

LE MONDE SELON GOOGLE

UNE COMMODITÉ PRESQUE BANALISÉE ; UN TUYAU DE PLUS, SOUVENT BRANCHÉ PAR LES MÊMES SOCIÉTÉS QUI APPORTENT EAU, TÉLÉPHONE OU TÉLÉVISION : LA TOILE D'INTERNET NE SUSCITE PLUS NI GRANDES ANGOISSES NI ENVOLÉES LYRIQUES. TOUT EST LÀ, À PORTÉE DE SOURIS, DANS LES MÉMOIRES GIGANTESQUES DES MOTEURS DE RECHERCHE... MAIS QUELS SONT LEURS CRITÈRES DE CHOIX ? QUELS BIAIS INTRODUISENT-ILS DANS LES REPRÉSENTATIONS ? DISSECTION DU PREMIER D'ENTRE EUX, GOOGLE.

PAR PIERRE LAZULY*

INTERNET, AVEC SES 3 MILLIARDS de pages, est souvent décrit comme la plus complète des encyclopédies : une incomparable documentation mise gracieusement à notre disposition, et des outils qui savent répondre dans la seconde à la moindre de nos interrogations. Les moteurs de recherche sont si performants qu'il suffit de quelques mots épars pour retrouver une information quand la mémoire nous fait défaut.

Ces outils incontournables sont, paradoxalement, de moins en moins nombreux : seules quatre entreprises américaines parviennent encore à proposer à un public mondial un service de qualité. Avant de prétendre aiguiller l'internaute dans un volume de données sans cesse croissant, il faut en effet pouvoir mobiliser des milliers d'ordinateurs pour parcourir la Toile et répertorier l'information disponible. Mais il faut surtout savoir en extraire les pages les plus pertinentes. C'est cette capacité, l'« intelligence » du moteur de recherche, qui en fera ou non le succès. Google l'a prouvé, en devenant en moins de trois ans le moteur de recherche le plus employé au monde : son approche novatrice lui permettait en général de proposer, dès la première page de résultats, l'information recherchée.

Le bouche-à-oreille fut immédiat : les initiés encourageaient leurs amis à utiliser ce moteur « génial », et Google allait passer de 10 000 requêtes par jour, début 1999, à plus de 200 millions au printemps 2003 : 53 % des requêtes mondiales lui sont désormais confiées, au point que nombre de ses 70 millions d'utilisateurs en viennent à assimiler l'Internet tout entier à cet incomparable outil. « Google est insensiblement devenu un instrument essentiel qui dépasse largement l'idée qu'on se fait généralement

* Auteur des *Chroniques du menteur* (menteur.com), et animateur de « L'autre portail » (rezo.net).

d'un moteur de recherche, résume le journaliste Francis Pisani. Je n'interrogeais plus seulement le moteur de recherche pour avoir accès à des sites contenant de l'information. Les résultats affichés en réponse à une question sont suffisants et les sites dont il donne la référence ne fonctionnent plus que comme instrument de vérification (1).»

La suprématie de Google n'est cependant pas sans soulever de réelles interrogations : comment un algorithme, si « génial » soit-il, peut-il choisir les dix réponses « les plus pertinentes » pour la requête « Irak », sur trois millions de pages contenant ce mot ?

Comme tout moteur de recherche, l'outil souffre en premier lieu d'une limitation importante : il ne peut proposer que l'information offerte au grand public. Si personne n'a jugé bon de publier d'article sur les moeurs du gypaète barbu (pour reprendre l'expression chère à Alexandre Vialatte), toute recherche sur ce thème restera vaine : en faisant « appel à Internet », on n'interroge pas l'ensemble des connaissances disponibles, mais seulement celles que différents contributeurs - universités, institutions, médias, particuliers... - auront choisi de proposer en libre accès (au moins pour un temps). La qualité de cette offre joue donc un rôle primordial dans la pertinence des réponses proposées.

