

Google : Historique et fonctionnement

Sommaire

I. Des origines à l'algorithme.....	1
I.1 Aux origines : Larry Page et Serguey Brin.....	1
I.2 L'algorithme de Google : le PageRank.....	2
II. Des services Google.....	2
II.1 Google Scholar ou rendre universellement accessible l'ensemble des connaissances disponibles.....	3
II.2 Google Books ou numériser l'ensemble des ouvrages de la planète.....	3
III. Le modèle Google.....	4
III.1 Une immense agence publicitaire.....	4
III.2 Une position hégémonique.....	5

I. Des origines à l'algorithme

I.1 Aux origines : Larry Page et Serguey Brin

Le moteur Google fut initialement conçu à Stanford par Serguey Brin et Larry Page, qui avaient alors pour projet d'indexer le fonds de leur bibliothèque universitaire.

Olivier ERTZSCHEID, 2005, p. 6

L'aventure Google commence en 1996 à l'université de Stanford, à l'époque où deux doctorants Larry Page et Sergeuy Brin se réunissent dans le cadre d'un projet de recherche pour travailler sur un moteur d'analyse des liens hypertextes¹. Le principe de ce moteur est de « déterminer le nombre de liens établis vers un site internet afin d'en déduire son importance relative »².

Les deux étudiants partent du postulat suivant : si un site dispose de nombreux liens externes pointant vers lui, c'est qu'il est reconnu comme étant une référence dans son domaine [...] ils développent alors un robot d'indexation (Googlebot) qui va parcourir la structure du web pour en étudier le maillage³.

Le principe algorithmique de Google, dénommé PageRank, est « la transposition au web des modèles bibliométriques » développés par Eugene Garfield dans les années 1960 pour des corpus de publications scientifiques⁴.

1 ROUSSEL, Guillaume, 2006, p. 3

2 ROUSSEL, Guillaume, 2006, p. 3

3 ROUSSEL, Guillaume, 2006, p. 3

4 ERTZSCHEID, Olivier, 2005, p. 6

I.2 L'algorithme de Google : le PageRank

Une des clés du succès du PageRank peut être vu dans le fait qu'il a été dès le départ conçu de façon à ne pas imposer de schéma organisationnel au web. Il a su épouser les premiers idéaux d'Internet en se présentant comme ne faisant rien d'autre que de rendre aux internautes le fruit de leurs actions et choix.

Cléo COLLOMB, 2015, p. 9

Dans l'article scientifique où ils exposent pour la première fois le principe du PageRank, « les fondateurs de Google indiquent qu'en sus du mode de comptabilité des liens entrants, chaque internaute agit comme un indexeur à chaque fois qu'il décide de créer un lien hypertexte »⁵. Le PageRank constitua, à l'époque de sa mise en œuvre, une véritable révolution technologique⁶.

Le référencement sur Google dépend du nombre et de l'importance des citateurs : chaque lien vers une page web est un vote [...] et si la firme de Mountain View se targue de faire appel aux contributions de la grande communauté d'Internet, c'est parce qu'elle se défend, dit-elle, d'intervenir stratégiquement et d'imposer un schéma organisationnel au web [...] À suivre les discours de Google, personne n'intervient donc stratégiquement ni manuellement pour ordonnancer le web : son organisation ne doit faire que refléter les actions et interactions des internautes⁷

Le moteur de recherche Google peut ainsi par exemple être compris comme une plateforme visant à organiser, capturer et valoriser l'intelligence de la foule⁸. Mais avec le PageRank, se dessine surtout les contours d'une algorithmie « mettant sous la coupe de la puissance calculatoire des moteurs, la moindre de nos interactions en ligne, le moindre de nos comportements connectés, la plus infime trace de nos plus éphémères conversations »⁹.

II. Des services Google

Un moyen utilisé par Google pour maintenir sa suprématie est « d'innover perpétuellement par le lancement de nouveaux services en ligne en s'appuyant sur sa technologie de recherche »¹⁰. La firme de Mountain View équipe ainsi ses activités relationnelles (Gmail, Drive, Youtube) augmentant « son capital soin, confiance et affects »¹¹. « Par son cycle incessant d'innovations, le moteur transforme les internautes en usagers-concepteurs, « ce qui est emblématique d'une nouvelle culture »¹²

5 ERTZSCHEID, Olivier, 2008, p. 15

6 ERTZSCHEID, Olivier, 2008, p. 7

7 COLLOMB, Cléo, 2015, p. 9

8 COLLOMB, Cléo, 2015, p. 14

9 ERTZSCHEID, Olivier, 2008, p. 13

10 ROUSSEL, Guillaume, 2006, p. 7

11 COLLOMB, Cléo, 2015, p. 8

12 GALLETZOT, Gabriel, 2009, p. 3

II.1 Google Scholar ou rendre universellement accessible l'ensemble des connaissances disponibles

Car pour Scholar comme pour l'ensemble des services déployés, Google se sert de son moteur comme d'un attracteur permettant d'amener un public captif considérable vers de nouveaux services.

