

Serge Proulx

Professeur titulaire. École des médias. Université du Québec à Montréal.
Professeur associé. Télécom ParisTech.

Google, grand architecte de l'information mondiale

S. Proulx et S. Couture (2008)

12

Attention, il s'agit d'un document de travail. Veuillez citer et vous référer à la version définitive :

S. Proulx et S. Couture (2008) *Google, grand architecte de l'information mondiale* in B. Badie, S. Tolotti, eds, *L'état du monde 2009*, La Découverte, Paris, p. 224-229.

Ce texte a été mis en ligne afin que les usagers du site Internet puissent avoir accès aux travaux de Serge Proulx. Les droits d'auteur des documents du site Internet [sergeproulx.info](http://www.sergeproulx.info) demeurent aux auteurs des textes et-ou aux détenteurs des droits. Les usagers peuvent télécharger et-ou imprimer une copie de n'importe quel texte présent sur [sergeproulx.info](http://www.sergeproulx.info) pour leur étude personnelle et non-marchande. Vous ne pouvez en aucun cas distribuer ce document ou l'utiliser à des fins lucratives. Vous êtes cependant invités à diriger les visiteurs vers [sergeproulx.info](http://www.sergeproulx.info) pour qu'ils accèdent aux textes.

Document téléchargé depuis <http://www.sergeproulx.info>

Google, grand architecte de l'information mondiale

Serge Proulx et Stéphane Couture

Laboratoire de communication médiatisée par ordinateur (LabCMO),

Université du Québec à Montréal (UQAM)

En septembre 2008, Google a fêté ses dix ans. En une décennie à peine, l'entreprise – cotée en bourse depuis 2004 – est devenue un acteur incontournable du Web, affichant un chiffre d'affaires de 16 milliards de dollars en 2007. Son moteur de recherche, le plus utilisé au monde, est au centre de la vie quotidienne des internautes : en 2007, il captait 87% du marché en France ; et 62% aux États-Unis. En décembre 2007, en termes d'audience mondiale, le site comptait 588 millions de visiteurs uniques. Comme trois-quarts des recherches sur Internet passent d'abord par Google, les autres moteurs de recherche apparaissent loin derrière, y compris Yahoo! (portail très populaire mais dont le chiffre d'affaires n'est pas à la hauteur de sa fréquentation). Comble de popularité, le nom Google est devenu un verbe : « to google » est répertorié dans des dictionnaires usuels comme signifiant « l'action de faire une recherche sur la Toile ».

Google, c'est pourtant aujourd'hui bien davantage qu'un simple moteur de recherche, si performant soit-il. Depuis 2003, l'entreprise s'est développée de façon tentaculaire : elle constitue aujourd'hui une plate-forme d'applications offrant en permanence de nouveaux services dont certains sont spectaculaires : *Google Maps* offre une cartographie gratuite de l'ensemble du globe ; *Google Print* vise la constitution d'une bibliothèque en ligne rassemblant tous les savoirs du monde ; *Google Scholar* met à disposition des internautes un ensemble impressionnant d'articles et références scientifiques... Et la liste s'allonge de jour en jour.

Mais c'est sur le marché des logiciels de bureautique que la prochaine bataille majeure sera livrée, dans un contexte de concurrence féroce entre Microsoft (qui, en 2008, a souhaité intégrer Yahoo!) et Google. Google et Adobe bousculent en effet l'hégémonie du géant informatique, largement dominant dans ce domaine – Microsoft a généré en 2007 un chiffre d'affaires de 51 milliards \$US et dégagé des profits de 14 milliards \$US – en inventant un