Or, si le nombre total de pages accessibles ne cesse de progresser, certaines sources institutionnelles ont volontairement appauvri leurs sites. Au lendemain du 11 septembre 2001, nombre de sites officiels aux Etats-Unis ont ainsi été expurgés de données « sensibles » - comme ce site de l'armée américaine qui présentait fièrement au grand public ses huit entrepôts d'armes chimiques (2), mais un bon nombre d'informations civiles ont aussi été retirées. Geographical Information Services a interdit l'accès à ses cartes du réseau routier (3), tandis que l'Etat de Pennsylvanie retirait les plans de ses infrastructures de télécommunications, de ses écoles et de ses hôpitaux (4). Sous couvert de lutte contre le terrorisme, certaines entreprises ont fait disparaître des informations que les groupes environnementaux n'avaient pu obtenir qu'au terme d'âpres batailles : des producteurs californiens d'électricité ont ainsi retiré les données relatives aux missions polluantes de leurs centrales électriques (5)...

L'effondrement de la nouvelle économie, en 2001, a lui aussi contribué à ce retrait : les éditeurs en ligne sont de plus en plus nombreux à réserver leurs articles à leurs seuls abonnés. Cette stratégie, visant à réaliser quelques recettes supplémentaires, a toutefois un revers : leur disparition de la Toile. Les sites auxquels on n'accède que sur abonnement (même gratuit) sont en effet ignorés par les moteurs. Aussi, même si le New York Times a publié le mois dernier une remarquable enquête sur les moeurs du gypaète barbu, celle-ci ne vous sera pas proposée. La majorité des articles de presse sont ainsi devenus virtuellement invisibles.

Dans le même temps, de nouveaux acteurs s'approprièrent le réseau. Des entreprises désireuses de s'assurer une visibilité en faisaient un support privilégié de leur communication ; des organisations militantes y découvraient un autre moyen de promouvoir leur cause... Les internautes, surtout, créaient toujours davantage de sites personnels : l'« autopublication », autrefois réservée aux plus technophiles, s'est démocratisée avec l'apparition d'outils plus simples à utiliser (lire l'encadré ci-dessous).

Au milieu des années 1990, confrontés à ce flux d'informations supplémentaires, deux

étudiants en informatique de l'université américaine de Stanford, Sergey Brin et Larry Page, suivent une intuition : un moteur de recherche basé sur l'étude mathématique des relations entre les sites produirait de bien meilleurs résultats que les techniques rudimentaires employées à l'époque. Persuadés que les pages les plus « pertinentes » sont celles qui sont les plus fréquemment citées (celles que les autres sites ont choisi de référencer à l'aide de liens hypertextes), ils décident d'en faire leur projet d'études et posent les bases d'un moteur plus « mathématique », qu'ils baptiseront Google à la création de leur entreprise, en septembre 1998.

Pour évaluer la « pertinence » des pages de la Toile, Brin et Page inventent le « PageRank », une échelle de valeurs propre à Google. La valeur d'une page Web y est sans cesse réévaluée en fonction du nombre de citations dont elle fait l'objet. Les sites isolés, destinataires d'aucun lien hypertexte, demeurent ainsi peu visibles, sans « légitimité ». Les sites abondamment cités deviennent en revanche, aux yeux de Google, des sites de référence. Cet algorithme original donne des résultats impressionnants.

Le système présente pourtant déjà un inconvénient : les sites nouvellement créés sont handicapés et ne deviendront visibles que s'ils parviennent à attirer l'attention de sites déjà bien établis. « PageRank repose sur la nature purement démocratique du Web », affirment les fondateurs de Google, qui doivent néanmoins concéder que « les votes issus de pages elles-mêmes importantes comptent plus et aident à rendre d'autres pages importantes ». Une démocratie étonnante, où les acteurs déjà influents disposent d'un droit de vote beaucoup plus important que les nouveaux entrants.

Chasseurs de renommée

UNE ANECDOTE RAPPORTÉE par Andrew Orlovski dans The Register est instructive (6). Le 17 février 2003, un article publié à la « une » du New York Times décrivait le mouvement de protestation antiguerre comme l'émergence d'une seconde superpuissance. « Les immenses manifestations antiguerre à travers le monde ce week-end, assurait-il, nous rappellent qu'il existe sans doute encore deux superpuissances sur la planète : les Etats-Unis et l'opinion publique. » Cette expression fut rapidement reprise par le secrétaire général des Nations unies, M. Kofi Annan. Une recherche sur l'expression « second superpower », confiée à Google les semaines suivantes, renvoyait bien à cette définition originale.

