adresse également un remerciement à Laurence Bompard et à Pascale Blaché pour leur aide tout au long de ma formation.

Enfin, je tiens à remercier ma compagne, ma famille, mes amis et mon associé pour le soutien continue qu'ils m'ont apporté.

PLAN

Introduction	P. 2
1 – Internet, un nouveau média ?	P. 4
1.1 Un outil conçu par l'armée	P. 4
1.2 Un moyen de communication grand public	P. 10
1.3. Le commerce électronique	P. 14
1.4 Internet dans l'entreprise	P. 20
2 – Le marché de la recherche de l'information	P. 23
2.1 <i>Altavista</i> , le moteur quantitatif	P. 24
2.2 <i>Google</i> , l'introduction de la pertinence	P. 27
2.3 <i>Yahoo</i> , de l'annuaire au moteur de recherche	P. 29
2.4 <i>Microsoft</i> , le géant du logiciel	P. 30
3 – Les nouveaux champs d'application de la recherche	P. 33
3.1 Le commerce électronique	P. 33
3.2 L'actualité	P. 33
3.3 La culture	P.36
3.4 La vie privée	P.39
4 – La guerre pour un monopole	P. 41
4.1 Le marché mondial	P. 41
4.2 L'exception française	P. 43
4.3 Vers un monopole ?	P. 44
4.4 Les dangers du monopole	P. 46
4.5 Les alternatives	P. 47
Conclusion	P. 50
Glossaire	P. 51
Bibliographie	P. 52
Annexes	P. 53

INTRODUCTION

Beaucoup ont considéré le crash de la « *net économie* » en 2001 comme un signe annonciateur de la fin d'Internet. Pourtant, quatre ans après, Internet occupe toujours plus de place dans l'économie mondiale, mais également dans notre vie quotidienne. Toutes les entreprises ont aujourd'hui compris qu'Internet n'est pas un simple effet de mode, mais qu'il va probablement continuer à changer nos habitudes de consommation. En l'espace de quatre ans, les applications d'Internet n'ont eu de cesse de s'élargir. Aujourd'hui, nous surfons sur la toile pour chercher une information, réserver un voyage, payer nos impôts, consulter des offres d'emplois ou encore débattre de sujets divers. Ce n'est probablement rien en comparaison de ce que les entreprises du secteur nous préparent. Déjà source d'actualité préférée des jeunes, demain, Internet constituera certainement notre principale source d'informations, devant la télévision, la presse et la radio.

Pour accéder à l'information que l'on recherche, noyée parmi les milliards de pages qui composent Internet, une catégorie de sites a su s'imposer : « les moteurs de recherche ». Trois sociétés se partagent aujourd'hui près de 80% des parts du marché mondial de la recherche : *Microsoft*, *Yahoo* et *Google*. Véritable porte d'entrée du web, les moteurs de recherche nous orientent vers l'information souhaitée.

Quels sont les enjeux du marché de la recherche d'information ? Quels pouvoirs possèdent ces trois entreprises en passe de devenir la première source d'information des 900 millions d'internautes que compte le web en 2005 ?

De plus n'est-il pas dangereux de laisser ces entreprises orienter près d'un habitant sur cinq de la planète sur des sujets aussi divers que l'actualité, la culture ou l'histoire ; à plus forte raison si un des trois acteurs arrivait à imposer un monopole ?

Afin d'essayer de répondre à cette problématique, nous allons dans un premier temps revenir sur la construction du « réseau des réseaux ». Nous découvrirons ensuite les différents acteurs de la recherche d'information. Puis nous étudierons l'étendue des capacités des moteurs de recherche qui leur confère cette position de porte d'entrée du web. Enfin, nous analyserons la répartition des parts de marché, les raisons qui pourraient nous conduire à un monopole et ses conséquences.

1. Internet, un nouveau média ?

1.1 Un outil conçu par l'armée

Pour comprendre comment internet est structuré, il est intéressant de connaître les origines du réseau des réseaux.

1.1.1 L'idée de Paul Baran

Il y a une trentaine d'années, RAND Corporation, le groupe d'experts américains le plus en vue de la guerre froide, fit face à un problème stratégique important.

- De quelle façon les autorités américaines pourraient-elles communiquer sans problème après une guerre nucléaire ?

L'Amérique, suite à une guerre nucléaire, aurait besoin d'un réseau de commandes et de contrôle, relié de ville en ville, d'état en état, de base en base. Peu importe à quel point ce réseau serait blindé et protégé, son câblage et ses commutateurs resteraient vulnérables à une explosion atomique. Une attaque nucléaire réduirait n'importe quel réseau en lambeaux.

- De quelle manière ce réseau serait-il commandé et contrôlé ?

Toute autorité centrale serait une cible évidente pour un missile ennemi. Le centre du réseau serait le premier endroit à viser. Le RAND s'est penché sur ce casse-tête dans le plus grand secret militaire, et a trouvé une solution audacieuse. La proposition du RAND, une idée de Paul Baran membre de ce groupe, fut rendue publique en 1964. En tout premier lieu, le réseau n'aurait «aucune autorité centrale». Par ailleurs, il serait conçu dès le départ pour opérer même si le réseau est en lambeaux.

- Des principes simples.

On assumerait en tout temps que le réseau est inefficace. Il serait conçu pour fonctionner sans cheminement logique pour transcender sa propre inefficacité. Tous les nœuds (nodes) dans le réseau seraient égaux entre eux, chaque nœud ayant autorité pour être l'auteur, passer et recevoir des messages. Les messages seraient divisés en paquets (packets) adressés séparément. Chaque paquet serait originaire d'un nœud source pour se diriger vers un nœud destination. Il serpenterait dans le réseau sur une base individuelle.

Le chemin emprunté par un paquet serait sans importance, seul le résultat final importerait. Un paquet serait lancé d'un nœud à un autre, plus ou moins dans la direction de sa destination, jusqu'à ce qu'il parvienne au bon endroit. Si de grosses sections du réseau étaient détruites, cela n'aurait tout simplement pas d'importance ; les paquets resteraient aéroportés et circuleraient par les nœuds qui auraient survécu.

Ce mode de livraison au hasard pourrait sembler inefficace en comparaison avec un système comme le téléphone, mais il serait extrêmement résistant.

1.1.2 Premiers essais : Arpanet

Durant les années 60, cet intrigant concept d'un réseau commutateur de paquets, décentralisé et à l'épreuve des explosions, ne fut pas pris au sérieux par le RAND, le MIT et la UCLA. Cependant, le laboratoire national de physique d'Angleterre construisit le premier prototype de réseau basé sur ces principes en 1968. Peu de temps après, l'agence sur les projets de recherche avancés du Pentagone décida de financer un ambitieux projet de grande envergure aux États-Unis. Les nœuds du réseau seraient constitués par de puissants ordinateurs. Ces machines, rares et très coûteuses, étaient un pré-requis à la mise en place par l'agence d'un réseau solide pour le bien des projets nationaux en recherche et développement.

À l'automne 1969, ce premier nœud fut installé à la UCLA. En décembre 1969, les quatre premiers nœuds du réseau naissant étaient en place. Il fut appelé ARPANET, du nom de son commanditaire, le Pentagone. Les quatre ordinateurs pouvaient transférer des données sur des lignes de transmission dédiées à haute vitesse. Ils pouvaient même être programmés à distance à partir d'autres nœuds. Grâce à ARPANET, les scientifiques et les chercheurs étaient en mesure de partager les ordinateurs sur de longues distances. Ce fut un service très pratique, le temps de calcul sur ordinateur, appelé temps d'ordinateur, était très précieux au début des années 70. En 1971, le réseau ARPANET était constitué de 15 nœuds ; en 1972, 37 nœuds.

Pendant la seconde année d'opération, un fait singulier devint clair. Les utilisateurs d'ARPANET faussaient le partage du temps d'ordinateur pour en faire un bureau de poste électronique dédié, très rapide et subventionné par le fédéral. Le principal trafic sur ARPANET n'était pas constitué par de l'informatique à distance mais était utilisé pour transmettre des nouvelles et des messages personnels. Les chercheurs utilisaient ARPANET pour collaborer sur des projets, pour échanger des notes de travail et éventuellement pour des bavardages plus informels. Les gens avaient leur propre compte personnel sur les ordinateurs d'ARPANET, et leur propre adresse personnelle pour le courrier électronique. ARPANET était moins utilisé pour l'informatique à distance que pour la communication de personnes à personnes.

C'était peu de temps avant l'invention des listes de diffusion, une technique de distribution ARPANET avec laquelle un seul message pouvait être expédié à un grand nombre de souscripteurs sur le réseau. Fait intéressant, une des premières listes de diffusion populaires fut « SF-LOVERS » destinée aux amateurs de science-fiction. Discuter de science-fiction sur le réseau n'avait rien à voir avec le travail et au grand damne de plusieurs administrateurs des ordinateurs ARPANET.

1.1.3 L'arrivée du TCP/IP

Durant les années 70, le réseau ARPANET prit de l'ampleur. Sa structure décentralisée permettait une croissance facile. Contrairement aux réseaux corporatifs standards, ce réseau pouvait accommoder plusieurs sortes de machines différentes. La seule contrainte de ce système résidait dans le fait que les machines devaient comprendre le langage des paquets commutateurs de ce nouveau réseau anarchique. Leur marque, leur contenu et même leur propriétaire n'avaient aucune importance.

Le protocole de communication originel du réseau ARPANET était connu sous le nom de NCP, «Network Control Protocol» (Protocole de contrôle réseau), mais avec le temps et l'amélioration des techniques, le NCP fut remplacé par un protocole de haut niveau, plus sophistiqué, connu sous le nom de TCP/IP. TCP, ou «Transmission Control Protocol» (Protocole de contrôle de transmission), converti les messages en suites de paquets à la source, puis les rassemble en messages à la destination. IP, ou «Internet Protocol» (Protocole Internet) gère l'adressage, s'assurant que les paquets voyagent de nœuds en nœuds et même de réseaux en réseaux constitués de différents standards.

Déjà en 1977, TCP/IP était utilisé par d'autres réseaux pour se relier à ARPANET. ARPANET continua de subir un contrôle serré, au moins jusqu'en 1983, lorsque sa branche militaire s'est rompue. Mais TCP/IP les reliait tous. Même ARPANET, malgré sa croissance, devint un voisin de plus en plus petit parmi la vaste galaxie toujours grandissante des autres machines reliées.

Alors que les années 70 et 80 s'égrenaient, plusieurs groupes sociaux différents se retrouvèrent en possession de puissants ordinateurs. Il était relativement simple de connecter ces ordinateurs dans le «réseau des réseaux» toujours en croissance. Alors que l'usage du TCP/IP était de plus en plus courant, des réseaux entiers avaient la possibilité de faire partie d'Internet, et y adhèrent dans le désordre. Depuis que le protocole TCP/IP était entré dans le domaine public et que toute la technologie de base était décentralisée, il était devenu difficile d'empêcher des personnes de se joindre au réseau. En réalité, personne ne désirait les empêcher de

se joindre aux ramifications de ce réseau complexe, connu plus tard sous le nom de «Internet».

Le branchement sur Internet ne coûte rien ou presque, puisque chaque nœud est indépendant et doit gérer ses propres finances et besoins techniques. De même que pour le réseau téléphonique, le réseau informatique devint de plus en plus utile alors qu'il embrassait des territoires constitués de gens et de ressources grandissants.

Un télécopieur n'est utile que si tout le monde en possède un. Jusqu'à ce que chaque personne possède son télécopieur, il demeure une curiosité. ARPANET fut une curiosité pendant quelques temps. Puis les réseaux d'ordinateurs devinrent une nécessité absolue.

1.3.4 Les années 80 : Internet

En 1984, la Fondation Nationale de la Science (National Science Foundation) s'impliqua par l'entremise de son Bureau sur l'informatique avancée (Office of Advanced Scientific Computing). Le nouveau NSFNET établit une allure cinglante dans l'avancement de la technique, connectant de nouveaux « super ordinateurs » plus performants et en utilisant des liens plus gros, plus rapides, mis à jour et étendus, en 1986, 1988, 1989. D'autres agences gouvernementales ont emboîté le pas : la NASA, l'Institut National de la Santé (National Institutes of Health), le département de l'énergie (Department of Energy), chacun d'eux gérant une stratégie numérique différente à l'intérieur de la confédération Internet.

Les nœuds de ce réseau des réseaux en croissance étaient constitués de plusieurs variétés. Des ordinateurs étrangers, et quelques ordinateurs américains, choisirent de se démarquer par leur localisation géographique. Les autres furent regroupés à l'intérieur des 6 domaines Internet de base : *gov*, *mil*, *edu*, *com*, *org* et *net*. Ces abréviations peu élégantes sont une des caractéristiques du protocole TCP/IP. *Gov*, *mil* et *edu* sont utilisés pour décrire les institutions gouvernementales, militaires et éducatives. *Com* désigne des institutions «commerciales», qui envahirent le réseau, entourées par une poignée d'institutions sans but lucratif. *Org* et *net* étaient originellement destinées pour les passerelles entre les réseaux.