modèle de bureautique en ligne baptisé *Software as a Service* (SaaS). Plus besoin d'acheter les produits Microsoft : les utilisateurs ont accès en ligne, gratuitement, aux logiciels mis à disposition par Google et Adobe. Cette confrontation entre le modèle traditionnel et propriétaire de Microsoft et le modèle de bureautique nomade et partagée de Google, qui s'appuie sur le développement des logiciels libres, pourrait voir s'imposer un nouveau modèle économique susceptible de mettre fin au monopole de Microsoft. C'est d'ailleurs bien pourquoi la firme de Redmond cherche aujourd'hui à investir dans les technologies de réseau. Cette stratégie de diversification de Google – et en particulier, sa volonté d'investir dans les usages nomades – traduit la finalité ultime de l'entreprise qui est, depuis sa création, « d'indexer l'Internet tout entier » (Vinton Cerf). L'établissement à l'échelle du globe du modèle des usages nomades signifierait ainsi la réussite pour Google de son projet d'organiser l'ensemble de l'information disponible sur la planète. Une simple connexion Internet suffirait à tout usager, à partir de tous types de terminaux (ordinateur, assistant numérique, téléphone), pour utiliser les applications de son choix (courriel, traitement de texte, agendas, jeux) puis sauvegarder l'ensemble des données personnelles ainsi produites dans les espaces mémoire de Google.

Une réussite fulgurante

Le nom de la firme est inspiré du terme mathématique *googol* qui désigne le nombre 10 à la puissance 100. L'appellation a été choisie pour symboliser la mission que s'est officiellement donnée l'entreprise : « organiser toute l'information du monde pour la rendre accessible et utile à tous », comme le dit son slogan marketing. En recourant au moteur de recherche et aux multiples services de Google (communication, bureautique, cartographie, socialisation, gestion de photos, échange de vidéos), l'internaute se met subtilement sous l'emprise du colosse informatique (et publicitaire) qui impose son dispositif de classement des informations du monde. L'expansionnisme mondial de Google traduit une volonté de domination économique somme toute classique mais cette domination n'est pas exclusivement économique dans la mesure où les biens sont de nature informationnelle. Or, l'information est au cœur des transformations contemporaines du mode de production. Dans ce contexte, la captation d'une quantité phénoménale d'informations par une seule entreprise pourrait

signifier la possibilité d'un contrôle plus profond (idéologique?) et plus ambigu (quelles que soient les intentions des fondateurs). Certains critiques vont jusqu'à craindre que les serveurs de Google ne deviennent un gigantesque « cerveau mondial » contrôlé par des individus sous l'emprise d'une idéologie posthumaniste et néoconservatrice. Mais nous n'irons pas jusque-là...

Tout a commencé en 1995-1996 quand deux étudiants de Stanford, Larry Page et Sergey Brin inventent l'algorithme *PageRank* permettant de définir l'ordre des pages dans les résultats de recherche. L'originalité de cet algorithme, par rapport aux précédents engins de recherche, est de prendre en compte le nombre de liens qui réfèrent aux documents indexés. Ainsi, plus une page est référée par d'autres sites web, plus cette page aura une grande visibilité sur Google. Cet algorithme donne ainsi des résultats plus pertinents que ses concurrents. Toutefois, des critiques ont mis en évidence la tendance de cet algorithme à favoriser un conformisme des résultats de recherche qui mettrait en danger la diversité culturelle. En effet, ce type de référencement augmenterait de manière récursive la popularité des sites considérés comme déjà populaires. Sans compter le fait que le classement prend aussi en compte une stratégie commerciale qui passe par la vente de mots-clés à des entreprises, facteur qui a davantage à voir avec la manipulation des recherches qu'avec la pertinence de contenu des sites en regard des requêtes.