Un universitaire de Harvard, James F. Moore, décidait alors d'allumer un contre-feu : le 31 mars, il inaugurait son site personnel par un article intitulé « La seconde superpuissance dévoile son beau visage (7) ». Un texte des plus anodins, dans lequel le terme de « seconde superpuissance » réapparaissait dans une version édulcorée propre à séduire même un membre du Parti républicain. D'autres « techno-utopistes » allaient alors se reconnaître dans sa pensée et, par leurs recensions très influentes sur le réseau, faire de son obscur article « la » référence sur le sujet. De fait, un mois plus tard, 27 des 30 premières réponses proposées par Google pour la recherche « second superpower » renvoyaient à sa version aseptisée. James Moore, expert en stratégie économique, technologie et leadership, savait ce qu'il faisait.

« Il aura fallu des millions de personnes de par le monde pour forcer la Grande Muette à

décrire le mouvement antiguerre comme la "deuxième superpuissance"; il aura suffi d'une poignée de "weblogues" (8) référençant son article pour que celui-ci, en vertu de l'algorithme PageRank de Google, bénéficie d'une légitimité telle que sa définition inoffensive écrase toutes les autres, constate Andrew Orłowski. Si le moteur de recherche était votre principale vision du monde, vous auriez du mal à croire que l'expression "deuxième superpuissance" puisse signifier autre chose. Son sens original a quasiment disparu. »

Pour Andrew Orłowski, l'anecdote révèle que « Google n'est pas "authentique", mais "synthétique" » : une recherche ne renvoie pas vers la principale référence sur le sujet, mais vers son acception la plus largement référencée. Un phénomène amplifié par la législation sur le droit d'auteur, qui interdit la publication en ligne des écrits qu'il protège. Si une recherche sur Raoul Vaneigem renvoie bien vers certains de ses textes (plusieurs de ses livres étant disponibles en ligne), vous n'obtiendrez en revanche, pour la plupart des auteurs, qu'une invitation à acquérir leurs ouvrages et, dans le meilleur des cas, la recension d'un livre faite par un internaute. Une situation que l'on pourrait comparer à celle d'une bibliothèque qui, devant renoncer au prêt gratuit de ses ouvrages, ne pourrait plus proposer que les fiches de lecture rédigées par ses adhérents.

Cette absence d'écrits de référence favorise les prises de position sur ce terrain de jeu idéologique. Et, curieusement, ce pouvoir symbolique - cette capacité de faire prévaloir sa perception d'un sujet donné - est l'un des rares à échapper encore à la structure habituelle de distribution du pouvoir. L'idéologie dominante n'y est pas surreprésentée, bien au contraire : une recherche lancée sur le nom d'un ministre de l'intérieur adepte des charters vous orientera vers des associations de défense des sans-papiers ; interrogé sur un homme d'affaires, Google pourra négliger ses communiqués de presse et raviver le souvenir des scandales financiers auxquels il a été mêlé. En réalité, le pouvoir d'influence des différents acteurs dépend surtout de leur degré d'appropriation du réseau : il ne suffit pas de développer un site, il faut également être capable de tisser des liens avec les autres sites et d'obtenir une reconnaissance de « ceux qui comptent » sur le réseau.

Si beaucoup bénéficient ainsi en toute bonne foi de la reconnaissance de leurs écrits, d'autres savent exploiter savamment les faiblesses de l'outil. Certaines agences se font ainsi une spécialité de réaliser, pour le compte de divers lobbies, des sites d'information dont le contenu pourrait à première vue laisser penser à des dépêches d'agence. Cette objectivité apparente suffit souvent à leurrer l'internaute qui, croyant à une information sérieuse, pourra être tenté de la référencer sur son propre site... et donc de lui accorder un pouvoir symbolique dont elle pourra profiter par la suite.

Certains sujets sensibles, comme les organismes génétiquement modifiés ou le conflit israélo-palestinien, font ainsi l'objet d'une lutte acharnée dans laquelle chacun s'efforce de rendre son idéologie « plus légitime » aux yeux de Google. Au point que le responsable d'un site Internet de référence - aux yeux de ce moteur de recherche tout au moins - a récemment eu la surprise d'être démarché par un intermédiaire commercial d'un genre

un peu particulier : « Je suis intéressé par l'achat de liens hypertexte sur votre site afin de promouvoir les sites de nos clients. Ces liens n'ont pas besoin d'être spécialement pris en évidence sur votre site, dans la mesure où nous n'en attendons pas de retombées directes en termes de visites. Votre site étant apprécié des moteurs de recherche, ces liens permettraient d'augmenter la visibilité de ces sites sur les moteurs. » Parmi ses clients, cet expert en renommée mentionnait sites financiers, agences de voyage et entreprises pharmaceutiques.