1.3.5 La fin d'Arpanet, la naissance du Web

ARPANET expire officiellement en 1989, victime de son succès éclatant. Peu d'utilisateurs s'en aperçurent car les fonctions d'ARPANET furent conservées et même améliorées par la suite. L'utilisation du standard TCP/IP dans les réseaux d'ordinateurs s'est progressivement généralisée. En 1971, seuls 4 nœuds constituaient ARPANET, aujourd'hui des dizaines de milliers de nœuds sont présents dans Internet, répartis dans plus de 42 pays, avec d'autres s'ajoutant chaque jour. Plusieurs millions de gens utilisent ce gigantesque réseau, père de tous les réseaux d'ordinateurs.

Internet est particulièrement populaire chez les scientifiques, et c'est probablement l'instrument scientifique le plus important de la fin du vingtième siècle. L'incroyable facilité d'accès qu'il offre aux données spécialisées et aux communications personnelles a grandement accru les recherches dans ce domaine.

Mais à ce stade, Internet est toujours, malgré son fort développement et sa mondialisation, un réseau complexe à utiliser, simplement disponible en mode texte. Bien heureusement, à la fin des années 1980, apparaissent l'hypertexte et le web, qui vont énormément faciliter l'accès et la circulation sur le réseau.

Le britannique Tim Berners-Lee, informaticien au CERN (Centre Européen de la Recherche Nucléaire), se base sur l'avancée technologique du Minitel français, pour développer son hypertexte sur le réseau. En effet, dès 1980, avec le Minitel, les français ont pris une certaine avance dans la Télématic. Il est possible, en appuyant sur un simple bouton, de circuler à travers le réseau minitel français et d'afficher des images.

L'Hypertexte consiste donc à pouvoir se déplacer sur le réseau, en de simples clics, afin de pouvoir transférer simplement des données, des images et du son.

Tim Berners-Lee mis donc au point un système d'informations organisées, basé sur l'Hypertexte (noté HTTP : Hyper Text Tranfert Protocol), dont les éléments les plus intéressants sont faciles d'accès : le World Wide Web.

En 1989, le Web fait donc son apparition. Pour accéder à une information, il n'est plus nécessaire d'utiliser une écran en mode caractères : le Web permet de sortir Internet d'un domaine restreint à des opérateurs techniciens spécialisés.

L'outil est donc graphique, puissant, très facile d'accès, beau et pas cher. Grâce à lui, vous pouvez visiter une exposition, lire votre journal, apprendre l'anglais, commander une pizza et que sais-je encore.

Le Web, autrement appelé World Wide Web (La toile d'araignée Mondiale : WWW), est aujourd'hui l'application la plus utilisée d'Internet. C'est lui qui a popularisé le vocable d'Internet au point que, largement utilisés par les médias, il est souvent restreint à la seule utilisation du Web.

Mais Tim Berners-Lee ne pouvant développer un navigateur permettant à tous les ordinateurs de parcourir son WWW, il lance un appel sur Internet. Mark Andreessen, étudiant à l'Université de L'Illinois et le futur fondateur de Netscape, se met à l'oeuvre. En 1993, la première version du navigateur Mosaic est en circulation. Internet touche enfin un large public.

1.2 Un moyen de communication grand public

La croissance d'Internet au début de 1990 fut spectaculaire, presque féroce. Internet s'est répandu plus rapidement que les téléphones mobiles et les télécopieurs. En 1993 sa croissance était de 20% par mois. Le nombre de machines hôtes avec des connexions directes TCP/IP double chaque année depuis 1988. Internet s'est détaché de ses origines militaires et d'institutions de recherche pour pénétrer les écoles élémentaires, les bibliothèques publiques et le secteur commercial.

Pourquoi les gens veulent-ils être «sur Internet» ? La raison la plus simple est la liberté. Internet est un des rares exemples d'une anarchie fonctionnelle, moderne et véridique. Il n'y a pas d'«Internet Inc.» Il n'y a aucun censeur officiel, aucun patron, aucun conseil d'administration, aucun actionnaire. En principe, chaque nœud peut parler d'égal à égal avec un autre nœud, dans la mesure où il respecte les règles du protocole TCP/IP, qui ne sont que des règles techniques et non sociales ou

politiques. Certains se sont battus pour une utilisation commerciale d'Internet, mais cette situation a été rapidement renversée alors que les compagnies fournissaient leurs propres liens.

L'anarchie d'Internet peut paraître curieuse ou même artificielle, mais cela a un certain sens. C'est un peu comme «l'anarchie» de la langue française. Personne ne loue cette langue et elle n'appartient à personne. En tant que francophone, nous avons la responsabilité de parler correctement français et d'en faire ce que nous voulons (même si les gouvernements fournissent une aide substantielle dans l'apprentissage de la lecture et l'écriture). Les personnes parlent, discutent, écrivent et cela fait évoluer la langue, elle est utile ; elle devient même fascinante et intéressante. Même si beaucoup de gens gagnent leur vie à utiliser, exploiter et enseigner la langue française, celle-ci demeure un bien public, une propriété publique. C'est la même chose pour Internet.

Internet est une institution qui résiste à l'institutionnalisation, ce réseau appartient à tout le monde et à personne à la fois.

Chaque groupe a ses propres revendications. Les gens d'affaires veulent qu'Internet repose sur de solides fondations financières, les gens du gouvernement veulent qu'Internet soit mieux contrôlé, les académiciens veulent qu'il soit consacré uniquement à l'éducation, les militaires le veulent sécuritaire et à l'épreuve de l'espionnage, et ainsi de suite.

Toutes ces sources de conflits demeurent étonnamment stables aujourd'hui, et Internet, jusqu'à maintenant, reste anarchique. Il y avait un temps où la haute vitesse du NSFNET et ses lignes à grande capacité étaient considérées comme «l'épine dorsale» d'Internet (Internet Backbone), et leurs propriétaires les utilisaient pour régner sur Internet. Mais aujourd'hui il en existe au Canada, au Japon et en Europe ; il y a même des épines dorsales privées appartenant à des entreprises et conçues spécialement pour transporter les données d'affaires. Aujourd'hui, même des ordinateurs personnels peuvent devenir des nœuds sur Internet.

Mais qu'est-ce qu'un individu peut faire avec Internet ? Fondamentalement quatre choses : le web, le courrier électronique, les groupes de discussion, l'informatique longue distance et le transfert de fichiers.

- **Le courrier électronique** ou l'email est considérablement plus rapide que la poste traditionnelle. Les utilisateurs américains d'Internet appellent cette dernière dédaigneusement le «snailmail» (la poste à pas de tortue). Le courrier électronique, lui, est semblable au télécopieur. C'est du courrier électronique instantané, mais vous n'avez pas à payer pour l'utiliser, du moins indirectement, et il est d'une portée mondiale. On peut également expédier des logiciels et des images numériques par le courrier électronique.

- **Les groupes de discussions** ou «newsgroups» forment à eux seuls un univers à part. C'est l'univers des nouvelles, des débats et des argumentations connu généralement sous le nom de «USENET». USENET est en réalité très différent d'Internet. Il est plutôt une énorme masse tourbillonnante de bavardages, de gens avides de nouvelles vagabondant sur Internet et prenant part à différentes discussions. USENET n'est pas un réseau physique, c'est un ensemble de conventions sociales. A tout moment il y a quelques 30,000 groupes de discussions dans USENET et leurs discussions génèrent près de 14 millions de mots dactylographiés chaque jour. Plusieurs journaux et publications sont publiés sur USENET et c'est gratuit.

Le courrier électronique et les groupes de discussions sont tous deux disponibles à grande échelle, même en dehors du noyau haute vitesse d'Internet lui-même. Les nouvelles et le courrier sont facilement accessibles par lignes téléphoniques standards.

- **L'informatique longue distance** fut l'inspiration originale d'ARPANET et c'est encore aujourd'hui un service très utile pour les programmeurs qui possèdent des comptes sur des ordinateurs très puissants. Les scientifiques bénéficient également de la puissance des super ordinateurs. Les bibliothèques offrent leurs catalogues de fiches électroniques gratuitement pour la recherche. D'énormes catalogues sur CD-

ROM sont de plus en plus disponibles par ce service. Et il existe une extraordinaire quantité de logiciels gratuits.

- **Les transferts de fichiers** permettent aux utilisateurs d'Internet d'avoir accès à des machines distantes pour la cueillette de programmes ou de documents. Plusieurs milliers d'ordinateurs sur Internet permettent à quiconque d'y accéder anonymement pour le transfert de fichiers publics, tout à fait gratuitement. Ce n'est pas une petite affaire puisque des livres entiers peuvent être transférés en quelques dizaines de minutes en utilisant un accès direct sur Internet. Aujourd'hui il existe plusieurs millions de ces fichiers publics disponibles à tous ceux qui en font la demande (et plusieurs autres millions de fichiers disponibles aux gens qui ont des comptes sur ces ordinateurs). Le transfert de fichiers sur Internet est devenu une nouvelle forme de publication, par laquelle le lecteur copie simplement le travail sur demande, dans la quantité qu'il ou elle désire, gratuitement.

Europe : les différents usages du Web par les internautes (Part des internautes, réponses multiples, en 2003)				
Usages	France	Allemagne	Royaume Uni	Moyenne UE 15
Courrier électronique	75 %	75 %	83 %	79 %
Chercher des nouvelles	66 %	74 %	66 %	74 %
Chercher des renseignements pour les voyages	57 %	66 %	74 %	63 %
Améliorer la formation, l'éducation	41 %	51 %	43 %	46 %
Chercher des renseignements sur la santé	26 %	40 %	44 %	38 %
Effectuer des opérations bancaires	30 %	33 %	30 %	31 %
Trouver des offres d'emplois	30 %	35 %	35 %	31 %
Réserver des spectacles	23 %	31 %	36 %	29 %
Participer à des forums	24 %	22 %	18 %	24 %
<i>Source : Eurobarometre</i>				
<i>Mis à jour le 24/04/2003</i>				

1.3 Le commerce électronique

Le commerce électronique, ou E-commerce, est probablement l'utilisation la plus récente d'Internet. En effet, c'est en 1997 qu'est née aux Etats-Unis la première boutique en ligne : la librairie *Amazon*. Son histoire est donc intéressante à ce titre. Fondée par Jeff Bezos, *Amazon* voit le jour en juillet 1995 à Seattle, sur la côte ouest des Etats-Unis. Quinze mois auparavant, au printemps 1994, Jeff Bezos fait une étude de marché pour décider du meilleur produit de consommation à vendre sur Internet. Dans sa liste de vingt produits marchands, qui comprennent entre autres les vêtements et les instruments de jardinage, les cinq premiers du classement se trouvent être les livres, les CD, les vidéos, les logiciels et le matériel informatique.

Anecdote dans les années 90, Internet représente aujourd'hui dans les pays développés près de 10% des transactions bancaires par carte de paiement. Avec une croissance de 30% par an, ce marché est au cœur de la stratégie des plus grosses entreprises mondiales.

Monde : le marché de l'e-commerce BtoC					
(En milliards de dollars)					
Sources	2000	2001	2002	2003	2004
eMarketer	60	101	167	250	428
Forrester Research	53	96	169	284	452
Gartner Group	-	-	-	380	-
IDC	59	118	-	213	54,2
Ovum	29	49	81	133	219

Source : eMarketer, Forrester Research, Gartner Group, IDC, Ovum *Mis à jour le 16/11/2001*

Sur tous les segments du commerce électronique destiné aux particuliers cinq catégories d'acteurs traditionnels et moins traditionnels se développent.

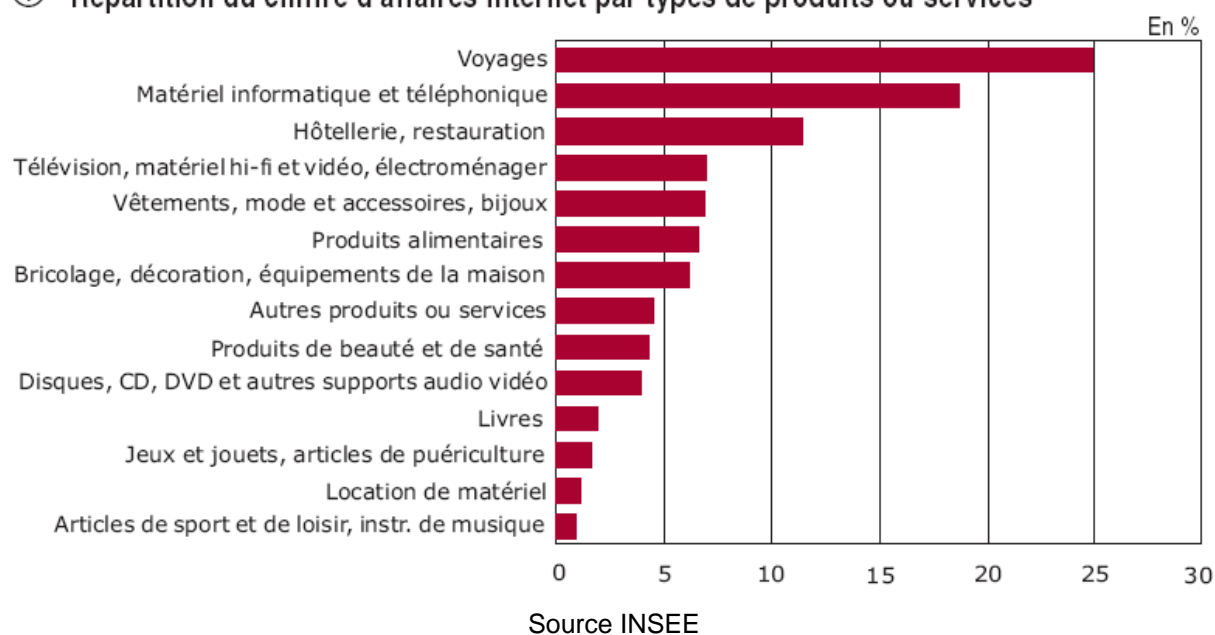
Les plus récents dits *pure players* n'ont qu'une activité purement électronique, sans aucun accès à un point de vente de biens ou de services. Les acteurs de la vente à distance ont modifié catalogue, prise de commande et paiement.