Une innovation importante des concepteurs du moteur de recherche a été leur décision de s'appuyer sur l'exploitation en réseau de plusieurs ordinateurs peu performants et peu coûteux plutôt que de recourir à des supercalculateurs comme le faisaient alors tous les engins de recherche. Cette architecture distribuée est aujourd'hui encore à la base des nombreux centres de données (*data center*) de Google : ces « fermes de serveurs » situés aux quatre coins du monde (Amérique, Europe, Asie) regrouperaient plus de 500 000 ordinateurs. Des pans entiers d'Internet sont répliqués sur ces grappes d'ordinateurs, condition nécessaire à l'établissement d'un système robuste, fiable et doté d'un temps de réponse exceptionnellement rapide. Ce système est le cœur technologique de l'entreprise : il suppose la réplication à l'identique des pages les plus consultées sur une multitude de serveurs, pages accessibles selon la situation géographique de l'internaute. La stratégie de Google consiste à diversifier les modes de réplication (serveurs et logiciels de conception différente, codage des algorithmes à l'aide de

langages de programmation différents) de manière à contrer le piratage éventuel de ses manières de faire. Ces secrets de fabrication de même que le nombre et la localisation secrète des *data center* constituent un avantage compétitif décisif pour l'entreprise.

Le nom de domaine « google.com » fait son apparition en 1997 mais c'est un an plus tard seulement que Page et Brin créent leur entreprise sous le nom de Google Inc., après avoir recueilli plus d'un million de dollars auprès d'amis, de membres des familles proches et d'investisseurs comme Andy Bechtolsheim, co-fondateur de Sun Microsystems. Les bureaux se trouvent alors dans un garage de Menlo Park (Californie), conformément à la mythologie de la Silicon Valley. Mais l'entreprise migre à Palo Alto dès 1999, pour finalement bâtir en 2003, à Mountain View, son célèbre siège social : le *Googleplex*.

Dès l'origine, Google était perçu comme l'un des plus importants moteurs de recherche du monde. Mais c'est en 2000 que le modèle économique de la firme est imaginé : l'entreprise entreprend de vendre de la publicité sur ses sites en plaçant des liens sponsorisés et des bannières ciblées qui s'affichent en fonction des requêtes des utilisateurs. Au premier trimestre de 2008, Google affichait un revenu de plus de 5 milliards US\$ dont 99% provenait des revenus publicitaires, ce qui démontre bien le rôle moteur de la publicité pour l'entreprise. En juin 2000, Google comptait déjà plus d'un milliard de pages indexées. Depuis 2005, Google a cessé de publier le nombre de pages indexées. Toutefois, au moment où Yahoo! annonçait avoir indexé plus de 20 milliards de pages, Google affirmait avoir indexé trois fois plus de pages que ses compétiteurs.

Le culte de l'innovation

Aux États-Unis comme en Europe, le mode de gestion interne et l'organisation du travail chez Google sont reconnus comme novateurs. Selon Mats Carduner, directeur général de Google France, la mise en place d'un cadre de travail agréable, d'une culture d'entreprise orientée vers l'innovation et d'un management informel favorisent la créativité des salariés : « *Je crois profondément au cercle vertueux de l'innovation* ». Ainsi, en donnant à chaque salarié la possibilité d'allouer 20% de son temps de travail à ses « projets personnels », l'entreprise a su créer un mode de gestion en résonance avec l'éthique du travail des ingénieurs informatiques. La promotion de ce « cercle vertueux » ne se réduit pas à la mise en place d'un environnement

de travail de qualité. Il prend parfois une dimension quasi-religieuse. Ainsi en est-il du rôle joué au sein de Google par Vinton Cerf (l'un des créateurs historiques du protocole TCP/IP à la base d'Internet), vice-président depuis septembre 2005 et présenté comme « chief Internet evangelist » de la firme. Cerf apparaît ainsi dans le rôle de visionnaire et d'ambassadeur itinérant de la firme à l'échelle du globe. Par exemple, il voudrait bien que l'entreprise puisse conserver la compatibilité des documents dans le temps de manière à ce que tout document soit toujours lisible indépendamment des changements de versions des logiciels. De même, Cerf se montre favorable au maintien de la neutralité de l'Internet, ce principe voulant que la même priorité soit accordée à la circulation de tous les types de données.

