

## Quand les mots valent de l'or

PAR FRÉDÉRIC KAPLAN

Le succès de Google tient en deux algorithmes : l'un, qui permet de trouver des pages répondant à certains mots, l'a rendu populaire ; l'autre, qui affecte à ces mots une valeur marchande, l'a rendu riche. La première de ces méthodes de calcul, élaborée par MM. Larry Page et Sergey Brin alors qu'ils étaient encore étudiants en thèse à l'université Stanford (Californie), consistait en une nouvelle définition de la pertinence d'une page Web en réponse à une requête donnée. En 1998, les moteurs de recherche étaient certes déjà capables de répertorier les pages contenant le ou les mots demandés. Mais le classement se faisait souvent de façon naïve, en comptabilisant le nombre d'occurrences de l'expression cherchée. Au fur et à mesure que la Toile s'étendait, les résultats proposés aux internautes étaient de plus en plus confus. Les fondateurs de Google proposèrent de calculer la pertinence de chaque page à partir du nombre de liens hypertextes pointant vers elle — un principe inspiré de celui qui assure depuis longtemps la reconnaissance des articles académiques. Plus le Web grandissait, plus l'algorithme de MM. Page et Brin affinait la précision de ses classements. Cette intuition fondamentale permit à Google de devenir, dès le début des années 2000, la première porte d'entrée du Net.

Alors que bien des observateurs se demandaient comment la société californienne allait pouvoir monétiser ses services, c'est l'invention d'un second algorithme qui a fait d'elle l'une des entreprises les plus riches du monde. A l'occasion de chaque recherche d'internaute, Google propose en effet plusieurs liens, associés à de courtes publicités textuelles, vers des sites d'entreprises. Ces annonces sont présentées avant les résultats de la recherche proprement dits. Les annonceurs peuvent choisir les expressions ou mots-clés auxquels ils souhaiteraient voir leur publicité associée. Ils ne paient que lorsqu'un internaute clique effectivement sur le lien proposé pour accéder à leur site. Afin de choisir quelles publicités afficher pour une requête donnée, l'algorithme propose un système d'enchères en trois étapes :

— L'enchère sur un mot-clé. Une entreprise choisit une expression ou un mot, comme « vacances », et définit le prix maximum qu'elle serait prête à payer si un internaute arrivait chez elle par ce biais. Pour aider les acheteurs de mots, Google fournit une estimation du montant de l'enchère à proposer pour avoir de bonnes chances de figurer sur la première page de résultats. Les acheteurs peuvent limiter leur publicité à des dates ou à des lieux spécifiques. Mais attention : comme on va le constater, le fait d'avoir l'enchère la plus haute ne garantit pas que vous serez le premier sur la page.

— Le calcul du score de qualité de la publicité. Google attribue à chaque annonce, sur une échelle de un à dix, un score, en fonction de la pertinence de son texte au regard de la requête de l'utilisateur, de la qualité de la page mise en avant (intérêt de son contenu et rapidité de chargement) et du nombre moyen de clics sur la publicité. Ce score mesure à quel point la publicité fonctionne, assurant à la fois de bons retours à l'annonceur et d'imposants revenus à Google, qui ne gagne de l'argent que si les internautes choisissent effectivement de cliquer sur le lien proposé. L'algorithme exact qui établit ce score reste secret, et modifiable à loisir par Google.

— Le calcul du rang. L'ordre dans lequel les publicités apparaissent est déterminé par une formule relativement simple : le rang est l'enchère multipliée par le score. Une publicité ayant un bon score peut ainsi compenser une enchère plus faible et arriver devant. Google optimise ici ses chances que l'internaute clique sur les publicités proposées.

Ce jeu d'enchères est recalculé pour chaque requête de chaque utilisateur — des millions de fois par seconde ! Ce second algorithme a rapporté à la firme de Mountain View la coquette somme de 9,72 milliards de dollars pour le troisième trimestre 2011 — un chiffre en croissance de 33% par rapport à la même période de l'année 2010 (1).

Le marché linguistique ainsi créé par Google est déjà global et multilingue. A ce titre, la Bourse des mots qui lui est associée donne une indication relativement juste des grands mouvements sémantiques mondiaux. La société propose d'ailleurs des outils simples et ludiques pour explorer une partie des données qu'elle collecte sur l'évolution de la valeur des mots. C'est ainsi que nous pouvons voir comment les fluctuations du marché sont marquées par les changements de saison (les mots « ski » et « vêtements de montagne » ont plus de valeur en hiver, « bikini » et « crème solaire » en été). Les flux et les reflux de la valeur du mot « or » témoignent de la santé financière de la planète. L'entreprise gagne évidemment beaucoup d'argent sur les mots pour lesquels la concurrence est forte (« amour », « sexe », « gratuit »), sur les noms de personnes célèbres (« Picasso », « Freud », « Jésus », « Dieu »), mais également dans des domaines de langue où la spéculation est moindre. Tout ce qui peut être nommé peut donner lieu à une enchère.