Les grossistes prennent aussi position sur ce nouveau marché. Enfin, d'autres grandes enseignes du voyage ou du commerce en magasin ont intégré internet comme un canal complémentaire de distribution.

Le chiffre d'affaires le plus important réalisé sur Internet concerne les voyages, suivis par le matériel informatique et l'hôtellerie. Ces transactions se sont multipliées grâce à l'effort entrepris pour gagner la confiance de l'internaute et lui offrir des services complémentaires. Fin 2003, les e-commerçants ont déjà effectué de larges investissements dans le domaine de la sécurité, et 90% d'entre eux offrent un système de sécurisation des transactions.

Les différents acteurs qui œuvrent dans l'activité de vente via Internet pour les particuliers (ou B to C) ont transformé la relation du client au commerçant en relation virtuelle, en offrant de nouvelles possibilités et de nouveaux services : passer sa commande alimentaire sans se déplacer, comparer des prix et des catalogues, voir la situation d'un hôtel sur un plan de ville à l'autre bout du monde avant de réserver, choisir des voyages bas prix à la dernière minute, acheter à un grossiste en se passant d'intermédiaires (graphique 1). Pour séduire leur clientèle, tous les e-commerçants ont dû investir soit pour se faire connaître et inspirer confiance, soit en offrant de nouveaux services. Ainsi les stratégies et les investissements sont-ils différents selon le type d'e-commerçants, déjà connu auparavant ou totalement nouveau.

① Répartition du chiffre d'affaires internet par types de produits ou services



① Les acteurs du commerce électronique

	Répartition du chiffre d'affaires internet (en %)	Chiffre d'affaires 2003 sur internet par entreprise (M€)	Part du chiffre d'affaires réalisé sur internet dans le CA total (en %)	Répartition du nombre d'entreprises (en %)
<i>Pure players</i>	8	1	100	17
Entreprises de la vente à distance *	25	15	12	28
Grossistes et industriels	9	45	3	15
Entreprises de services *	48	100	4	25
Click et magasins	10	22	6	15
Ensemble	100	40	5	100

Source INSEE

Nous l'avons vu, les pure players sont des entreprises exclusivement spécialisées dans le commerce électronique ; elles ne disposent pas de réseau de distribution physique (exemples : amazon.fr, lastminute.com). Deux autres types d'acteurs n'avaient traditionnellement que peu de contacts directs avec la clientèle : les entreprises de vente à distance ou vente sur catalogue, antérieures à l'essor d'Internet, qui ont intégré celui-ci comme canal supplémentaire de leurs ventes (exemple : 3suisses.fr) ; les grossistes et les industriels, pour qui Internet est un moyen d'accéder directement aux particuliers (exemple : dell.fr).

D'autres catégories d'entreprises pratiquaient et pratiquent toujours le contact direct avec la clientèle : les agences de voyages, qui ont ouvert des sites de vente de services en ligne (exemple : accorhotels.com) ; les détaillants traditionnels (dits aussi «click et magasins» de l'anglais click and mortar) qui ont vu dans le commerce en ligne un canal complémentaire de distribution de leurs produits (exemple : fnac.fr).

Parmi les pure players, un sur deux déclare avoir atteint le seuil de rentabilité en 2003, malgré d'importantes dépenses en publicité-marketing, qui atteignent 41% des coûts liés à l'activité de l'entreprise sur Internet. Ces entreprises sont orientées vers la vente d'une catégorie spécifique de produits. Elles déclarent mettre en avant sur leur site Internet le prix, le choix et l'originalité des produits. Leur stratégie dédiée à un segment de la vente par internet implique une priorité forte donnée au contenu du site : les pure players intègrent plus souvent l'évolution et l'enrichissement du contenu de leur site ainsi que sa maintenance technique. À l'opposé, ils sous-traitent davantage l'hébergement du site, la préparation de la commande et le paiement en ligne. Il s'agit de petites unités : 79% d'entre elles n'emploient aucun salarié. Leur activité est concentrée sur la vente de matériel informatique ou téléphonique (28%), la vente de matériel de télévision, hi-fi, vidéo et électroménager (22%), la vente de produits de beauté et de santé (20%) et les produits culturels tels que les CD, DVD et livres (15%).

L'essor d'internet a transformé la vente à distance

De 1999 à 2003, le développement du commerce électronique a transformé la vente à distance. L'activité de vente sur Internet des entreprises traditionnelles de la vente par correspondance s'est fortement accrue. Ainsi, fin 2003, les e-commerçants au sein de la vente à distance sont 4 fois plus nombreux que fin 1999. Leur chiffre d'affaires sur Internet a été multiplié par 6, passant de 150 millions d'euros en 1999 à 900 millions en 2003. La part des ventes qu'ils réalisent via Internet passe de 5% à 12%.

En 1999, la vente sur Internet était principalement concentrée sur les livres, les disques et supports audio-vidéo (CD, DVD, VHS), et les vêtements et accessoires de mode. En 2003, la gamme s'est étoffée : les vêtements, accessoires et bijoux représentent 24% des ventes contre 13% en 1999, les produits liés au bricolage et à l'équipement de la maison 19%, la télévision, le matériel hifi et vidéo et l'électroménager 16%, le matériel informatique et la téléphonie 11%. Enfin, les disques, CD, DVD et autres supports audio et vidéo ne représentent plus que 12% des ventes en 2003, contre 48% en 1999.

Grossistes et industriels privilégient l'image de marque

En offrant un nouveau débouché et un contact direct avec le particulier, Internet bouscule la stratégie commerciale des grossistes et des industriels. Internet leur permet d'atteindre le particulier pour lui vendre des biens sans passer par les détaillants. Pour 60 % de ces entreprises, la rentabilité du site Internet est déjà atteinte, son positionnement concurrentiel étant centré sur l'image de marque.

Le matériel informatique ou téléphonique représente 28 % du chiffre d'affaires sur Internet de ces nouveaux e-commerçants, devant les produits de beauté et de santé (17%) et les produits alimentaires boissons incluses (14%). Les priorités sont la relation-client, le marketing et la rationalisation des coûts. Ces entreprises, disposant de structures informatiques et logistiques plus importantes, intègrent plus fréquemment l'hébergement du site internet, le transport des marchandises et le paiement en ligne. À l'inverse, l'évolution et l'enrichissement du contenu du site sont souvent confiés à des spécialistes externes.

Voyages et réservations particulièrement adaptés à internet

Les ventes de services en ligne à destination des particuliers représentent une part importante du commerce électronique sur Internet. Les ventes de voyages et de réservations d'hôtels sont les plus importantes parmi ces ventes de services en ligne. Les agences de voyages et les entreprises de l'hôtellerie ont trouvé avec Internet un moyen efficace de faire connaître et de vendre leurs offres au plus grand nombre. La technologie multimédia d'Internet est particulièrement adaptée pour que le client

puisse choisir et réserver à distance la chambre d'hôtel ou le voyage qui lui convient le mieux, y compris en dernière minute. De son côté, l'entreprise peut mettre en ligne des offres au meilleur prix et en temps réel, notamment pour les réservations de dernière heure : la mise à jour des prix est quotidienne pour 84 % des sites.

Dans ce secteur, 80% des entreprises déclarent une activité de vente sur Internet déjà rentable en 2003. Sur les sites de ces entreprises, 11% des visiteurs effectuent une transaction, un taux nettement supérieur à celui des autres sites de vente. Les hôtels et agences de voyages présents sur Internet ont réalisé en 2003 près de 10% de leur chiffre d'affaires par ce media : 67% dans la vente de voyages et 31% dans la réservation de chambres d'hôtels. La faiblesse des coûts logistiques permet à ces e-commerçants de dépenser plus que les autres acteurs en publicité et marketing (32% des coûts liés à l'activité de l'entreprise sur Internet) et dans la relation avec le client.

Un canal supplémentaire pour les détaillants traditionnels

Les entreprises dites « click et magasins » ont développé une activité commerciale sur Internet en complément de leur réseau de points de vente physiques. Seuls 47% des e-commerçants détaillants déclarent avoir atteint le seuil de rentabilité de leur activité Internet : la rationalisation des coûts et le marketing constituent, avec la logistique, les priorités du moment. Leur positionnement concurrentiel, qui bénéficie déjà d'une enseigne connue, est centré avant tout sur la fidélisation de leur clientèle internaute et la fiabilité de leur site de vente.

Les secteurs les plus représentés sont les grandes surfaces alimentaires, le commerce de matériel informatique et téléphonique et le commerce de détail d'articles culturels et de loisirs. Les grandes enseignes du commerce spécialisé sont ainsi présentes sur le réseau des réseaux et offrent la commande en ligne.

La place des ventes de produits alimentaires (près d'un tiers de l'activité) illustre l'activité croissante des filiales Internet des grands groupes de la grande distribution alimentaire. À la différence des pure players qui sous-traitent la préparation de la commande, 98 % des e-commerçants intègrent cette fonction dans leur activité.

1.4 Internet dans l'entreprise

Le e-commerce est l'utilisation la plus visible d'Internet. Mais de nombreuses autres applications confèrent une nouvelle dimension aux entreprises.

1.4.1 Outil de communication externe

L'utilisation de la messagerie électronique permet d'établir une meilleure relation et à moindre coût avec les clients, les fournisseurs, les filiales ou une force de vente. La mise en place d'un serveur permet aux clients ou aux fournisseurs de venir chercher instantanément l'information qu'ils souhaitent (catalogue, prix, remises...). Le résultat en est la rapidité des transmissions et la réduction du coût d'impression et d'acheminement. La généralisation des boîtes aux lettres électroniques permettant de recevoir des informations en quantité et sous différentes formes est la source d'importantes économies. L'entreprise traitant avec des clients internationaux ou des filiales a également intérêt à mettre en place un système très économique de téléphonie par Internet. Aujourd'hui les entreprises peuvent recruter avec efficacité sur Internet. Il existe un nombre important de sites concernant l'emploi. Du moteur de recherche au site d'annonces, la quantité des candidatures disponibles, permet une sélection rapide, qualitative et internationale à un coût incomparable.

1.4.2 Outil de communication interne

Les entreprises qui possèdent un intranet sont beaucoup plus réactives et peuvent répondre plus efficacement aux clients. Internet comme l'intranet accélèrent les flux d'informations. Ils permettent donc une meilleure réactivité par rapport à la concurrence. La mise en place d'une base de données de documents commerciaux, marketing, comptables ou techniques permet d'optimiser la circulation de l'information. Les documents sont disponibles sur le réseau interne tout en filtrant le type d'information disponible, chacun disposant de la même source, évitant de faire circuler de l'information qui ne serait plus à jour.

1.4.3 Outil de Marketing

Pour une entreprise, la démarche de promotion d'un site web fait partie d'une démarche marketing. Les annuaires et les moteurs de recherche sont utilisés pour trouver des sites mais aussi des produits, donc des entreprises. En maîtrisant bien l'utilisation d'Internet, les outils de référencement et les moteurs de recherche, les entreprises peuvent trouver une cible consommatrice de proximité ou internationale. En diffusant des informations à forte valeur ajoutée sur leur serveur les entreprises captent l'attention d'une cible de clients potentiels. L'utilisation d'un formulaire électronique rempli par les visiteurs permet de les qualifier. L'entreprise accroît ainsi un fichier qui lui permet ensuite de déclencher des opérations marketing : lancement d'études de marché ciblées, diffusions d'informations complètes sur les produits, organisations de promotions personnalisées selon le profil et les achats des clients. L'utilisation du Push-média permet de systématiser ce type d'opération. D'autre part, il existe aujourd'hui des méga-bases qui permettent de faciliter le commerce one to one ou business to business. Le cybermarketing développe le concept du sur-mesure de masse. Les possibilités d'Internet permettent de proposer à chaque client un service ou produit conçu selon ses propres besoins ou préférences. Sur Internet, ces sur-mesure de masse peuvent concerner les quatre composantes du marketing mix. La valeur ajoutée de tels systèmes est très élevée et permet de fidéliser les clients.