Par ailleurs, Google soutient massivement les développeurs du logiciel libre à tel point que certains analystes y voient là « l'arme secrète » de l'entreprise. Dès 1996, les prototypes du moteur de recherche utilisaient, comme aujourd'hui, des serveurs fonctionnant sous GNU/Linux. L'entreprise compte parmi son personnel salarié plusieurs développeurs du libre (navigateur Firefox, serveur Apache, système de partage de fichiers Samba). De même, le programme *Summer of Code* – mobilisant des étudiants en informatique reliés à des projets en libre – permet à Google de prendre connaissance des projets les plus novateurs du domaine de manière à pouvoir les intégrer immédiatement s'ils s'accordent à sa mission de « grand architecte de l'information mondiale ». Alors que Google utilise déjà gratuitement une bonne partie de la bande passante financée par les opérateurs de télécommunications, cet engagement envers le libre permet à Google d'utiliser un nombre important de logiciels de façon gratuite.

Ce « culte de l'innovation », surtout lorsqu'il prend un caractère évangélique, rend possible la mobilisation d'une armée d'informaticiens zélés construisant passionnément l'infrastructure assurant « l'organisation de toute l'information du monde »... Ce mode de fonctionnement n'est pas sans lien avec des formes émergentes du capitalisme informationnel qui se déploient aujourd'hui grâce aux réseaux numériques et qui permettent aux firmes de capter à coût presque nul les idées neuves et les projets innovants formulés par les internautes ordinaires (*crowdsourcing*).

Libertés sous caution

Si le moteur de recherche est le coeur de métier de l'entreprise, le développement de services tels le système de messagerie *Gmail* ou la suite bureautique en ligne *GoogleApps*, ou l'acquisition d'entreprises – comme YouTube en 2006 – lui donnent une position stratégique bien plus centrale encore. Le développement des sites de réseaux sociaux (Friendster, MySpace, Facebook...), donnent ainsi aux services de Google une place grandissante. Comme en témoigne *Orkut*, un réseau social de Google très utilisé au Brésil. En raison de l'importance croissante de ces sites de socialisation, Google a décidé de lancer *OpenSocial*, une boîte à outils commune dont l'objectif est de favoriser la compatibilité entre ces différents sites. Plusieurs portails se sont déjà joints à l'initiative, comme Friendster, LinkedIn, MySpace et Yahoo!.

Dans ce contexte, de nombreux analystes s'inquiètent des tendances monopolistiques de Google. Ils soulignent notamment les dangers que la concentration de ses activités pourrait faire peser sur la liberté d'expression et la protection de la vie privée. Google conserve en effet pendant dix-huit mois des données personnelles concernant les internautes afin de pouvoir leur proposer des publicités ciblées. Cette collecte d'informations se fait au moment de l'activation du moteur de recherche (identification des thèmes, habitudes de navigation, auto-description sur les sites de socialisation) et des achats en ligne, mais aussi via l'usage du courrier électronique *Gmail*, le contenu même des messages étant collecté et analysé. En 2007, le groupe *Privacy International* a d'ailleurs classé Google, bon dernier parmi une vingtaine d'entreprises en matière de respect de la vie privée. Le rachat de DoubleClick – une régie publicitaire ayant développé des outils sophistiqués de traçage de données personnelles – a également suscité l'inquiétude aux États-Unis et en Europe. Cette acquisition, la plus importante depuis YouTube, permettra à Google de consolider sa position dans le domaine de la publicité en ligne et d'augmenter significativement sa capacité à recueillir et exploiter les données personnelles des internautes. Cette situation provoque la jalousie de Microsoft et Yahoo!, mais inquiète les organisations de défense des consommateurs, qui soulignent la quantité déjà phénoménale d'informations personnelles détenues par Google.

