Incontournable et sans contenu…

http://google.fr

Par Boris Beaude



Si l'Internet avait un centre, Google n'en serait pas loin. En 1995, alors étudiants à Stanford, Lary Page et Sergey Brin n'imaginaient pas encore que leur rencontre serait à l'origine du site internet le plus utilisé en 2004, à l'exception de celui de Microsoft. Le succès de Google est tel, qu'il est à l'origine d'un verbe aux État-Unis (*to google* : rechercher sur Google) et depuis peu en France, même s'il reste peu usité (*googler*). On *google* ainsi ses collègues, ou soi-même pour les plus narcissiques. Réputé pour la sobriété de son interface, sa rapidité et la pertinence de ces résultats, ce site s'est imposé rapidement comme le meilleur moyen de trouver une information sur l'Internet, au point d'être sollicité plus de 200 millions de fois par jour. Son interface, proposée en 88 langues, permet de chercher en 35 langues parmi plus de 4 milliards de pages. En cela, Google est devenu incontournable, en se contentant de relier efficacement des choses existantes plutôt qu'en en créant de nouvelles. Inspiré de l'ordre de grandeur mathématique Googol (10 à la puissance 100), néologisme inventé par le neveu du mathématicien Edward Kasner, Google fait référence à la quantité considérable d'informations que doit analyser le moteur de recherche.

Malgré leurs divergences d'opinions sur un grand nombre de sujets, Lary et Sergey partageaient un intérêt prononcé pour l'analyse de données. Ils étaient convaincus, l'un et l'autre, que la production croissante d'information s'accompagnerait d'une exigence accrue, quant aux moyens d'y accéder rapidement et sélectivement. Ils développèrent alors un nouvel algorithme, qui s'avérera rapidement être d'une efficacité sans égal. La rencontre de David Filo, lui aussi de Stanford et cofondateur de Yahoo! et d'Andy Bechtolsheim,

fondateur de Sun Microsystems, allait parfaire les ingrédients de cette *succes story*. Le premier allait les encourager à créer leur société et le second leur apporta les 100 000 dollars nécessaires à une levée de fonds encore plus importante. Lary et Sergey renoncèrent à leur thèse de doctorat et en 1998, avec un investissement de départ d'un million de dollars, Google inc. était née.

Le succès de Google, dit-on, repose sur la pertinence de l'algorithme utilisé pour classer les résultats d'une recherche. Appelé « PageRank », cet algorithme devenu célèbre fait référence à Lary Page et non, comme on le supposerait, aux pages ainsi référencées. Comme beaucoup d'algorithmes efficaces, celui à l'origine d'un tel succès se résume à quelques principes élémentaires. À la différence des moteurs les plus courants, Google ne se limite pas à la recherche de la pertinence des résultats en fonction de l'occurrence des mots-clés présents dans une page. Afin d'obtenir un classement plus pertinent des résultats, Google prend aussi en compte les relations qui existent entre les pages. Un lien vers une page est considéré comme une forme de reconnaissance de son contenu. Afin de rendre cette démarche encore plus efficace, le « PageRank » d'une page sera d'autant plus important que les sites qui y font référence sont eux-mêmes importants et qu'ils font référence à peu de sites. En peu de temps, Google s'est imposé comme le moteur de référence à tel point qui occupe 60 % de cette activité à l'échelle mondiale. L'efficacité opérationnelle de cette méthode, qui repose sur les processus d'autorégulation de la théorie de l'information conforte les tenants de la cybernétique et de sa prétendue capacité à intervenir dans la régulation de l'information, au-delà des automates. Pour s'en convaincre, il n'est pas inutile de reprendre la formulation de Google pour décrire leur algorithme : « PageRank est un champion de la démocratie ». En d'autres termes, Google utilise le contenu de l'Internet lui-même pour ordonner le classement des résultats d'une recherche, excluant de ce fait un arbitrage personnel quant à la valeur ou non du contenu d'une page.

On retrouve dans cette formulation et dans cette démarche l'esprit même de cette société, qui cultive une culture d'entreprise marginale, réservant le vendredi aux recherches personnelles, proposant gratuitement des jus de fruits bio dans les bureaux, faisant la promotion de la voiture électrique auprès de ses employés et se défendant de ne pas proposer un moteur de recherche qui