1.4.4 Outil commercial

Internet est un support global grâce à la convergence des médias sur ce même support. De surcroît, il permet de supprimer les contraintes géographiques et de sous-tendre des stratégies de juste à temps ou de temps réel. Internet a donc une forte valeur ajoutée en matière commerciale. Il permet d'ouvrir les portes de l'exportation aux PME/PMI. Il facilite la mise en place d'une activité internationale par une disponibilité permanente de l'offre, un meilleur ajustement à la demande, une circulation plus rapide des produits et la réduction des stocks et des invendus. On note cependant que cette stratégie nécessite une logistique parfaite, c'est-à-dire des livraisons assurées en temps voulu. Internet permet d'élargir son marché et de

positionner l'entreprise sur le marché interentreprises et/ou grand public. C'est-à-dire le commerce en ligne avec le grand public, et d'autre part les relations commerciales avec les clients professionnels et les fournisseurs. En résumé Internet est un nouveau canal de distribution pour les entreprises.

1.4.5 Outil de service après vente

Internet grâce à son interactivité facilite les contacts avec les consommateurs. En participant aux groupes de discussion ou en multipliant les sondages les entreprises restent à l'écoute de leur marché. L'entreprise analyse en temps réel des informations qui lui permettront d'améliorer la qualité de ses services comme de ses produits. Il est également possible à l'entreprise d'optimiser son service après vente par la mise à disposition 24h/24 de listes de questions/réponses qui permettent aux clients de résoudre des problèmes courants ou de suivre par exemple le routage d'un colis.

1.4.6 Outil de veille stratégique

Avec Internet, l'entreprise renforce la connaissance de son environnement. Le réseau permet d'organiser une veille efficace sur le secteur d'activité de l'entreprise. A l'aide des outils d'Internet (robots et moteurs de recherches) il est devenu plus facile de localiser au niveau mondial plusieurs fournisseurs d'un produit ou d'un service, aussi pointu ou spécifique soit-il. On peut se faire ainsi une idée plus précise de l'offre en comparant les prix et la qualité des produits et négocier à l'échelle mondiale. Internet permet ainsi d'ajuster les prix. Il permet également de baisser les prix à la vente ou d'accroître les marges par la désintermédiation.

2. Le marché de la recherche d'information

Nous l'avons vu, de par sa construction, sur Internet, rien n'est défini quant à l'organisation des sites, de leurs pages ou de leur contenu. Il n'y a aucune organisation ou hiérarchie comme dans une bibliothèque par exemple.

Pour accéder à l'information que l'on recherche, noyée au sein de plusieurs milliards de pages, deux solutions différentes se sont développées : les annuaires et les moteurs de recherche.

Les annuaires, dont le plus connu est Yahoo, sont fondés sur un principe similaire aux pages jaunes : des personnes parcourent le web et classent les sites manuellement dans un système d'organisation hiérarchique de catégories.

Les moteurs de recherches sont eux mis à jour par des programmes informatiques, appelés robots, parcourant le web de liens en liens et indexant l'ensemble des sites dans une base de données.

Au niveau de l'utilisateur, la recherche sur un annuaire se fait en sélectionnant des catégories et sous catégories. Sur un moteur de recherche, l'internaute saisit un ou plusieurs mots-clefs, et le robot lui retourne les sites les plus pertinents qu'il a indexés.

Les moteurs de recherche sont donc beaucoup plus exhaustifs, mais la qualité des résultats est directement fonction de la pertinence du robot.

Si au départ, les annuaires ont connu un grand succès, les progrès techniques des moteurs de recherche ont peu à peu grignoté les parts de marché jusqu'à les faire quasiment disparaître aujourd'hui.

2.1 Altavista, le moteur quantitatif

Ce n'est pas le premier, mais certainement celui qui a rendu populaire cette technique de recherche. Nous sommes en train de parler d 'Altavista, le moteur de recherche né chez Digital, qui peut se vanter d'avoir été le premier à dépasser le cap des 100 millions de pages web recensées.

Le projet AltaVista est né pour démontrer la puissance du processeur Alpha, qui est le cœur des serveurs Digital, et de l'architecture cluster avec laquelle ils sont intégrés. La première version en bêta-test fut mise en ligne le 15 décembre 1995 et en moins d'une année le nombre d'utilisateurs passa de 500 000 à plus de 25 millions.

Ce succès est dû à son archive très complète et sa vitesse de réponse qui est estimée selon Digital à 0.7 seconde en moyenne pour chaque recherche.

La procédure d'enregistrement de pages Web dans l'archive d'AltaVista est également très sophistiquée. Son robot appelé Scooter, parcourt le Web et insère dans la base de données près de 10 millions de nouvelles pages par jour. Le moteur de "référencement" est capable d'examiner plus d'un milliard de caractères par heure, et peut également identifier la langue d'origine des pages (caractéristique unique des moteurs de recherche) permettant ainsi à l'utilisateur de mieux paramétrer sa recherche.

Dans le cadre d'une stratégie d'expansion, AltaVista a conclu d'importants accords avec des sociétés de télécommunications, afin de dupliquer le site dans d'autres pays. Le plus récent est le groupe espagnol Telefonica qui a mis sur pied le site en exclusivité pour le marché ibérique et pour le continent américain, excepté les Etats Unis et le Canada. Précédemment étaient déjà opérationnels d'autres sites "jumeaux" fruits d'accords avec Telstra en Australie, Alam Teknocrat en Malaisie et avec la suédoise TELIA ; cette dernière ayant permis la réalisation d'AltaVista avec une interface dans toutes les langues européennes, y compris le français, et cela

depuis septembre 1998. Néanmoins, c'est depuis mars 1999 que la version allemande a été activée gérant plus de 6 millions de documents en provenance de sites allemands, autrichiens et suisses.

C'est au cours de l'année 1997 que Digital a mis à disposition la technologie AltaVista dans un environnement local, afin que n'importe quel utilisateur puisse bénéficier de ses puissantes fonctions pour analyser les données sur ses propres ordinateurs. Bien évidemment le marché que Digital vise est celui des sociétés et de leur réseau informatique, pour lesquels il a mis au point des versions sophistiquées d'AltaVista capables de "tourner" soit sur Unix avec une architecture Digital Alpha, soit dans un environnement Windows NT sur Alpha et Intel. Une version personnelle pour l'environnement Windows 98 est également disponible.

C'est en fin 1997 qu'AltaVista intègre une fonction qui permet d'accéder à un site, tout en traduisant le contenu des pages dans la langue désirée. Cette technologie est réalisée par la société Systran et AltaVista l'a donc intégrée dans son moteur de recherche. Son fonctionnement est très simple : lorsque l'on effectue une recherche, près de chaque ligne qui affiche le résultat de la recherche, apparaît l'option "Translate"; il suffit de cliquer dessus, laisser AltaVista accéder au site qui nous intéresse, et traduire le contenu dans la langue que l'on a choisie et nous présenter la version traduite de la page, tout en gardant le même format et graphisme.

Parmi les différents moteurs de recherche, AltaVista était le seul n'ayant pas un annuaire de sites répertoriés par typologie, appelé communément directory. En janvier 1999 ce manque a été comblé, et sur la base d'un accord avec LookSmart, voilà que "Browse by Subject", une liste de 250 000 sites subdivisés sur la base de 16 000 arguments, a été créé. Durant la même période, suite à un accord avec la société américaine Switchboard, AltaVista a également activé les fonctions People Search (recherche de personnes par e-mail) et Business Search (recherche de sociétés et produits). Par ailleurs, grâce à la société iName, AltaVista a pu également offrir aux internautes la possibilité d'obtenir gratuitement une adresse e-mail.

Le début de l'année 1998 a été marqué par le rachat de Digital par la société Compaq. La stratégie adoptée par Compaq, en annonçant en janvier 1999 qu'allait naître une société indépendante appelée Altavista Company, était celle de donner une plus grande puissance au moteur de recherche et offrir plus de services, en changeant ainsi la direction que Digital avait prise, qui consistait à utiliser AltaVista comme un moyen de promotion et de divulgation de ses services adressés aux sociétés. En effet, Compaq a tout de suite créé une vaste gamme de nouveaux produits disponibles sur AltaVista. Outre au déjà mentionné service de traduction, examinons en détails certaines fonctions exclusives d'AltaVista.

En octobre 1998, les résultats des recherches se sont enrichis de l'utilisation du service AskJeeves lequel propose, en relation à la recherche effectuée par l'utilisateur, une série de réponses aux questions les plus souvent posées. On a par la suite introduit un contrôleur orthographique en quatre langues (dont le français) des termes introduits, en considérant qu'une recherche sur cinq contient une erreur, selon les estimations d'AltaVista. Toujours dans le domaine des recherches, AltaVista réussit dorénavant à interpréter automatiquement les termes qui peuvent être considérés comme une phrase unique, par exemple "San Francisco", évitant ainsi de devoir les insérer entre guillemets.

Mais début des années 2000 Altavista a été victime de son succès : le moteur a été dépassé par les millions de sites indexés dans sa base de données. Il commençait à être difficile de trouver son bonheur parmi les résultats de recherche.

2.2 Google, la recherche qualitative

Avant l'arrivée de Google, il était assez aléatoire de réussir à trouver réponses à sa recherche sur le Web. Google est arrivé à nous proposer un degré de pertinence très impressionnant et il facilite grandement les recherches, notamment pour les non-initiés. Il est acquis que Google est devenu plus qu'un moteur de recherche car il y a une vraie signification sociale derrière l'influence de ce moteur, allant jusqu'à devenir un verbe.

Tout d'abord, le terme "Google" est un jeu de mots sur "googol" qui a été inventé par Milton Sirota, neveu du mathématicien Edward Kasner. Un googol est le chiffre 1 suivi de 100 zéros. Il suggère la mission de Google qui est d'organiser l'immense quantité d'informations disponible sur le Web.

Les fondateurs de la société, Larry Page et Sergey Brin, firent connaissance à l'université de Stanford, U.S.A. Leur premier projet commun fût un outil de recherche appelé le BackRub. Vous pouvez consulter le résultat de cette recherche sur le site de l'Université de Stanford. Cet outil se basait sur les liens qui pointaient vers un site pour évaluer sa pertinence et sa popularité. Déjà, le marketing viral contribua au succès de BackRub auprès des étudiants, en utilisant le bouche à oreille comme moyen de se faire connaître. Il y a aussi les photos des débuts de Google, qui sont éloquentes.

A l'époque, il existait d'autres moteurs de recherche qui étaient capables de scanner une grande portion du Web pour construire un index et trier les pages qui correspondaient à certains mots-clefs. En revanche, les résultats affichés pour une recherche n'étaient pas très pertinents. Il fallait souvent tâtonner pour trouver la bonne information que l'on recherchait.

La réussite de Google tient principalement au fait que Larry Page et Sergei Brin ont trouvé le moyen de déterminer les pages qui étaient les plus pertinentes par rapport

aux critères de recherche. C'est un algorithme mathématique appelé le PageRank qui est à la base du succès de Google, car il arrive à donner un degré d'importance aux pages Web très efficace. C'est le PageRank qui distingua Google de ses concurrents dès le départ.

Quand le robot de Google (GoogleBot) visite une page Web, il analyse la structure de la page et suit les liens vers les autres pages Web. Ainsi, Google a une excellente vue d'ensemble du réseau et trouve quelles sont les pages qui sont liées entre elles. Surtout, il évalue le degré de pertinence d'une page en fonction du nombre d'autres pages qui lui font un lien. De plus, les liens qui viennent de pages qui ont une grande importance, parce qu'un grand nombre de pages font pointer un lien vers elles sont très bien perçus. Il y a une circularité évidente à ce schéma et Google a trouvé le moyen d'évaluer tout cela pour donner une note PageRank de 0 à 10 à la page évaluée.

Jusqu'au milieu de l'année 1998, Larry Page et Sergey Brin continuèrent de travailler sur leur moteur de recherche. Ensuite, la recherche de partenaires commença. Les principaux acteurs du Web ne montrèrent que peu d'intérêt pour Google. Leur ami et fondateur de Yahoo, David Filo, comprit l'intérêt de Google, mais il leur conseilla de démarrer leur propre société d'outil de recherche au lieu de chercher à se joindre à un portail existant.

Un des fondateurs de Sun Microsystems, Andy Bechtolsheim, leur donna un gros coup de pouce en investissant 100 000 \$. Finalement, grâce aux amis et la famille, les deux compères réunirent un million de dollars pour démarrer Google Inc. qui s'installa à Menlo Park, Californie en Septembre 1998. Craig Silverstein fût le premier employé de Google Inc., en tant que Directeur de la technologie.

Dès le départ, Google répondait à 10 000 recherches par jour. Au début de l'année 1999, Google répondait à 500 000 recherches par jour, mais surtout prenait son essor au sein des plus gros portails du Web, comme Yahoo, AOL et Netscape. Dès le milieu de l'année 1999, Google gérait 3 millions de recherches par jour et des millions de dollars de revenus.

Le déménagement dans le local actuel, Googleplex, pris place à cette période et se singularise par des innovations au niveau de l'espace de travail et une ambiance informelle.

En 2000, Google devint le plus gros outil de recherche au monde, avec un index de 1 milliard de pages au début de l'année. Depuis le début, Google fait un gros effort pour sortir des innovations marketing et techniques, comme les Adwords, Google Image Search, ou la Google Toolbar.