Des organisations citoyennes s'inquiètent également des formes de censure exercées par Google dans certains pays. Amnesty International a déjà dénoncé le filtrage par le moteur chinois de Google de mots-clé tels que *droits de l'homme*, *réforme politique*, *Tiananmen* ou

Falun Gong. Certains actionnaires sont également intervenus, en assemblée générale, pour suggérer à la direction de l'entreprise de développer une attitude plus respectueuse de la liberté d'expression. En avril 2007, le fond de pension de la Ville de New York (détenant alors près de 200 millions de dollars d'actions de Google) a proposé que l'entreprise mène une politique favorisant le libre accès à Internet, en insistant sur le fait que les données personnelles ne devaient pas être stockées dans un pays qui pratique le contrôle de la Toile et que l'entreprise devait résister légalement aux demandes de censure. La proposition a été rejetée. En mars 2008, dans la même veine, des actionnaires ont suggéré de créer un comité consultatif chargé d'informer la direction de Google sur les conséquences éventuelles de ses décisions en matière de droits de l'homme. Le Conseil d'administration de l'entreprise a recommandé aux actionnaires de voter contre cette proposition. Si le mot d'ordre de l'entreprise est toujours « Don't be evil » (« Ne fais pas le mal »), cette attitude semble de plus en plus difficile à concilier avec les impératifs commerciaux et capitalistes qui participent de son ascension retentissante.

Références

Battelle, John (2005) *The Search. How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, Portfolio (Penguin Group), New York.

Baudriller, Marc et al. (2008) "Comment Google a changé le monde", *Challenges*, 123, 8-14 mai, p. 48-59.

Cassin, Barbara (2007) *Google-moi. La deuxième mission de l'Amérique*, Albin Michel, Paris.

Le Crosnier, Hervé (2008) « Mouvements tectoniques sur la Toile », *Le Monde diplomatique*, mars, p. 19.

Vise, David A. et Mark Malseed (2005) *The Google Story*, Macmillan, New York.

Liens (URL):

Présentation de l'entreprise :

<http://www.google.com/intl/fr/corporate>

« Google. 2008. Google France nommé meilleur environnement de travail en 2008 par l'Institut Great Place to Work® »

http://www.google.com/press/20080321_greatplace.html

« SearchEngine Land. 2007. Google is 10 years old ? Finding the Real Google Birthday »

<http://searchengineland.com/070914-104722.php>

Résultats financiers (premier trimestre 2008) :

<http://www.webrankinfo.com/actualites/200804-resultats-financiers-google-1er-trimestre-2008.htm>

« Googlinside: le retour du bâton de l'affaire DoubleClick »

<http://www.googlinside.com/index.php/2007/04/26/110-google-le-retour-de-baton-de-l-affaire-doubleclick>

« Le nouveau monde version Google »

<http://hebdo.nouvelobs.com/p2205/articles/a332473.html>

« Un algorithme qui juge de la conformité »

<http://www.fluctuat.net/blog/1617-Un-algorithme-qui-juge-de-la-conformite>

« Le logiciel libre est l'arme secrète de Google contre Microsoft »

<http://framablog.org/index.php/post/2008/01/24/Microsoft-Google-logiciel-libre>

« Privacy International donne un carton rouge à Google »

<http://www.zdnet.fr/actualites/internet/0,39020774,39370136,00.htm>

« Google (again) Opposes Anti-Censorship and Human Rights Proposals »

<http://michaelzimmer.org/2008/03/26/google-again-opposes-anti-censorship-and-human-rights-proposals/>

Pour en savoir plus :

Google :

<http://www.google.com>

Entrée Wikipedia de Google :

<http://fr.wikipedia.org/wiki/Google>

Google Stories :

<http://www.google-stories.com>

Googlinside :

<http://www.googlinside.com/index.php>

La section Google du carnet de Michael Zimmer, chercheur au Yale Law School :

<http://michaelzimmer.org/category/search-engines/google/>

The Googlization of Everything. Livre rédigé par Siva Vaidhyanathan :

<http://www.googlizationofeverything.com>