valorise plus les offres commerciales que la recherche elle-même. Google revendique ne pas être une société comme les autres, défendre des valeurs morales et ne pas respecter les règles du capitalisme. L'introduction de Google en bourse fin août dernier fut représentative de cette revendication. Alors que la règle veut que toutes communications soient interdites avant une introduction, Lary Page et Sergey Brin accordèrent une interview au magazine Playboy, soutenant que la devise de leur entreprise est : « *Don't be evil* » et « *Make the world better* ». Alors que la règle veut aussi que les introductions soient gérées par l'intermédiaire des banques, le choix de la société fut celui d'une technique du 17e siècle : « l'adjudication à la hollandaise ». Cette méthode consiste en la mise aux enchères des actions, sans que les différents investisseurs ne soient informés des autres offres. Au terme de l'enchère, les actions sont réparties du plus offrant au moins offrant, les derniers n'étant pas servis. Cette méthode, marginale lors des introductions, est celle qui est souvent utilisée par les États (dont la France) lors d'un emprunt. Comment ne pas y voir une provocation, alors même que cette introduction s'est faite par l'intermédiaire de leur propre site Internet, et que si des analyses au jour le jour l'ont perçu comme un échec — la valeur d'introduction étant nettement inférieure à l'attente initiale —, d'autres n'oublient pas que si Google avait procédé selon le mode d'introduction classique, la valorisation auraient probablement été plus de moitié inférieure. En l'espace de deux semaines, l'action avait déjà crû de près de 15%. Cette culture de l'utopie savamment entretenue et alimentée par un goût pour la dissidence se retrouve dans des détails qui entretiennent le mythe. Par exemple, beaucoup ignorent que Google est aussi une calculatrice scientifique de qualité. Ainsi, la requête 1+2*3 donne *7* comme résultat, *cos(2*pi)* donne *1* et *5 !* donne *120*. Mais pour ne pas s'arrêter en si bon chemin, *MMIV – X* donne *MCMXCIV* et *radius of earth* donne *6378,1 kilometers*. De même, des passionnées adeptes du *Google Bombing*, une méthode qui consiste en la manipulation des résultats, parviennent à créer des configurations surprenantes. Ainsi, le premier site qui répond à la requête « miserable failure » est celui de la bibliographie de Georges Bush sur le site de la Maison Blanche. Cela ne durera probablement pas !

Un tel succès interroge sur l'intérêt des investisseurs pour une société qui propose un service gratuit, dont les frais de fonctionnement ne sont pourtant pas du tout négligeables. Pour fonctionner, Google utilise plus de 15 000 ordinateurs

répartis sur plusieurs sites à Santa Clara, San José et Palo Alto en Californie, Herndon et Sterling en Virginie, Washington dc, Dublin en Irlande ou Bangalore en Inde. Actuellement, Google est moins développé en Asie, où son utilisation est encore relativement faible. Malgré l'automatisation des recherches, Google emploie près de 1500 personnes, dont des chercheurs et des ingénieurs de très haut niveau. Sur les six premiers mois de 2004, Google a investi plus de 80 millions de dollars en recherche et développement et plus de 100 millions en marketing. Le total de ces frais de fonctionnement s'élevait à près d'un milliard de dollars alors qu'ils n'étaient « que » de 370 millions en 2003 et 33 millions sur l'ensemble de 2000. Cette très forte croissance des dépenses répond à celle de la demande, à mesure que le contenu de l'Internet s'enrichit et que le nombre d'utilisateurs augmente. On l'aura compris, le coût des investissements permettant d'offrir une telle qualité de service est considérable et il ne peut être envisagé sans que des revenus importants soient dégagés de cette activité.

Sur ce point, Google présente aussi des singularités à même de le distinguer de ses concurrents. Les revenus de la société reposent sur deux activités. La vente de leur technologie dans le cadre de réseaux d'entreprise (intranet) ou de sites web dont le contenu exige un outil de recherche interne puissant et la vente d'espaces publicitaires. Cette dernière activité est de loin celle qui a connu la plus forte croissance, au point de représenter aujourd'hui 98% des revenus, dont la moitié par l'intermédiaire de sites partenaires. Si la vente de leur technologie correspondait en 2001 à près d'un quart des revenus de la société, cette activité représente aujourd'hui seulement 2% des revenus. En revanche, la part la plus importante des revenus est réalisée aux Etats-Unis. Sur les six premiers mois de 2004, seulement un petit tiers des revenus provenait de l'étranger, malgré leur augmentation croissante depuis 2001. En proposant un service de recherche très performant et en proposant des offres commerciales très discrètes, Google est parvenu à rendre cinq fois plus efficace la présence de partenaires commerciaux sur les pages de résultats. En cela, l'interface très épurée de Google n'enlève rien à la rentabilité de leur investissement. Ainsi, en respectant les règles de la « nouvelle économie », Google parvient à réaliser un chiffre d'affaire de plus d'un milliard de dollars sur les six derniers mois et à dégager des bénéfices considérables. En offrant un service gratuit dont l'activité ne consiste qu'à valoriser ce qui existe déjà, Google s'inscrit dans une forme très

aboutie de cette économie peut-être toujours en devenir, mais certainement enterrée trop vite en 2000 par ceux-là mêmes qui l'ont portée aux nues, ou par ceux qui n'ont jamais voulu y croire.