A la fin de l'année 2000, Google se voyait le témoin de 100 millions de requêtes par jour. Depuis, Google n'a de cesse de progresser et il domine actuellement le domaine de la recherche sur Internet outrageusement.

Au début d'Internet, il était de mauvais goût de penser à la manière de gagner de l'argent. Les revenus initiaux de Google provenaient de la vente de bannières et des licences vendues aux autres moteurs de recherche. Le tournant arriva en 2000, quand les Adwords furent lancés. Ce système de liens sponsorisés basé sur des enchères de mots-clefs apparaît à côté des résultats de recherche naturels. Désormais, Google est en bourse (GOOG) et les impératifs semblent changer pour la société Google Inc., mais les concurrents semblent vouloir affronter Google sur tous les fronts (MSN, Yahoo, Amazon et d'autres à venir). Le marché de la recherche d'information n'est pas encore consolidé et les avantages techniques et marketing de Google ne seront pas de reste pour affronter la concurrence.

2.3 Yahoo, de l'annuaire au moteur

Comme nous l'avons évoqué, Yahoo n'était au départ qu'un simple annuaire. La qualité de son classement lui a permis de devenir l'un des premiers sites Internet mondial. Pour enrichir son annuaire, Yahoo a choisi d'ajouter en 1999 un moteur de recherche, fourni par Google.

Mais au fil des mois, les internautes ont délaissé la partie annuaire pour la partie moteur de recherche, puis progressivement se sont directement rendus sur Google, sans passer par Yahoo.

Conscient de ce problème, et devant la perte régulière de parts de marché, Yahoo a pris la décision de développer son propre engin de recherche puis revenir dans la course de la recherche d'information.

Ainsi en mars 2003, Yahoo rachète la société Inktomi puis Overture en juillet de la même année.

Inktomi était notamment le pourvoyeur discret des recherches de MSN.

Quant à Overture, star des technologies publicitaires dites de « positionnement payant », il venait d'acheter les deux moteurs de recherche Altavista et All The Web. Yahoo se trouvait donc à la tête d'un énorme potentiel capable de concurrencer Google.

Durant plus d'un an, le doute a plané concernant le lancement du nouveau moteur de recherche : serait-ce l'intégration d'All The Web dans Yahoo ou bien de l'ancien leader Altavista ?

La réponse fut donnée le 18 février 2004, date à laquelle Yahoo annonce le lancement de son Yahoo Search Technology, vraisemblablement, un nouveau moteur, basé sur la meilleur des trois technologies acquises.

2.4 Microsoft, le géant du logiciel

Entreprise fondée aux Etats-Unis en avril 1975 par deux informaticiens, Bill Gates et Paul Allen, Microsoft est aujourd'hui le premier éditeur mondial de logiciels. Passé en moins de trente ans de 2 à plus de 56 000 employés, le chiffre d'affaires de Microsoft s'élevait en 2003 à 32.2 milliards de dollars, soit une augmentation de 13.5% par rapport à 2002, pour un bénéfice net de 9.9 milliards de dollars. L'entreprise compte aujourd'hui 70 filiales présentes dans le monde entier et ses logiciels sont disponibles dans plus de 30 langues différentes.

Son logiciel phare, le système d'exploitation *Microsoft Windows* équipe en 2004 plus de 90% des ordinateurs vendus dans le monde.

Il fut un temps où Microsoft comptait faire de son MS-Network (MSN) une alternative à Internet et en même temps un moyen pour dépasser les leaders des services online comme America On Line (AOL) et CompuServe. A quelques années de distance, suite à l'échec commercial de ce service, Microsoft présente à nouveau

MSN sous un autre jour, en suivant la tendance des services sur Internet, en tant que portail online.

Voici donc MSN Search qui plus qu'à être un moteur de recherche est un catalogue de liens vers les services de Microsoft comme par exemple Expedia ou Carpoint. En arrivant sur MSN on trouve bien évidemment la barre classique pour insérer les mots désirés pour une recherche, mais l'objectif est finalement celui de raccorder les diverses informations mises en ligne par Microsoft et ainsi mettre un peu d'ordre parmi les différentes facettes du projet inachevé du premier MSN.

Le projet initial, dont on commença à parler déjà au début de l'année 1997, prévoyait la naissance d'un moteur de recherche classique, basé sur une technologie entièrement développée par Microsoft ; un nom codifié du projet fut également choisi : Yukon. Malgré tout, c'est en novembre 1997 que l'on annonça que le moteur de recherche aurait utilisé la technologie Inktomi, déjà utilisée par HotBot et plus récemment par Snap et GoTo.

Il aura fallu 10 mois après l'annonce officielle avant que MSN Search voie le jour, en débutant avec une version beta en septembre 1998 qui s'agrandit durant le mois d'octobre.

Lorsqu'un internaute utilise le moteur de recherche Microsoft, la requête formulée est transmise à Inktomi Search Engine qui effectue un travail de classification des sites et propose les résultats correspondant à la requête formulée. Cette classification est transmise au moteur de recherche qui présente les résultats obtenus sur son interface. Les résultats affichés par le moteur de recherche Microsoft sont classés par Inktomi.

Ce n'est que quelques années plus tard, probablement suite au rachat d'Inktomi par Yahoo, que l'idée initiale de développer une technologie moteur propre à Microsoft refait surface.

La technologie Inktomi est utilisée par Microsoft jusqu'en novembre 2004, date à laquelle la société lance en France un bêta test du moteur de recherche MSN search. Désormais MSN Search s'appuie sur une technologie de recherche développée en interne. MSN Search est doté d'un nouvel algorithme de recherche et d'un index initial comptant plus de 5 milliards d'URL. En ce qui concerne l'interface de l'outil de recherche, Microsoft opte pour la sobriété au même titre que son principal concurrent : Google.

Mais ce n'est que quelques mois plus tard que MSN dit officiellement adieu à Yahoo. Le 1^{er} février, le PDG de Microsoft en personne, Bill Gates, annonce le lancement officiel dans le monde (6 langues, 24 pays) de l'intégration de sa technologie "maison" au sein de MSN.

Pour accompagner le lancement de MSN Search, Microsoft a prévu une vaste campagne publicitaire à travers le monde. S'il ne souhaite pas préciser les montants investis, les ambitions sont, elles, clairement affichées. La campagne doit toucher 90% des Américains, Britanniques et Japonais, avec notamment des spots TV diffusés lors du Super Bowl, de la remise des Oscars ou des Grammy Awards.

3. Les nouveaux champs d'application de la recherche

Si jusqu'à présent les moteurs de recherche se cantonnaient à trier les milliards de pages présentes sur Internet, aujourd'hui, de nouveaux champs d'application, bien plus larges, s'ouvrent à eux.

3.1 Le commerce électronique

En décembre 2002, Google lance en version bêta, *Froogle*, « un hypermarché mondial en ligne ». *Froogle* permet de localiser un produit commercialisé sur le web. On peut effectuer une recherche par mots-clefs ou en naviguant parmi la quinzaine de rubriques thématiques (Vêtements, Automobile, Informatique, Electronique...)

La recherche avancée permet d'affiner la requête sur le nom du produit, les occurrences, la catégorie et d'indiquer un intervalle de prix.

Froogle permet à Google de mettre un pied dans le marché lucratif du commerce en ligne. Il se place également en concurrence directe avec les comparateurs de prix.

Quand à MSN et Yahoo, ils choisissent plus simplement d'intégrer à leur site *Kelkoo*, le comparateur de prix leader en France.

Mais devant le succès grandissant de *Froogle*, Yahoo réagit et annonce en juillet 2003 le lancement de « Yahoo! Product Search ».

Si pour le moment Google ne demande aucune participation financière aux sites marchands présents sur son moteur, Yahoo, lui, ne propose que des sites partenaires ou financièrement liés à Yahoo Inc.

3.2 L'actualité

Une des faiblesses des moteurs de recherche résidait dans leur délai de mise à jour. En effet, l'indexation des milliards de pages présentes sur Internet nécessite beaucoup de temps. En 2004, le délai moyen de mise à jour de l'index était d'un mois. Cela signifie qu'un site traitant d'un fait d'actualité risquait de ne pas être visible sur Google avant un mois. Réduire ce délai pour l'ensemble des 100 millions de sites

présents dans la base de données des moteurs aurait nécessité un accroissement de leurs ressources informatiques colossal.

La solution a été trouvée par Google qui en Mars 2002 a lancé un nouveau service « Google News ». Google est parti du postula que certains sites sont peu mis à jour alors que d'autres, les sites d'actualités ont de nouveaux articles toutes les minutes. Google a donc sélectionné les sites d'informations les plus populaires et les a inclus dans sa section « Google News ». La mise à jour de ces sites est donc quasi permanente et permet à Google de fournir des résultats pertinents pour les recherches touchant à des faits d'actualité.

Mais Google n'en reste pas là et va profiter de son outil pour rattraper son retard sur le contenu éditorial. En effet, depuis toujours Yahoo met à disposition une section « Actualité » qui propose, comme un journal, une sélection d'articles. Ce travail est fait manuellement par des responsables éditoriaux.

Mais si on pouvait résumer la philosophie de Google en un mot, ce serait probablement « automatisation ». Effectivement, aucun des services qu'il propose à l'heure actuelle ne nécessite une intervention humaine. Tout est géré par des programmes informatiques complexes. Google a donc repris son principe d'automatisation et a développé des logiciels capables de classer et regrouper des articles en catégories. La pertinence de ce classement est proche du travail réalisé par les équipes de Yahoo.

MSN ne restera pas longtemps à l'écart de cette guerre et annonce en juillet 2004 le lancement de son propre moteur d'actualités nommé Newsbot.

Lors de cette annonce, la firme de Bill Gates prétend aller plus loin que les sites déjà existants pour proposer aux utilisateurs une information qui prendra en compte les articles qu'ils ont déjà lus. Ceci devrait permettre de ne fournir que l'information qui les intéresse. Newsbot se démarquerait ainsi de Google News qui regroupe les informations par thème.

Les répliques de Google ne se font pas attendre puisqu'en mars 2005, Google propose une nouvelle option sur son service d'actualités permettant de sélectionner uniquement les articles proches de nos centres d'intérêts.

Quand à Yahoo, un communiqué de presse a annoncé récemment la préparation d'une nouvelle version de son portail d'actualités.

MSN News

Web Actualités Images Sports

paramètres aide

Actualités

msn Search - France

Page d'accueil - Le top 10 - Monde - France - Économie/Bourse - Sports - Divertissements - Sciences/Santé - Technologies

Page d'accueil de Newsbot

Le CAC 40 retombe vers les 4.000 points

Les Echos - Il y a 1 heure

Le CAC 40 s'effrite de 0,29%, à 4 007,71 points. Une consolidation prévisible après cinq séances de gain. L'indice s'est notamment apprécié de 0,6%, hier, clôturant, pour la première fois...

Articles similaires (4)

Le CAC 40 finit en hausse de 0,80% à 4 019,35 points

La Tribune

Bourse de Paris: le CAC 40 a de nouveau franchi la barre des 4 000 points

La Tribune

Hausse de +0,58% à 4 011,14 pts pour le CAC 40

Boursorama

George Bush se félicite de cette 'victoire majeure'

Le Monde - Il y a 11 heures



LE MONDE 05.05.05 14h11

Washington de notre envoyée spéciale es Américains se sont remis à espérer qu'Oussama Ben Laden soit à leur portée. Mercredi soir 4 mai, les chaînes de télévision...

Reuters/ISTRAKISTAN

Édition et traduction simultanées (E)

Tony Blair remporte un 3eme mandat historique avec la victoire de son parti travailliste aux legislatives jeudi

France 2 - Il y a 5 heures



M. Blair a dit vouloir 'répondre avec sagesse' aux attentes des électeurs qui ont donné aux travaillistes, selon les sondages, de 66 à 80 sièges de majorité dans la nouvelle assemblée...

Photos et articles similaires (9)

Reuters/ RUSSELL BOYCE

QUELLE EUROPE VOULONS-NOUS ? Une Union des peuples et des personnes

Par Christine Boutin

Le Figaro - Il y a 7 heures

Référendum 2005 International Les forces de sécurité irakiennes, cibles des attaques Sharon a participé à la Marche des vivants 20 000 personnes pour une Marche des Vivants en hommage aux...

Yahoo News

Yahoo! My Yahoo! Mail

Search the Web Search

YAHOO! NEWS Sign In New User? Sign Up News Home - Help

Tell us what you think about the new Yahoo! News! [Send Feedback](#) | [Discover More](#)

News Home U.S. Business World Entertainment Sports Tech Politics Science Health Most Popular Index

Photos Opinion Local News Sub News Comics Weather Full Coverage Audio/Video

Search: All News Search Advanced

Blair's Labour Party Wins Re-Election

AP - 42 minutes ago

LONDON - Tony Blair, elected to a historic third term as Britain's prime minister, acknowledged Friday that the Iraq war was "a deeply divisive issue" that hurt his Labour Party, but he said Britons were ready to "move on" and focus on the future.

Slideshow: British Elections

Video: Exit Polls: Blair Wins Historic Third Term

THE NEW YAHOO! NEWS

We've redesigned Yahoo! News to make it easier for you to find the news you want from a wider range of sources.