Google a réussi à étendre le domaine du capitalisme à la langue elle-même, à faire des mots une marchandise, à fonder un modèle commercial incroyablement profitable sur la spéculation linguistique. L'ensemble de ses autres projets et innovations technologiques — qu'il s'agisse de gérer le courrier électronique de millions d'utilisateurs ou de numériser l'ensemble des livres jamais publiés sur la planète — peuvent être analysés à travers ce prisme. Que craignent les acteurs du capitalisme linguistique ? Que la langue leur échappe, qu'elle se brise, se « dysorthographie », qu'elle devienne impossible à mettre en équations. Quand le moteur de recherche corrige à la volée un mot que vous avez mal orthographié, il ne fait pas que vous rendre service : le plus souvent, il transforme un matériau sans grande valeur (un mot mal orthographié) en une ressource économique directement rentable. Quand Google prolonge une phrase que vous avez commencé à taper dans la case de recherche, il ne se borne pas à vous faire gagner du temps : il vous ramène dans le domaine de la langue qu'il exploite, vous invite à emprunter le chemin statistique tracé par les autres internautes. Les technologies du capitalisme linguistique poussent donc à la régularisation de la langue. Et plus nous ferons appel aux prothèses linguistiques, laissant les algorithmes corriger et prolonger nos propos, plus cette régularisation sera efficace.

Pas de théorie du complot : l'entreprise n'entend pas modifier la langue à dessein. La régularisation évoquée ici est simplement un effet de la logique de son modèle commercial. Pour réussir dans le monde du capitalisme linguistique, il faut cartographier la langue mieux que n'importe quel linguiste ne sait le faire aujourd'hui. Là encore, Google a su construire une stratégie innovante en développant une intimité linguistique sans précédent avec ses utilisateurs. Nous nous exprimons chaque jour un peu plus au travers d'une des interfaces de la société ; pas simplement lorsque nous faisons une recherche, mais aussi quand nous écrivons un courrier électronique avec Gmail ou un article avec Google Docs, quand nous signalons une information sur le réseau social Google +, et même oralement, à travers les interfaces de reconnaissance vocale que Google intègre à ses applications mobiles. Nous sommes des millions chaque jour à écrire et à parler par son biais. C'est pourquoi le modèle statistique multilingue qu'il affine en permanence et vers lequel il tente de ramener chaque requête est bien plus à jour que le dictionnaire publié annuellement par nos académiciens. Google suit les mouvements de la langue minute par minute, car il a le premier découvert en elle un minerai d'une richesse extraordinaire, et s'est doté des moyens nécessaires pour l'exploiter.

La découverte de ce territoire du capitalisme jusqu'ici ignoré ouvre un nouveau champ de bataille économique. Google bénéficie certes d'une avance importante, mais des rivaux, ayant compris les règles de cette nouvelle compétition, finiront par se profiler. Des règles finalement assez simples : nous quittons une économie de l'attention pour entrer dans une économie de l'expression. L'enjeu n'est plus tant de capter les regards que de médiatiser la parole et l'écrit. Les gagnants seront ceux qui auront pu développer des relations linguistiques intimes et durables avec un grand nombre d'utilisateurs, pour modéliser et infléchir la langue, créer un marché linguistique contrôlé et organiser la spéculation sur les mots. L'utilisation du langage est désormais l'objet de toutes les convoitises. Nul doute qu'il ne faudra que peu de temps avant que la langue elle-même s'en trouve transformée.

FRÉDÉRIC KAPLAN

Chercheur à l'École polytechnique fédérale de Lausanne, auteur de *La Métamorphose des objets*, FYP Editions, Limoges, 2009, et, avec Georges Chapouthier, de *L'Homme, l'Animal et la Machine*, CNRS Editions, Paris, 2011.

(1) « Google Q5 2011 : \$9.72 billion in revenue, \$2.73 billion in net income [http://techcrunch.com/2011/10/13/google-q5-9-72-billion-in-revenue-2-73-billion-in-net-income/] », TechCrunch, 13 octobre 2011.

---

Mot clés: Langue Internet Commerce Marketing Entreprise Publicité Capitalisme Technologie Informatique Technologies de l'information

TRADUCTIONS DE CET ARTICLE

ESPAÑOL [Cuando las palabras valen oro \(es\)](#)