C'est sans doute face à toutes les questions à caractère politique, sur lesquelles des points de vue radicalement différents coexistent sur la Toile, que Google montre vraiment ses limites : ses critères mathématiques peuvent privilégier de facto certaines opinions et accorder une pertinence indue à des écrits ne reflétant que l'opinion de quelques-uns. L'assise et la surreprésentation dont ont bénéficié les « premiers arrivés » sur le Réseau, la densité des liens qu'ils entretiennent (notamment à travers le phénomène essentiellement américain des weblogs), désignent - mathématiquement - les « maîtres à penser » actuels de Google. Certes, l'outil a brillamment fait ses preuves pour les questions techniques ou pratiques. Mais il est des domaines où la pertinence échappe aux algorithmes.

PIERRE LAZULY.

(1) Francis Pisani, « Ma vie Google », *netsurf*, <http://www.netsurf.ch/archives-2001/0110/Ol1024n.html>

(2) « Security Concerns Prompt Army to Review web Sites Access », *Defense Information and Electronics Report*, 26 octobre 2001, <http://www.fas.org/spp/news/2001/10/dier102601.html>

(3) *San Francisco Chronicle*, 5 octobre 2001.

(4) *The News Media and the Laïe*, revue du Reporters Committee for Freedom of the Press, automne 2002, <http://www.rcfp.org/news/mag/cgi?26-4/foi-stateage>

(5) Jean-Pierre Cloutier, « Crise : sites Web censurés, modifiés, amendés », *Les (1) coniques de Cybérie*, 30 octobre 2001, <http://cyberie.gc.ca/chronik/20011030.html>

(6) Andrew Orlowski, « Anti-war slogan coined repurposed and Googlewashed... in 42 days », *The Register*, 4 mars 2003, <http://www.theregister.co.uk/content/6/30087.html>. Traduction française <http://www.uzjine.net/article1960.html>

(7) James F. Moore, « The Second Superpower Rears its Beautiful Head », <http://cyber.law.harvard.edu/people/jmoore/secondsuperpower.html>

(8) Internaute disposant d'un site personnel sur lequel il publie ses propres écrits, Lire Francis Pisani, « Internet saisi par la folie des weblogs », *1.1 monde diplomatique*, août 2003.

Monopoly ?

DANS L'UNIVERS DE LA NOUVELLE ÉCONOMIE, GOOGLE EST UN OVNI. C'est une start-up qui a réussi. Alors que la majorité d'entre elles n'ont pas survécu à l'éclatement de la bulle Internet, Google a su trouver le plus rapidement une viabilité financière. Son chiffre d'affaires, estimé à près de 1 milliard de dollars pour l'année 2003 (1), repose sur la vente de deux types de services. Sa technologie de recherche, tout d'abord, qu'elle fournit clé en mains à des clients tels que Yahoo !, Cisco, Vodafone ou l'armée américaine. Mais aussi, de plus en plus, son expertise en matière de publicité ciblée : les messages publicitaires affichés sur Google - de simples textes, mais correspondant précisément à la recherche effectuée - sont en effet en moyenne cinq fois plus cliqués que les bandeaux de publicité traditionnels.

Alors que tous y voient le candidat idéal pour une introduction en Bourse, Google refuse encore d'entrer à Wall Street, persuadé que ses ambitions technologiques seraient contrariées par une gestion à court terme imposée par les actionnaires. « Nous sommes

plutôt bons, mais nous sommes encore loin d'être parfaits. Nous ne le serons pas avant un long moment », affirme Larry Page (2). Peut-être rêve-t-il aussi de faire de Google un nouveau Microsoft et de ne s'introduire en Bourse que lorsqu'il disposera d'un quasi-monopole. Comme une très longue partie de Monopoly.

(1) « Google Announces Corporate Search Customers ». Reuters. 12 août 2003. (2) « Inside Google », *San Jose Mercury News*, 4 mai 2003.