Discover More

Personalize News Home Page: Add/Remove News Categories | Change Layout | Weather

Top Stories

AP Reuters | AFP | Los Angeles Times | USATODAY.com | CSMonitor.com | My Sources

- Car Bomb Kills 14 South of Baghdad AP - 18 minutes ago
- Army Demotes General in Abu Ghraib Scandal AP - 18 minutes ago
- Employers Boost Hiring, Jobless Rate Steady AP - 32 minutes ago
- North Korea May Test Nuclear Device AP - 1 hour, 26 minutes ago
- White House Tries to Allay Russian Anger AP - 2 hours, 58 minutes ago

All Top Stories from AP

FULL COVERAGE

- Iraq
- Pakistan
- U.S. Military
- Space Shuttle
- Federal Reserve

All Full Coverage

PHOTO HIGHLIGHT



Google News

Google Actualités standard | Version texte

Web Images Groupes Annuaire Actualités

Recherche Actualités Rechercher sur le Web

Actualités France BÉTA 500 sources d'information mises à jour en continu

À la une

International

France

Economie

Science/Tech

Sports

Culture

Santé

Adopter Google Actualités comme page de démarrage

Alertes Actualités

À propos de Google Actualités

Jean-Pierre Raffarin cherche à construire son avenir politique

Le Monde - Publié depuis 1 heure

POUR son troisième anniversaire à Matignon, vendredi 6 mai, Jean-Pierre Raffarin préfère songer à l'avenir. A son avenir. Il ne devait donc pas y avoir de cérémonie particulière, vendredi, pas de dossier...

Raffarin à Matignon: trois bougies et bouge de là ? Libération

Ce qu'il dit des politiciens Le Figaro

Nouvel Observateur - Boursier.com - Ici.Camas - Le Point - et 35 articles connexes >>

Le Fatah remporte les municipales

TF1 - Il y a 7 heures

"Le Fatah a obtenu 59,9% des suffrages exprimés, contre 33,3% au Hamas", a annoncé vendredi matin aux journalistes Firas Iari, directeur général du Comité supérieur pour les élections des collectivités locales...

Le Hamas dénonce la corruption du Fatah Le Figaro

Le Fatah résiste au Hamas lors des municipales palestiniennes Angolapress

Nouvel Observateur - Libération - Le Monde - Stop Infos - et 141 articles connexes >>

Gérez des gâchetiers de la RATP

Nouvel Observateur - et 9 articles connexes >>

Virus: Sotomayor vous invite à la Coupe de Monde 2006

Futura Sciences - et 19 articles connexes >>

RAR Honda renonce à en appeler de sa suspension

Canal - et 191 articles connexes >>

Rob Dylan se livre en librairie

Adoccy - et 18 articles connexes >>

Napoléon, mort du cancer

Swissinfo - et 17 articles connexes >>

Dans l'actualité

- Tony Blair
- Simone Veil
- Florence Aubenas
- Buenos Aires
- Milan AC
- Britney Spears
- Max Mosley
- Black Hawk Down
- Ligue des Champions
- Michael Schumacher

International >> modifier France >> modifier

3.3 La culture

3.3.1 Livres

En collaboration avec de grandes institutions américaines et britanniques, les fondateurs du moteur de recherche lancent un programme de numérisation de millions d'ouvrages qui seront accessibles sur le Web.

Six ans après la création de Google, les cofondateurs de la firme californienne, Sergey Brin et Larry Page, annoncent, fin 2004, le lancement d'un programme de numérisation de millions d'ouvrages. Ceux-ci seront accessibles gratuitement aux internautes du monde entier. Ce projet, Google Print, peut déjà compter sur la collaboration des bibliothèques de trois universités américaines, Harvard, Michigan et Stanford, de la New York Public Library et de la Bodleian Library, rattachée à l'université d'Oxford, au Royaume-Uni.

Cette dernière débute son programme avec la numérisation de son fonds du XIXe siècle : un million et demi d'ouvrages tombés dans le domaine public, dont de nombreux titres inaccessibles au grand public. L'opération devrait être bouclée en trois ans. Les premiers livres seront affichés sur la Toile dès juin 2005 et Oxford envisage déjà de poursuivre l'opération avec d'autres fonds.

Il faudra en revanche six années à la bibliothèque de l'université du Michigan pour numériser l'ensemble de son fonds, sept millions d'ouvrages. Et à peu près autant pour celle de Stanford (huit millions de volumes). La bibliothèque d'Harvard se contente de scanner dans un premier temps quarante mille de ses quinze millions de volumes, sélectionnés au hasard. Même rythme à la New York Public Library (dix-

sept millions de volumes) qui devrait numériser, en six mois, cent vingt mille titres tombés dans le domaine public

Depuis n'importe quel ordinateur, les internautes pourront consulter intégralement ces ouvrages, en mode texte. En revanche, ils n'auront pas la possibilité de les télécharger, de les sauvegarder ni de les imprimer depuis le site Google, qui a désactivé les fonctions copie et impression. La numérisation des pages en mode texte est une opération longue et onéreuse. L'opérateur de la Silicon Valley affirme avoir conçu et développé une robotique portable et rapide pour scanner les textes. Mais il se refuse à décrire cette technologie : "Un secret commercial !" Sauf pour préciser que le travail de numérisation sera effectué sur place, dans chacune des bibliothèques, et que la manipulation des livres ne les endommagera pas plus qu'une consultation par un lecteur. Google conservera une copie digitale de chaque œuvre et en remettra une seconde à ses partenaires, qui pourront l'utiliser comme bon leur semble, y compris autoriser le téléchargement et l'impression de certains textes à partir de leur propre site Web.

La plupart des bibliothèques impliquées avaient déjà lancé des programmes de numérisation de leurs collections, mais avec une technologie moins perfectionnée et, surtout, à une échelle plus faible, faute de moyens. Or la société Google, introduite en Bourse avec succès en août 2004, dispose des fonds nécessaires pour financer cette opération. La firme réglera la totalité de l'addition, mais refuse de préciser ses coûts, estimés de 150 à 200 millions de dollars.

Comme l'accès à cette bibliothèque virtuelle est entièrement libre pour les lecteurs-internautes, on se demande ce qui pousse Google dans cette aventure. D'autant plus que la firme californienne, qui vit de ses revenus publicitaires, n'envisage pas d'introduire des fenêtres de publicité sur les pages des livres - tout en n'excluant pas des accords commerciaux à venir avec des éditeurs, des auteurs, des libraires ou des bibliothèques. "Nous gagnons de l'argent en améliorant notre moteur de recherche et en augmentant la masse et la qualité d'informations accessibles pour nos usagers, explique Suzan Wojcicki, qui pilote ce programme au siège de

Mountain View (Californie). Nous sommes passés d'un index d'un milliard de pages à huit milliards aujourd'hui."

Les bibliothèques d'Europe, d'Amérique latine ou d'Asie seront-elles sollicitées pour participer à cette aventure ?

Avec ce projet, Google affirme son avance technologique. Aucun projet similaire n'est officiellement prévu chez les deux grands rivaux.

3.3.2 Encyclopédies

Dans ce domaine, c'est Microsoft qui détient une longueur d'avance.

En effet, en 1993, Microsoft annonce la sortie de Microsoft Encarta, la première encyclopédie multimédia pour ordinateur personnel. Elle comprend des milliers d'articles assortis d'images, de séquences sonores et animées, d'illustrations, de graphiques, de photographies, un atlas, une ligne du temps interactive, le tout sur un seul CD-ROM.

Au fil des années, Microsoft a amélioré son encyclopédie, en l'enrichissant de nombreux documentaires. Aujourd'hui, Encarta est l'encyclopédie la plus vendue dans le monde.

Forte de ce succès, la firme de Redmond a progressivement mis en ligne son encyclopédie sur le site MSN, puis l'a intégrée dans son moteur de recherche.

Pour palier à cette absence de contenu, c'est en *Wikipedia* que Google et Yahoo vont trouver leur salut.

Wikipédia est un projet encyclopédique associatif dont tout visiteur peut modifier les pages à volonté. Il permet non seulement de communiquer et diffuser des informations rapidement, mais de structurer cette information pour permettre d'y naviguer commodément. Ce projet a commencé le 15 janvier 2001 en anglais. Au mois de mai 2005, l'ensemble du projet comportait environ 1 300 000 articles en plus de 100 langues différentes.

Cette encyclopédie étant gérée par une association, les deux géants de la recherche vont utiliser un moyen subtil pour s'attirer la sympathie des fondateurs de *Wikipédia*.

En effet, le 11 février 2005, Google propose d'aider la Fondation *Wikipedia*. Une opération séduction qui permettrait à *Wikipedia* de profiter des serveurs et de la bande passante maison. En échange de quoi, même si les termes de cette proposition sont encore flous, les articles de l'encyclopédie libre pourraient servir à alimenter le moteur Google.

Moins de deux mois après, Yahoo et la Fondation *Wikipedia* annoncent officiellement la conclusion d'un accord qui servira les deux parties. Au terme de ce rapprochement, Yahoo! va fournir une capacité d'hébergement « significative » à Wikipedia. On n'en connaît pas les détails techniques, mais le gîte sera ouvert dans des serveurs situés dans un centre en Asie. De son côté, Yahoo! va expérimenter la mise à disposition du contenu des articles de l'encyclopédie libre dans ses raccourcis en français.

Si l'initiative est concluante, elle sera alors étendue à d'autres langues en Asie, Amérique latine, Europe et Etats-Unis.

3.4 Vie privée

Nous l'avons vu, les moteurs de recherche ne se contentent plus de fournir un accès rapide à l'information présente sur les sites web, mais nous permettent aujourd'hui d'accéder à la culture et à l'actualité.

Pour se rendre indispensables, les moteurs se sont lancés en 2004 dans une nouvelle guerre : la recherche d'informations dans nos données personnelles.

C'est Google qui a commencé en lançant en octobre 2004 Desktop Search. Ce logiciel installé sur le PC de l'utilisateur lui permet de chercher parmi ses documents de la même façon qu'il recherche sur le web avec Google.

Google Desktop Search amène la puissance de Google à nos informations personnelles, sur notre propre ordinateur. Les utilisateurs peuvent ainsi rechercher à travers un large éventail de fichiers et d'information telles que des e-mails dans Microsoft Outlook et Outlook Express; des fichiers dans Microsoft Word, Microsoft Excel, Microsoft PowerPoint, et textes; l'historique des sites web dans Internet Explorer; et des conversations dans la messagerie instantanée d'AOL.

Une fois de plus, Google ne reste pas longtemps seul sur ce créneau. En novembre 2004 MSN lance MSN Desktop Search à grand renfort de communication. Moins de deux mois après, Yahoo ferme la marche en annonçant Yahoo Desktop Search.

En mai 2005, les moteurs de recherche se sont imposés comme une véritable porte incontournable pour accéder à tous types d'informations.

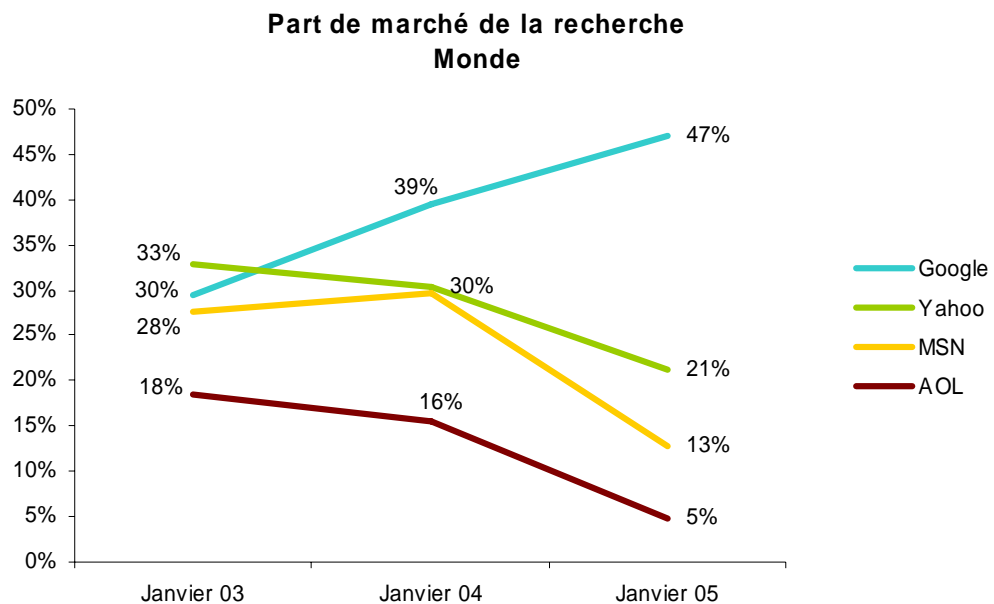
Que ce soit pour rechercher un site web, consulter l'actualité, lire un livre, trouver une définition dans l'encyclopédie ou encore rechercher un document texte sur son ordinateur.

Cette large palette de possibilités va encore s'enrichir de fonctionnalités, notamment avec l'évolution des périphériques mobiles.

