

Aux Etats-Unis, Google modifie son algorithme en réponse aux webspams

Description

En communiquant sur une modification « majeure » de son algorithme de recherche, baptisée Panda, Google a voulu rappeler qu'il était toujours à la pointe de l'innovation dans la recherche, où la concurrence est de plus en plus forte. En même temps, cette modification, nécessaire face aux webspams, témoigne de la manière dont les résultats du moteur de recherche influent sur les contenus du Web.

En 2010, les critiques se sont multipliées contre le moteur de recherche Google, notamment parce qu'il accorderait dans ses résultats trop de place à des contenus sans grande valeur. En même temps, Google est de plus en plus concurrencé, sur le marché des moteurs de recherche, l'alliance entre Bing et Yahoo! étant désormais effective aux Etats-Unis, mais également comme porte principale d'Internet, où Facebook rivalise avec lui en termes d'audience. Pour Google, améliorer les performances de son moteur de recherche pour satisfaire les internautes est donc un impératif, lequel est d'autant plus impérieux que Google réalise l'essentiel de ses revenus grâce à la vente des liens sponsorisés. En annonçant, le 24 février 2011 sur The Official Google Blog, dans un post d'Amit Singhal et Matt Cutts, procéder à un changement profond de son algorithme, Google a ainsi confirmé la nécessité pour le moteur de proposer des résultats d'un genre nouveau.

En même temps, cette évolution est liée au succès de Google qui reste à ce jour le principal outil de navigation sur Internet, à tel point que de nombreux sites ont cherché à en exploiter les failles pour récupérer facilement des audiences grâce à un bon référencement. En effet, pour Google, la modification de l'algorithme aura pour effet de « rétrograder les sites de mauvaise qualité » : « les sites qui apportent peu de valeur pour les internautes, qui copient des contenus d'autres sites, ou qui ne sont tout simplement pas très utiles ». Autant dire que cette modification vise les sites qui se sont construits pour être référencés par Google, les agrégateurs de contenus étant précisément cités. Les fermes de contenus qui produisent des contenus en fonction des mots clés les plus recherchés par les internautes peuvent également entrer dans cette catégorie, même si elles ne sont pas explicitement visées par Google (voir REM n°17, p.54). Google envisage en fait de redonner la priorité aux contenus exclusifs et aux sites apportant une certaine valeur ajoutée, « ceux qui ont des contenus originaux, et des informations comme de la recherche, des études détaillées, des analyses réfléchies, etc. ». Enfin, le changement de l'algorithme a pour conséquence de modifier le classement de 11,8 % des sites référencés, certains devant être rétrogradés quand d'autres seront promus.

La mesure, qui s'applique pour l'instant aux Etats-Unis, devrait être étendue au monde entier où Google contrôle, le plus de 66 % des recherches. Sur le marché américain, elle permet à Google de

s'affirmer face à Bing, mais face également à d'autres moteurs de recherche ayant développé une stratégie originale, comme Blekko qui, sensible aux critiques adressées par les internautes contre les *webspams*, – ces sites bien référencés mais de peu d'intérêt, – a opté pour filtrer les fermes de contenus de ses résultats. Ce «*changement majeur*» de l'algorithme de Google a suscité l'inquiétude parmi les fermes de contenus, au premier rang desquelles Demand Media aux Etats-Unis, mais également chez Yahoo! ou AOL, qui misent tous les deux sur la production massive de contenus pour capter une audience leur permettant d'assurer des débouchés à leurs activités de régie. Quelques jours après l'entrée en vigueur de l'algorithme, les résultats étaient contrastés. Demand Media, l'une des plus grandes fermes de contenus aux Etats-Unis, constatait que son audience ne chutait guère, celle-ci étant même en progrès. Elle a sans doute bénéficié du déficit de visibilité constaté pour certains de ses concurrents. D'autres ont vu, en effet, leur référencement dans Google se dégrader et leur audience chuter, le moteur de recherche étant le principal apporteur d'audience des fermes de contenus et des agrégateurs : selon la société allemande Sistrix, qui étudie le trafic généré par les mots clés, la ferme Suite101, pourtant réputée pour la qualité de ses contributions, ou encore AssociatedContent de Yahoo! voyaient leur référencement chuter de plus de 90 % dans les résultats de Google. Ce changement d'algorithme aura sans doute pour effet d'entraîner une reorganisation des fermes de contenus pour lesquelles retrouvent, dans Google, la place privilégiée qu'elles étaient parvenues à occuper.

Parce que les réactions ont été nombreuses à la suite de la modification de l'algorithme de Google, le moteur de recherche a par ailleurs communiqué, deux mois plus tard, une liste de conseils pour les éditeurs de sites, qui devrait les aider à être bien référencés si les contenus qu'ils mettent en ligne sont de bonne qualité. Parmi ces critères, la confiance accordée à l'information est citée en premier, ce qui donne ici un avantage aux marques de presse par rapport aux contributions d'inconnus sur les agrégateurs et les forums. La duplication du même contenu sur le même site ou sur plusieurs sites est considérée comme l'indication d'une moins grande valeur de l'information, les articles produits par les fermes de contenus étant notamment publiés sur de très nombreux sites. Enfin, d'autres critères comme le nombre de signatures par site sont pris en compte : un journal en ligne aura quelques signatures de journalistes, un agrégateur aura presque autant de signatures que d'articles. Là encore, les sites de presse en ligne semblent favorisés.

Enfin, Amit Singhal et Matt Cutts précisent, dans le blog Google, que cette modification de l'algorithme n'est pas liée au programme Blocklist du navigateur Chrome, lequel propose à ses utilisateurs de bannir les sites indésirables des pages de résultats. Ils notent également que les sites bannis par les internautes sont également ceux que le nouvel algorithme retrograde, mais cette fois-ci de manière neutre et automatique. La corrélation vient ici confirmer le lien étroit entre les performances de la technique et les attentes des utilisateurs. Elle rappelle en même temps que l'accès aux contenus est décidé par un algorithme qui, même s'il évolue, pourrait bien avoir déjà eu le temps d'habituer les internautes à souhaiter et à écarter certains sites.

Sources :

- « Finding more high-quality sites in search », The Official Google Blog, post de Amit Singhal et Matt Cutts, www.google.com, 24 février 2011.
- « Google part à la chasse des « pollueurs du Web » », www.france24.fr, 25 février 2011.
- « Google opère un nouveau réglage pour retrograder les sites de mauvaise qualité », *La Correspondance de la presse*, 28 février 2011.
- « Les copieurs n'ont plus droit de citer sur Google », Sandrine Cassini *La Tribune*, 28 février 2011.
- « Google modifie son moteur pour faire le ménage sur la toile », M.Q. *Les Echos*, 28 février 2011.
- « Google veut enrayer la percée des sites de mauvaise qualité », Cécile Ducourtieux et Xavier Ternisien, *Le Monde*, 8 mars 2011.
- « More guidance on building high-quality sites », Amit Singhal, Google Webmaster Central Blog, May 6, 2011.

Categorie

1. Ailleurs
2. Repères & tendances

date créée

20 mars 2011

Auteur

alexandrejoux