Cependant, en s'inscrivant dans un cercle vertueux, alors qu'augmente la demande, les revenus, mais aussi les investissements, Google est de plus en plus dépendant de son activité commerciale, sans laquelle tout le reste s'effondrerait. L'émergence ou le développement d'un concurrent pourrait remettre en cause ce fragile équilibre. L'arrivée prochaine de Microsoft sur ce marché ne laisse pas indifférent les dirigeants de Google, qui voient d'un mauvais œil la possible intégration d'un moteur de recherche à Windows. Par passion, et probablement, aussi, par anticipation, le rêve de Lary et Sergey prend aujourd'hui une forme beaucoup plus complexe, poussant à l'extrême l'idée de relier toutes formes d'informations numériques disponibles, au point d'en avoir en partie le contrôle. La société d'aujourd'hui n'a plus rien de la chambre de Larry. Depuis, ce dernier et Sergey se sont entourés de personnalités, dont la spécialité n'est pas l'innovation technologique mais stratégique ou financière. Dirigée par Eric E. Schmidt depuis 2001 et accompagné d'une multitude de spécialistes en commerce, management ou marketing, Google n'a plus rien d'une société de fond de garage. Aujourd'hui, la société cherche un équilibre entre la qualité du service offert et la possibilité de générer des revenus croissants.

Or, le service offert est aujourd'hui tellement vaste, qu'il est généralement en grande partie méconnu ou peu utilisé. C'est pourtant en travaillant sur cet approfondissement et sur la diversification de l'offre que Google envisage de rester compétitif. Au-delà de la qualité de son moteur de recherche et de la discrétion de l'offre publicitaire, le site propose des services souvent peu utilisés ou très novateurs. Outre les traductions automatiques de piètre qualité et de plus en plus courantes, Google est le seul moteur qui propose l'accès à une copie de la page recherchée. Cette fonction, parfois contestée, permet d'accéder à une page, même si cette dernière n'est momentanément pas accessible ou si elle a été effacée sur le site d'origine. Google permet aussi d'obtenir la liste des sites qui présentent des similarités avec un autre site ou de chercher tous les sites qui font référence à un site particulier (link:espacestemps.net). Enfin, il est possible

de restreindre sa recherche à un site (internet site:lemonde.fr), à des images ou à un ensemble de langues. Mais, finalement, il s'agit bien là de l'offre la plus connue du moteur, tant elle répond, en somme, à ce que l'on peut en attendre.

Le développement de Google s'est surtout fait sur d'autres fronts. Peu à peu, Google a étendu son service au-delà des pages web. En référençant aussi les documents pdf, word ou excel, le site a ouvert des perspectives nouvelles, ce type de documents contenant souvent des informations nombreuses et pertinentes. Aussi, Google propose des recherches dans les newsgroups. Le moteur propose en l'occurrence l'accès aux archives des messages échangés par le réseau Usenet depuis 1981, soit plus de 800 millions de messages. Ce réseau, très utilisé par les informaticiens dans les années 80 avant de toucher un plus large public à la fin des années 90, reste méconnu du grand public. Google, propose aussi un accès plus classique aux pages web selon le classique « annuaire », qui classe les pages par thèmes, comme le font d'autres portails comme Yahoo, mais là encore, sans intervention humaine puisque le site utilise l'Open Directory Project, qui est un système gratuit de référencement, reposant sur la bonne volonté de chacun pour référencer les sites les plus intéressants. Là encore, l'idéologie cybernétique n'est pas loin. Enfin, le site propose un « portail » qui présente une synthèse de l'actualité et des liens vers les articles d'origine, à partir de plus de 500 sources, dont Reuter, *Libération*, *Le Figaro*, *L'Humanité*, *Le Temps*, *L'Express*, *Le Nouvel Observateur*, mais aussi le *New York Times* et d'autres journaux du monde entier. Ainsi, il est possible de faire une recherche à partir d'un ou plusieurs mots-clés dans la majeure partie des journaux d'actualité mondiale. Il reste au lecteur à faire son choix entre *Le Parisien*, *Le Point*, *L'Humanité* ou *Le Figaro*. Dans cette logique, un autre service, pour l'instant restreint aux États-Unis, permet de comparer les prix proposés pour un produit entre une multitude de commerces en ligne, au même titre que les nombreux sites offrant déjà ce service, tel que Kelkoo ou Monsieurprix. La stratégie de Google semble évidente : rendre l'information disponible sur l'Internet la plus accessible possible, et au même titre que les chaînes privées de télévision, se rémunérer du détournement de l'attention ainsi obtenue, vers des offres de partenaires commerciaux.