4. La guerre pour un monopole

4.1 Parts de marché mondiale

C'est l'avance technologique (la technologie du Page Rank) qui a permis à Google d'acquérir régulièrement des parts de marchés depuis sa création en 1999 aux dépens de Yahoo et MSN. En Mars 2005, Google concentrait 47% des recherches, contre 21% pour Yahoo et 14% pour MSN.



Source : Nielsen

Cependant, depuis le début de l'année le taux de croissance de Google se réduit peu à peu. En effet, ses deux principaux rivaux ont quasiment rattrapé la majorité de leur retard et fournissent aujourd'hui des résultats de recherche d'un niveau de qualité sensiblement équivalent à ceux de Google.

La guerre des innovations s'est donc déplacée aujourd'hui sur les services annexes. En ce qui concerne la recherche, nous avons vu les nouveaux défis de la recherche d'informations personnelles, dans l'actualité et dans le domaine culturel.

Mais c'est également dans le domaine de la communication et plus particulièrement dans la fourniture d'adresses email et de messagerie instantanée que la guerre s'est intensifiée ces derniers mois. En effet, fournir une boîte aux lettres aux utilisateurs est un moyen de les attirer une fois de plus sur leur portail.

Dans ce domaine, Yahoo et MSN ont une avance indéniable. MSN d'abord avec son service de mail Hotmail, racheté en 1997, et son logiciel de messagerie instantanée MSN Messenger, en font le leader de la communication.

Mais Yahoo est également en bonne place avec son portail « My Yahoo » qui fournit un très grand nombre de services à ses utilisateurs.

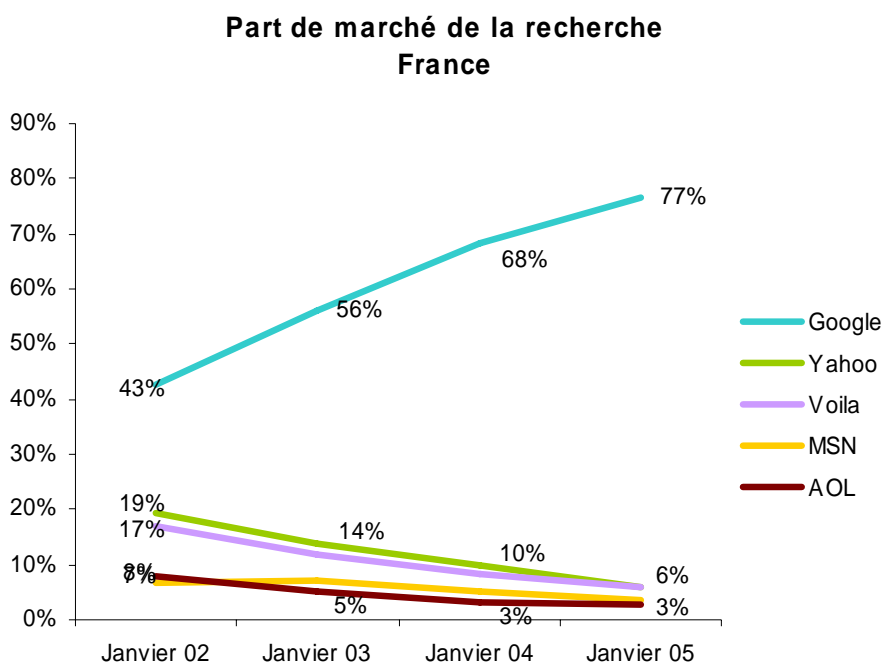
Google, très en retard sur ce secteur, n'avait qu'une solution pour pousser les utilisateurs de Yahoo et MSN à se convertir, proposer un service réellement innovant. Le pari sera réussi à travers le service Gmail : Un giga octet d'espace de stockage, là où les concurrents ne dépassent 10 Mo. Basé exclusivement sur du marketing viral, le plan de communication du lancement de Gmail a été également une réussite : obtenir un compte Gmail, équivaut à entrer dans une soirée privée et branchée : il faut se faire inviter. Google ne l'a pas encore lancé officiellement, mais l'a d'abord offert à quelques privilégiés. Puis il a permis à ces "élus" d'inviter leurs amis à s'inscrire à leur tour. Chaque utilisateur a d'abord eu droit à deux invitations, puis à six nouvelles quelques semaines plus tard. Et le bouche à oreille a fonctionné : ceux qui avaient un compte Gmail ont chanté les louanges du service sur de nombreux weblogs, et de plus en plus d'internautes ont désiré en avoir un. Même si le nombre d'utilisateur de Gmail est encore très loin de celui de MSN et Yahoo, Google comble petit à petit son retard.

4.2 L'exception française

En France, la situation est très différente. Nous sommes face à un quasi monopole de Google qui culmine au mois d'avril à près de 78% des recherches. Cette différence peut s'expliquer par plusieurs facteurs. Tout d'abord les français ont accédé massivement à Internet bien plus tard que dans les principaux pays développés. L'arrivée d'Internet dans les foyers et entreprises français se situe aux débuts des années 2000, date à laquelle Google était déjà bien implantée. En France, Yahoo et MSN n'ont pas eu la prime au premier entrant comme ce fut le cas aux USA notamment.

Ensuite, certains fournisseurs d'accès à Internet, point de démarrage de la plupart des internautes, ont fait le choix de proposer Google comme moteur de recherche par défaut. Par exemple Free, deuxième fournisseur d'accès français a signé un accord avec Google pour lui donner une visibilité très importante sur son portail free.fr. Cela contribue à la notoriété de Google.

Ces deux facteurs ont permis à Google d'acquérir en moins de deux ans près de 80% des parts de marché.



Source : Xiti

4.3 Vers un monopôle ?

Nous l'avons vu, les performances des trois moteurs sont aujourd'hui sensiblement équivalentes. La bataille va donc probablement se poursuivre à un autre niveau : la communication et l'intégration.

Sur ces deux points, Microsoft possède des atouts indéniables.

Tout d'abord, au niveau de la communication, la firme de Redmond possède une capacité financière sans équivalent. Cela lui permet de mettre en place des campagnes de publicités d'une ampleur phénoménale. Revenons sur le lancement de MSN Search. Microsoft s'est offert l'événement planétaire le plus diffusé à la télévision : la finale du championnat de football américain, le SuperBowl. Le spot de 30 secondes y est facturé près 2 millions d'euros.

Pour accroître sa puissance financière, alors qu'il était jusqu'à présent entièrement autonome, Google a dû s'introduire en bourse en 2004. Cet apport de trésorerie lui permettra peut-être de communiquer, lui qui n'a encore jamais fait de campagne de publicité, se basant uniquement sur le bouche à oreille et le marketing viral.

Deuxième avantage considérable pour Microsoft : l'intégration.

En effet, près de 90% des ordinateurs sur la planète sont équipés du système d'exploitation de Microsoft : Windows. Cette situation monopolistique a déjà permis au géant de Redmond de remporter des guerres de parts de marché, notamment celle des navigateurs web, logiciel incontournable pour accéder à Internet.

Il est intéressant de revenir sur cette histoire qui pourrait se reproduire sur le marché de la recherche.

Au milieu de l'année 95, le navigateur de Netscape régnait en maître, car il n'avait pas de concurrent sérieux. Ses parts de marché de l'époque tournaient autour de 80%. Microsoft prend alors conscience de l'importance du Web et crée son propre navigateur, Internet Explorer. La première version fut distribuée gratuitement avec Windows 95. Netscape Navigator coûtait alors 39,95 \$.

Ainsi, dès 1996, les parts de marché du navigateur de Netscape commencent à décroître, au rythme de 1% par mois environ. Selon la justice américaine, Microsoft

abuse alors de sa situation de monopole dans le domaine du système d'exploitation pour imposer son navigateur aux constructeurs de micro-ordinateurs. Mais il faut également souligner que, à partir de la version 4, Internet Explorer l'emporte sur le plan technique, et que cette avance ne se démentira plus par la suite. Sa première grande innovation est d'avoir permis un accès aux services en ligne MSN directement depuis Windows, l'OS prenant toutes les opérations en charge. Avec la mutation des services propriétaires en World Wide Web, l'OS le plus répandu sur le marché devient également le moyen le plus simple pour accéder à Internet. Des liens vers Internet sont désormais présents sur toutes les applications Microsoft, et Internet Explorer 4 est complètement intégré à Windows.

Au début de l'année 2000, les parts de marché mondiales du navigateur de Netscape étaient de 18%. Au milieu de l'année 2001, elle descend vers les 12%, mais elle était stabilisée. Cette dernière baisse coïncide avec l'apparition de la version 6 du navigateur de Microsoft. La chute va alors s'accélérer de manière exponentielle. Une étude sur l'année 2001 crédite Navigator de 4,5% de parts de marché, contre 94,2% à Explorer.

Aujourd'hui, même si de nouveaux acteurs se sont lancés dans la bataille, Internet Explorer domine largement le secteur des navigateurs Internet avec près de 90% des parts de marché.

Ce nouveau monopole n'est pas sans lien avec le marché de la recherche. Effectivement, Microsoft utilise son navigateur pour favoriser l'utilisation de son moteur de recherche : un bouton intégré dans le logiciel permet de lancer une recherche sur MSN Search en un clic. Cette intégration devrait encore s'intensifier avec la prochaine version du navigateur. Microsoft risque de mettre encore plus en avant son outil de recherche maison, au grand damne de Google et Yahoo.

Nous l'avons vu, Microsoft possède de nombreux atouts pour étendre son monopole sur les systèmes d'exploitation à la recherche d'information.

Est-ce que les innovations technologiques de Yahoo et de Google leur permettront d'éviter cela ?

4.4 Les dangers d'un monopole

Dans tout secteur d'activité, le monopole caractérise l'entreprise qui est seule à vendre un certain produit, et il fait référence à une forme de concurrence imparfaite.

Il en résulte une diminution de l'innovation et de la baisse des prix.

Par exemple en France, sur le marché de la recherche sur Internet, le monopole actuel de Google le place en média incontournable pour toute entreprise d'e-commerce. En effet, un visiteur sur trois qui accède à un site de vente en ligne provient de Google. Les firmes qui veulent se développer sur le web doivent donc acheter de la visibilité sur les moteurs de recherche américains au travers de son service de « liens sponsorisés ». Si un monopole mondial était établi, nous assisterions probablement à une envolée des prix.

Deuxième risque plus grave encore, l'abus d'une position dominante pourrait conduire l'acteur en question à développer son propre site d'e-commerce et étendre ainsi son monopole à d'autres secteurs d'activités.

Si par exemple la recherche « voyage » conduit exclusivement vers un site de voyage, les autres acteurs du voyage en ligne perdraient leur principale source de trafic. Si l'on considère que la réservation de voyage est en passe de devenir majoritairement « online », c'est tout le secteur du voyage qui pourrait passer entre les mains du moteur de recherche leader.

Mais nous pouvons aller plus loin dans cette analyse car le monopole d'une entreprise sur le marché de la recherche d'information pourrait avoir des conséquences autrement plus graves qu'économiques.

En considérant que les moteurs de recherche tendent à constituer notre principale source d'information et d'accès à la culture, un monopole sur ce secteur pourrait affecter en profondeur notre perception et notre accès à l'information en éliminant des contenus plus critiques et plus objectifs. Une telle situation, pourrait entraîner des

dérives telles que la diffusion d'une pensée unique, la manipulation des esprits, la diffusion de mensonges orientés par une volonté unique, une propagande idéologique.

4.5 Les alternatives

Face à ces problèmes de concentration de pouvoir sur un sujet aussi sensible que l'accès à l'information, plusieurs projets associatifs ou gouvernementaux sont en passe d'émerger.

4.5.1 Culture

Google Print est certainement le projet le plus controversé. Pour rappel, le service en question est un moteur de recherche capable de retrouver une information dans des livres préalablement numérisés. Cela permettra également de procéder à un éventuel achat en ligne si l'extrait intéresse. Evidemment, c'est là aussi une plateforme idéale pour Google en termes d'image de marque. Et, AdWords obligent, cette plateforme constitue un nouveau panneau publicitaire pour les liens sponsorisés de Google.

Ce marché de la librairie en ligne est un secteur où les intérêts mis en jeu dépassent allégrement le strict cadre financier. Ainsi Jean-Noël Jeanneney, président de la Bibliothèque Nationale de France, expliquait il y a peu que « A partir du moment où on numérise ces 14 millions de livres, il y a un sentiment de puissance, [cela] pose un problème car il y a forcément un choix vu d'Amérique, et choisir, c'est forcément colorer. Ce n'est pas neutre ». Son analyse avait été publiée dans un article dans le quotidien *Le Monde*.

Cette analyse a été rapidement suivie d'effets puisque les chefs d'état de six pays membres (France, Pologne, Allemagne, Italie, Espagne et Hongrie) ont demandé à la Commission Européenne et au Conseil de l'Europe la mise en place rapide d'un débat sur une future « bibliothèque numérique européenne ». La lettre est disponible depuis le site de l'Elysée : « Le patrimoine des bibliothèques européennes est d'une

richesse et d'une diversité sans égales. [...] S'il n'est pas numérisé et rendu accessible en ligne, ce patrimoine pourrait, demain, ne pas occuper toute sa place dans la future géographie des savoirs » prévient-on.