Olivier ERTZSCHEID, 2005, p. 6

Google Scholar est un service qui met à disposition (en texte intégral ou non) « des ouvrages et des articles scientifiques, parfois numérisés par Google lui-même, parfois « récupérés » dans des entrepôts existants (archives ouvertes) »¹³. Le fait est indéniable, une écrasante majorité d'universitaires utilisent le moteur Google pour leurs recherches d'articles scientifiques¹⁴.

La première motivation de Google pour le secteur de la diffusion scientifique s'inscrit d'abord dans une logique capitaliste [...] La seconde tient à la stratégie et au positionnement de la firme [...] Si l'on peut considérer comme gagnée – par Google – la bataille de l'information (aussi bien en termes d'accès, de diffusion, que de modèle économique, publicitaire associé), la prochaine bataille à remporter est celle de la connaissance¹⁵

De manière générale, les ambitions de Google s'accroissent mal d'humilités¹⁶. Pour Eric Schmidt, l'un des ses dirigeants, « le rôle de Google « rendre l'information mondiale accessible et utile », c'est à dire organiser, hiérarchiser et diffuser les différentes données à sa disposition (pages web, médias mais aussi annonces publicitaires) à l'aide de programmes et de formules de calculs¹⁷.

II.2 Google Books ou numériser l'ensemble des ouvrages de la planète

L'énorme atout de Google Books est d'être un simple appendice du moteur Google

Alexandre MOATTI, 2012, p. 4

Pour Google, présenter le catalogue le plus large possible « est d'abord l'occasion de conforter une position déjà outrageusement dominante de guichet unique d'accès à l'information, puis à la connaissance »¹⁸. Annoncé le 15 décembre 2004 par Google, le projet Google Books vise la numérisation de 15 millions de livres en six ans en partenariat avec les plus importantes bibliothèques américaines¹⁹. Initialement et bien maladroitement dénommé Google Print ; le service Google Books englobe ainsi la numérisation d'ouvrages encore sous droits et concerne l'ensemble

13 ERTZSCHEID, Olivier, 2006, p. 3

14 ERTZSCHEID, Olivier, 2005, p. 2

15 ERTZSCHEID, Olivier, p. 4

16 ERTZSCHEID, Olivier, p. 5

17 ROUSSEL, Guillaume, p. 6

18 ERTZSCHEID, Olivier, 2006, p. 1

19 MOATTI, Alexandre, 2012, p. 3

de la politique de numérisation de Google²⁰. Il ne faut d'ailleurs pas oublier que Google développe sa stratégie commerciale à destination des piliers de la culture scientifique que sont les bibliothèques²¹.

Face aux ambitions de Google, le vibrant plaidoyer lancé par Jean-Noël Jeannenay, à l'époque président de la BNF, en faveur d'une bibliothèque numérique européenne soulevait un grand enthousiasme²². Mais le projet Europeana, censé contrer ce qui était vu comme une hégémonie culturelle de Google dans le domaine de la diffusion de connaissances sur Internet a débouché plus sur un catalogue disparate de médiathèques que sur un service centralisé de bibliothèque²³. Sur Google Books est appliqué la recette du moteur de recherche : simplicité et sobriété²⁴. Finalement, l'une - Europeana - est un site de passage, où l'on est attiré par curiosité, alors que l'autre est devenu un outil de travail régulier²⁵.

La stratégie de Google, y compris dans le projet Books est fondée sur l'idée que les moteurs sont - et seront toujours à l'avenir - la porte d'accès à Internet, la difficulté d'Europeana s'explique parce qu'un site ne peut en aucun cas rivaliser avec un moteur indexant des millions de sites²⁶

III. Le modèle Google

III.1 Une immense agence publicitaire

Le modèle économique de Google est celui d'une agence média. Plus de 80 % de ses revenus sont directement issus de son offre publicitaire (liens sponsorisés).

Olivier ERTZSCHEID, 2006, p. 4

Comme sur Facebook, la gratuité des services Google pour les particuliers est compensée par la gigantesque manne publicitaire offerte par les professionnels dont l'objectif premier est d'être bien référencé. Or, dans une économie numérique au sein de laquelle le web s'affirme comme un média de masse, la « monétisation de l'ensemble des services offerts est inéluctable »²⁷. Le modèle publicitaire adopté est un succès puisqu'il permet à Google de générer d'incroyables bénéfices : « la société devient bénéficiaire pour la première fois en 2001 avec 7 millions de dollars de bénéfices »²⁸.

Ce modèle économique présente un avantage par la diversité de la clientèle. Ce système d'achat de liens commerciaux est utilisé aussi bien par de très petites structures que par de grands groupes. Google ne dépend donc pas de

20 ERTZSCHEID, Olivier, 2006, p. 3

21 ERTZSCHEID, Olivier, 2005, p. 6

22 MOATTI, Alexandre, 2012, p. 3

23 MOATTI, Alexandre, 2012, p. 4

24 MOATTI, Alexandre, 2012, p. 5

25 MOATTI, Alexandre, p. 4

26 MOATTI, Alexandre, 2012, p. 5

27 ERTZSCHEID, Olivier, 2006, p. 4

28 ROUSSEL, Guillaume, 2006, p. 3

quelques gros clients mais bien d'une multitude de sociétés de toutes tailles²⁹.