Actuellement, Google développe son activité au-delà de la recherche d'informations, en s'intéressant aussi à la mise en relation des internautes entre eux. Les récentes acquisitions et le développement de nouvelles offres telles que Gmail pour la messagerie électronique, Blogger pour les Blog, le contesté Orkut pour les « réseaux sociaux » et Picasa et Hello pour le tri et le partage de photos numériques sont autant de nouvelles orientations de la société dont l'activité semble de plus en plus orientée vers le contrôle et la valorisation de l'information, la plus vaste possible et sous toutes ses formes. La société a aussi annoncé qu'elle propose un système de calcul distribué (Google Compute), permettant d'allouer la capacité de calcul de son ordinateur aux partenaires de notre choix lorsqu'on ne l'utilise pas.

Google, en l'espace d'une demi-décennie, est devenu un site incontournable, dont l'activité principale consiste en la mise en relation de ressources numériques réparties dans le monde entier et dont les revenus reposent sur la mise en relation des utilisateurs de ce service avec des partenaires commerciaux. En limitant le classement des résultats à un algorithme qui utilise l'Internet lui-même pour en évaluer le contenu, cette société est au cœur des problématiques contemporaines de l'autorégulation des systèmes complexes et de la valeur accordée à la « neutralité » de tels systèmes, qui se présentent comme la résultante de la somme des choix individuels.

Pourtant, un tel succès ne va pas sans quelques remises en cause. Conscient qu'une société repose sur des valeurs qu'un tel système n'est pas en mesure de respecter, il est de plus en plus courant que des exceptions à la règle ne soient pas exclues. Par exemple, de nombreux sites sont sur une « black-list », lorsqu'ils ne respectent pas des valeurs fondamentales, et ce, selon le pays de consultation. S'il est admis que cette mesure est inefficace, elle présente en revanche l'intérêt de respecter les règles établies de la vie collective, en ne reportant pas le problème sur un arbitrage technologique. Aussi, les sollicitations commerciales sont de plus en plus surveillées, suite à quelques erreurs telles que la proposition de compagnies aériennes lors d'une recherche sur les attentats du 11 septembre. Les résultats d'une recherche sur « avion 12 août » ou « avion 11 septembre » diffèrent. Dans le premier cas, le site propose des liens vers des partenaires commerciaux alors qu'ils sont absents dans la

deuxième requête, qui elle, renvoie à l'actualité. Le sens accordé à cette date ne dépend pas du célèbre algorithme de Google, mais bien de filtres qui ont été mis en place pour éviter les associations polémiques. Avec le développement de Gmail (messagerie), de Blogger (Blog) et de Orkut (réseaux sociaux), le problème soulevé est d'un autre ordre. D'un point de vue contractuel, l'utilisateur de ces services s'engage à accepter que le contenu de ses échanges ou de ce qu'il produit puisse être analysé automatiquement par des robots afin de cibler au mieux le lien avec des partenaires commerciaux de Google. Ces services reposant sur des données nominatives associées à des idées, privées pour la messagerie et publiques pour le Blog ou les réseaux sociaux, l'automatisation de l'analyse de ces contenus va au-delà de ce qui se fait actuellement avec les cookies pour les bannières publicitaires. Si beaucoup répondent à cela qu'ils n'ont rien à cacher à des robots, participant ainsi à la formalisation de l'intelligence collective défendue par le philosophe Pierre Lévy [*L'intelligence collective. Pour une anthropologie du cyberespace*, La Découverte, Paris, 1994], d'autres pensent qu'un tel système confère un pouvoir qui ne doit pas être laissé entre toutes les mains, voyant dans la société de l'information et dans les technologies de l'information et de la communication une supercherie dont les fondements peuvent être comparés aux régimes les plus totalitaires, comme l'a souligné à plusieurs reprises le sociologue Philippe Breton [*L'utopie de la communication*, La découverte, 1992]. En cela, les plus grands défenseurs de la liberté d'expression et de la libre circulation, en poussant leur logique à l'extrême, mettent peu à peu en place un dispositif de surveillance de la circulation des idées sans équivalent dans l'histoire de l'humanité. « *Don't be evil* » disent-ils ! Bien que provocante, cette formulation, qui n'est pas innocente, n'est-elle pas l'essence même de cette problématique ? À défaut de faire le bien, ces pionniers ne pensent-ils pas faire le moindre mal ?

Google.

D'autres moteurs de recherche :

Yahoo!, Voilà, MSN, Alltheweb, Dir.com, Eo, Kartoo

Résumé

Si l’Internet avait un centre, Google n’en serait pas loin. En 1995, alors étudiants à Stanford, Lary Page et Sergey Brin n’imaginaient pas encore que leur rencontre serait à l’origine du site internet le plus utilisé en 2004, à l’exception de celui de Microsoft. Le succès de Google est tel, qu’il est à l’origine d’un ...