Outre l'initiative en cours de la BNF, nommée Galica, divers projets de numérisation ou de retranscription germent, bourgeonnent et fleurissent depuis bien longtemps sur le net. Et ce, bien avant que Google ne brandisse son scanner-bulldozer. On citera ainsi le projet Gutenberg et ses 15 000 eBooks (dont certains en français). Le site des Classiques des Sciences Sociales qui rassemble 1 372 œuvres originales de 459 auteurs différents, tombées dans le domaine public québécois ou publié avec l'accord des intéressés. La bibliothèque Athena qui offre des centaines de textes d'auteurs francophones à disposition de tous. Le site Ebooks Libres et Gratuits, fruit du travail d'un groupe francophone international. Citons encore l'Association des Bibliophiles Universels ou ABU et son catalogue d'auteurs classiques... Certes, cela n'a rien de comparable aux chiffres avancés par le moteur de recherche (14 millions d'ouvrages traités en 6 ans) mais cela permet déjà de se constituer gratuitement une belle bibliothèque. En attendant.

4.5.2 Moteur de recherche

Il en est de même pour la recherche d'information.

Alors que l'hégémonie de Google se fait de plus en plus grande, il semblerait que l'Europe prenne conscience du problème. En effet, la France et l'Allemagne ont annoncé à l'occasion du conseil des ministres franco-allemands du 26 avril dernier la mise en place du projet "Quaero", destiné à développer des "outils intégrés de gestion des contenus multimédias", dont un moteur de recherche.

Selon le site du groupe de travail franco-allemand "le projet se propose de produire les machines adaptées à la recherche et à l'interprétation des informations multimédias" et donc de rendre accessible l'usage des contenus multimédias à une plus large partie de la population.

"Le projet a pour objet de créer une avancée déterminante en introduisant des technologies de rupture au service de la production, la gestion, et l'utilisation des contenus numériques multimédias. Les principales avancées technologiques portent sur la transcription, l'indexation, et la traduction automatique de documents audiovisuels multilingues, la reconnaissance et l'indexation d'images, les moteurs de recherche multimédias thématiques..."

On peut saluer la coopération du moteur français Exalead au projet. Selon Francois Bourdoncle, PDG : "Notre ambition est de parvenir à développer un outil unique pour la recherche de documents en mode texte, audio et vidéo (...) Un outil qui nous permette de mener la recherche d'information de façon unifiée et homogène, là où la plupart des moteurs restent aujourd'hui cantonnés à la recherche en mode texte".

Mais l'Europe saurât-elle faire preuve de la réactivité nécessaire pour faire face à aux acteurs privés déjà en place ?

Conclusion

Simple réseau informatique développé par l'armée dans les années 70, Internet s'est rapidement immiscé dans la vie civile jusqu'à être utilisé aujourd'hui par près d'un habitant sur six de la planète. Si la première utilisation du réseau des réseaux fut la communication par le biais des emails, il est aujourd'hui utilisé dans de très nombreuses circonstances, personnelles comme professionnelles. La plupart des applications d'Internet commencent par l'emploi d'un moteur de recherche.

Que ce soit pour rechercher un emploi, acheter un billet d'avion, suivre l'actualité ou consulter un livre, le point de départ sur Internet se trouve être l'un des trois grands acteurs de la recherche : Yahoo, Microsoft et Google.

Le combat acharné que se livrent ces trois géants risque de déboucher à terme sur un monopole. Mais plus que pour tout autre secteur d'activité, un monopole sur la recherche d'information pourrait engendrer des conséquences néfastes sur l'économie, mais aussi et surtout sur l'accès à l'information.

Si par le passé, les lois anti-trust n'ont pas eu la réactivité nécessaire pour éviter la mise en place de monopole, notamment dans l'affaire Netscape, nous sommes en droit d'attendre que sur un sujet aussi sensible, l'intervention des juridictions compétentes sauront conserver un équilibre concurrentiel.

D'autre part, nous pouvons espérer que les projets alternatifs comme ceux menés par la BNF ou l'Europe contribueront à préserver cet équilibre.

Enfin n'oublions pas qu'Internet est en perpétuelle évolution et qu'une nouvelle révolution technologique pourrait à tout moment modifier fondamentalement la place des moteurs de recherche dans l'accès à l'information.

Glossaire

AdWords : Système publicitaire du moteur de recherche Google qui affiche des annonces texte ciblées selon la recherche effectuée.

ARPANET: Advanced Research Projects Agency Network.

MIT : Massachussets Institute of Technology Centre universitaire américain célèbre pour ses nombreuses découvertes technologiques.

RAND : La Rand Corporation est une institution américaine à but non lucratif qui a pour but d'améliorer la politique et le processus décisionnel par la recherche et l'analyse.

Système d'exploitation (OS) : Ensemble cohérent de logiciels permettant d'utiliser un ordinateur et tous ses périphériques. Il assure le démarrage de celui-ci et fournit aux programmes applicatifs les interfaces pour contrôler les éléments de l'ordinateur.

UCLA : University of California Los Angeles. Principale université de Californie.

Weblog : (mot-valise anglais, contraction de web et log) un site web proposant un journal en ligne tenu par une ou plusieurs personnes. Dans son usage francophone comme anglophone, weblog est fréquemment raccourci en blog.

Bibliographie

Sites Web

Abondance : <http://www.abondance.com/>
DSI : <http://www.dsi-info.ca/>
Journal Du Net : <http://www.journaldunet.com/>
Le Monde : <http://www.lemonde.fr/>
PC Inpact : <http://www.pcinpact.com/>
Secrets 2 Moteurs : <http://www.secrets2moteurs.com/>
Search Engine Watch : <http://searchenginewatch.com/>
Search Engine Optimization : <http://www.seoachat.com/>
Université Rennes 2 : <http://www.uhb.fr/>
Webmaster World : <http://www.webmasterworld.com/>

Ouvrages

«Quand Google défie l'Europe»

Jean-Noël Jeanneney - Mille et une nuits – 2005

« The Search » : The Inside Story of How Google And Its Rivals Changed Everything»

John Battelle – Portfolio – 2005

« Microsoft Secrets »

Michael A. Cusumano - Free Press – 1998

“Google And The Mission To Map Meaning And Make Money”

Bart Milner - Electric Book Company – 2004

Annexes

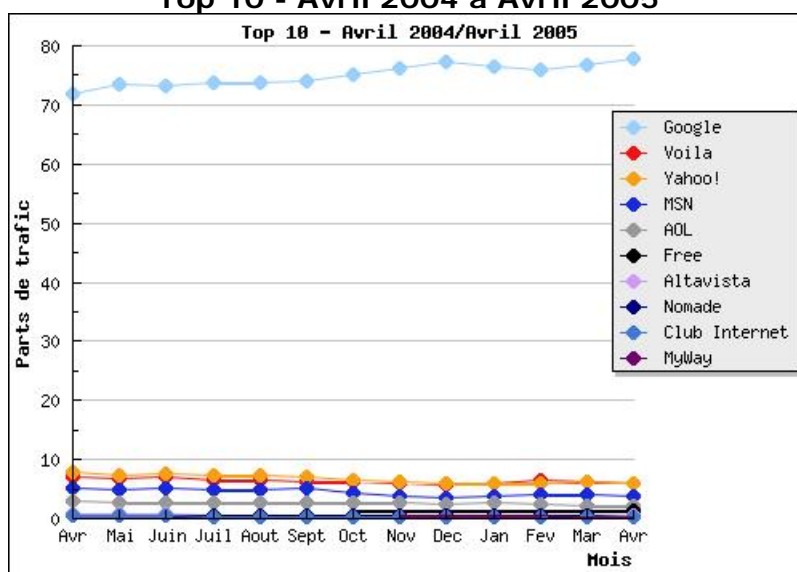
Annexe 1 – page 54 : Intégralité des parts de marché en France.

Annexe 2 – page 55 : Intégralité des parts de marché dans le monde.

Annexe 3 – page 56 : Donnée financière de Microsoft

Annexe 1

Top 10 - Avril 2004 à Avril 2005



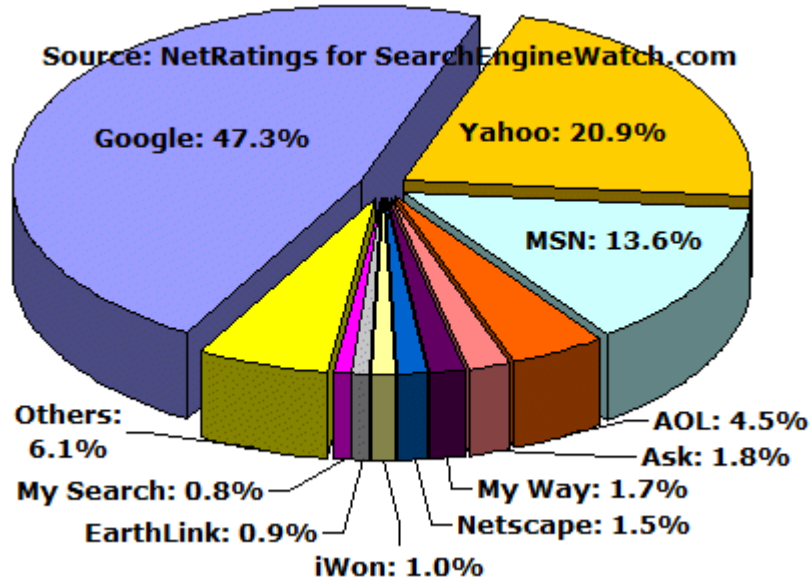
TOP 15 (Avril 2005)

Outils de recherche	% de trafic généré	Tendance
1 - Google	77.75 %	▲(+1.08)
2 - Voilà	5.89 %	▼(-0.46)
3 - Yahoo!	5.84 %	▼(-0.39)
4 - MSN	3.84 %	▬(-0.11)
5 - AOL	2.23 %	▬(-0.05)
6 - Free	1.42 %	▬(+0.06)
7 - Altavista	0.47 %	▬(-0.04)
8 - Nomade	0.37 %	▬(-0.04)
9 - Club Internet	0.35 %	▬(0)
10 - MyWay	0.35 %	▬(-0.03)
11 - Lycos	0.33 %	▬(-0.02)
12 - All The Web	0.30 %	▬(-0.02)
13 - Netscape	0.20 %	▬(-0.02)
14 - MySearch	0.17 %	▬(0)
15 - Virgilio	0.11 %	▬(+0.03)

Source Xiti 2005

Annexe 2

Share Of Searches: March 2005



Note that the figures are search-specific but probably not necessarily web-search specific. More about this is [explained](#) on the comScore search engine ratings page. Also note that some companies own more than one search site. This means searches at different sites may be combined into one overall figure for the company's entire network. The notes below provide more information of what's in each share. Remember, in all cases, only activity by those in the US is measured, even if those in the US go to a site run by a company outside the US, such as [Google UK](#).

- **Google:** Shows searches at any Google-branded web site such as [Google.com](#) or [Google Images](#).
- **Yahoo:** Shows searches at any Yahoo-branded web site such as [Yahoo.com](#) or [Yahoo Local](#). Does NOT include searches at the Yahoo-owned sites of [AltaVista](#), [AllTheWeb](#) and [Overture](#).
- **MSN:** Shows searches at any MSN-branded web site such as [MSN Search](#).
- **AOL:** Shows searches at any AOL-branded web site such as [AOL Search](#). Does NOT include searches at AOL-owned [Netscape Search](#) (this is on the chart separately) or probably other AOL-owned sites.
- **Ask:** Shows searches at [Ask Jeeves](#) but not Ask Jeeves-owned [Teoma](#). Does NOT show searches at the Ask Jeeves-owned sites of [MyWay.com](#), [iWon](#) and [My Search](#). These are shown on the chart separately.
- **Earthlink:** Shows searches at [Earthlink](#).
- **Other:** Shows searches that occur at other search sites not named on the chart. Any site not listed on the chart can be assumed to have a share that is less than the smallest named site (IE, My Search is the lowest named search site with a 0.8% share. Other sites will have shares less than this).

Annexe 3

DONNEES FINANCIERES Microsoft Corp – EURO

(en millions, excepté les revenus par action)

Exercice se terminant le 30 juin	2000 ⁽¹⁾	2001 ^(1, 2)	2002 ^(1, 3)	2003 ^(1, 4)	2004
	€	€	€	€	€
Recettes	22,467	28,041	32,003	31,334	30,957
Bénéfices d'exploitation	10,771	12,992	9,333	9,292	7,592
Revenus avant les modifications de convention comptable	9,220	8,559	6,042	7,331	6,865
Bénéfices nets	9,220	8,143	6,042	7,331	6,865
Bénéfices dilués par action avant les modifications	0.83	0.76	0.54	0.67	0.63
Bénéfices dilués par action	0.83	0.73	0.54	0.67	0.63
Dividendes en espèces par action	-	-	-	0.08	0.13
Encaisse et investissements à court terme	25,159	36,748	39,031	42,400	49,802
Total de l'actif	54,651	68,415	70,595	70,654	75,937
Capitaux propres	43,734	54,994	55,379	56,113	61,501