Si « beaucoup reconnaissent que la société de Mountain View a su imposé un nouveau modèle en matière de communication d'entreprise », ce la s'est traduit par un double discours qui oscille « entre culture d'entreprise et valeurs culturelles universelles »³⁰. D'un côté, la confiance, la liberté des internautes sont érigées en valeurs, mais de l'autre le discours affiché par Google procède bien « d'une rhétorique publicitaire masquée faite de glissement de sens »³¹.

Il faut aussi avoir à l'esprit que, désormais, « ce qui fait le prix inestimable de la firme de Mountain View, ce sont moins les publicités qu'elle vend que les résultats aux requêtes des internautes qu'elle ne vend pas »³². Aujourd'hui, ce qui vaut de l'or, « ce n'est pas tant ce que l'on parvient à vendre aux internautes moyennant des publicités toujours plus ciblées, que les données qu'ils fournissent eux lorsqu'ils agissent, désirent, contribuent, partagent, vivent³³. Les données personnelles représentent en quelque sorte le pétrole du XXI^e siècle.

III.2 Une position hégémonique

L'objectif principal de Google est de rester le leader sur le marché de la recherche d'information. C'est la base même de son succès et perdre cet avantage concurrentiel pourrait s'avérer fatal pour le moteur de recherche.

Guillaume ROUSSEL, 2006, p. 5

« En quelques années seulement, Google est devenu une véritable passerelle entre les ressources Internet et les internautes s'imposant à la fois comme la référence d'accès à l'information et un exemple de réussite économique³⁴. Le leader mondial de la recherche a su ainsi conquérir un marché stratégique et concurrentiel : « l'accès à l'information sur le web »³⁵. Évidemment, sa position hégémonique dans le paysage occidental, sa capacité à capter et à monétiser des pans entiers de nos activités et de nos vies numériques, en font un dispositif qui mérite d'être interrogé³⁶. « Le marché du Search » comme l'appelle les américains est ainsi outrageusement dominé par Google, avec très loin derrière lui, Yahoo ! Et Microsoft, en France comme outre-atlantique »³⁷.

Néanmoins, face à l'omniprésence du moteur de recherche, des alternatives existent : Qwant, DuckDuck Go, IxQuick. L'association Framasoft a réalisé une suite de logiciels (Framadate, Framacalc, etc.) qui se présente comme autant d'alternatives libres et éthiques aux services de Google.

29 ROUSSEL, Guillaume, 2006, p. 5

30 GALLETZOT, Gabriel, 2009, p ; 4

31 GALLETZOT, Gabriel, 2009, p.

32 COLLOMB, Cléo, 2015, p. 8

33 COLLOMB, Cléo, 2015, p. 8

34 ROUSSEL, Guillaume, 2006, p. 2

35 ROUSSEL, Guillaume, 2006, p. 2

36 ERTZSCHEID, Olivier, 2008, p. 3 et GALLETZOT, Gabriel, 2009, p. 1

37 ERTZSCHEID, Olivier, 2008, p. 7

Bibliographie

- Collomb, Cléo. « Google, une entreprise de mise en foule de la multitude », novembre 2015. <https://hal.archives-ouvertes.fr/hal-01253751>.
- . « Un communisme numérique ». In *Thémat'IC 2014 Communautés en ligne et constructions identitaires*, édité par IUT Robert Schuman, 13-29. Strasbourg, France: Patricia Caillé, 2014. <https://hal.archives-ouvertes.fr/hal-01290140>.
- Ertzscheid, Olivier. « D'un numérique l'autre : des moteurs, des libraires, et des usages. » *cahiers de la librairie*, n° 5 (1 mai 2006): 33-42.
- . « Google Scholar, Google Print ... un nouvel entrant pour la diffusion de la recherche. » *Perspectives documentaires en éducation* 62 (2005). https://archivesic.ccsd.cnrs.fr/sic_00090140/document.
- . « Moteurs de recherche : des enjeux d'aujourd'hui aux moteurs de demain. », octobre 2008, 59-89.
- Gallezot, Gabriel, et Brigitte Simonnot. « Introduction (au livre l'entonnoir) : Du flou cognitif au flou sémiotique : ce qui se joue dans l'entonnoir de Google. », 2 juin 2009, 19-27.
- Moatti, Alexandre. « Bibliothèque numérique européenne : de l'utopie aux réalités ». *Réalités industrielles. Annales des mines*, novembre 2012, 43-46.
- Roussel, Guillaume. « Google, ou comment s'imposer comme un point de passage obligé ». In *Conférence DocSoc 2006 - Semaine du document numérique*. Fribourg (Suisse): ADBS éditions, 2006. https://archivesic.ccsd.cnrs.fr/sic_00110306.