

Déréférencement : l'avis de Google

[Retour au sommaire de la lettre](#)

Nous en avons parlé le mois dernier, et la nouvelle a parcouru le microcosme du référencement, Google a dernièrement « blacklisté » un certain nombre de sites appartenant à des sociétés de référencement françaises. Le mois dernier, divers acteurs du « milieu » nous ont donné leur avis à ce sujet. Ce mois-ci, c'est Matt Cutts, porte-parole technique de Google US, qui répond - avec un peu de retard indépendamment de notre volonté et dont nous nous excusons auprès de nos lecteurs - à nos questions.

Pourquoi Google a-t-il déréférencé des sociétés françaises de référencement, il y a quelques semaines de cela ? Quels sont les critères qui font que le moteur de recherche « blackliste » un site ? Jusqu'à maintenant, Google était resté très discret à ce sujet. Abondance a donc posé dix questions sur ce sujet à Matt Cutts, « software engineer » chez Google et très souvent porte-parole technique du moteur.

Matt Cutts a rejoint Google en tant que "Software Engineer" en janvier 2000. Auparavant, il travaillait à son "Ph.D. in computer graphics" à l'université de Caroline du Nord à Chapel Hill. Il a publié des articles sur le "bureau du futur" ainsi que sur la reconnaissance d'objet dans les images compressées. Matt a écrit l'appliquatif "Safe Search", le filtre familial utilisé par Google. Il travaille également pour le Ministère de la Défense américain et a également étudié des projets pour une société de jeux vidéo. Pour en savoir plus sur son parcours : <http://www.cs.unc.edu/~cutts/>.

Matt Cutts a accepté de répondre à nos questions (merci à lui) afin de clarifier la position du moteur au sujet du spamdexing. Une pierre supplémentaire à ajouter à l'édifice...



- Certains SEOs ont dernièrement disparu de l'index de Google, leur PageRank étant réduit à zéro. Est-ce une action délibérée de Google et pourquoi ?

La qualité de l'index et des résultats que nous fournissons à nos utilisateurs est de la plus haute importance pour nous. Google ne tolère pas le spam et encourage ses utilisateurs à dénoncer tout type de fraude qui pourrait nous aider à mettre en place de nouveaux algorithmes visant à bloquer toute tentative de ce type. De plus, lorsqu'un site met en place des tentatives continuelles et répétitives de spam, nous bannissons de notre index ses pages et / ou ses auteurs. Nous ne tolérons pas les sites qui s'engagent dans des procédures de spamming ou d'autres techniques visant à altérer leurs classements et à induire les internautes en erreur. Parfois cependant, des problèmes techniques, comme le fait qu'un site soit « down » lorsque nos robots désirent le visiter, peut amener une absence temporaire d'une source d'information dans nos résultats. Cependant, il est clair que toute violation de nos « quality guidelines » sera, à l'avenir, sanctionnée par un blacklisting dans l'index.

NDLR : le moins que l'on puisse dire est que la réponse de Matt Cutts est d'une clarté limpide :-)

Pour avoir accès aux « quality guidelines » de Google, allez aux urls suivantes :

- En français :

<http://www.google.fr/intl/fr/webmasters/guidelines.html>

<http://www.google.fr/intl/fr/webmasters/seo.html>

- En anglais :

<http://www.google.com/webmasters/guidelines.html>

<http://www.google.com/webmasters/seo.html>

- Tous les SEOs et sites web dernièrement bannis en France l'ont-ils été pour la même raison ?

Non, pas obligatoirement. A chaque fois que Google effectue une action sur un site, cela se fait selon une procédure de « cas par cas ». Mais nous ne pouvons pas parler ici de cas individuels, désolé.

- Pourquoi n'avoir pas privilégié la voie du dialogue avant l'action de blacklistage ?

Traditionnellement, Google n'informe jamais les sites qui sont identifiés comme spammeurs. Il y a plusieurs bonnes raisons à cela.

Premièrement, contacter chaque site ou webmaster qui met en place des techniques douteuses et leur demander de changer leurs habitudes est un travail énorme qui serait quasi impossible à mettre en œuvre.

D'autre part, certains spammeurs « malicieux » pourraient mettre à profit le fait d'être sûr à 100% d'avoir été pris « en flagrant délit » de violation des critères de notre politique de qualité.

Nous avons donc plutôt choisi d'afficher, sur notre site (NDLR : voir adresses ci-dessus), cette politique afin d'aider les webmasters à comprendre les actions principales qui nous amèneront à supprimer de notre index les sites qui mettent en place les techniques les plus « extrêmes ».

Sur notre site, certaines informations (notamment

<http://www.google.com/webmasters/guidelines.html#quality>) sont en ligne depuis plusieurs années, et indiquent clairement que, par exemple, tout texte caché, tout lien caché, toute technique de cloaking ainsi que des tentatives sournoises de redirection, notamment à base de Javascript, ayant pour but de montrer des contenus différents aux internautes et aux moteurs de recherche peuvent être susceptible de blacklistage.

- Beaucoup disent en France que des sociétés utilisant des techniques de spam bien plus importantes sont encore dans l'index de Google. Ces observateurs ressentent donc une certaine injustice. Que répondez-vous à cette remarque ?

Nous tentons de faire de notre mieux pour toujours augmenter et surveiller la qualité de notre index. Nous sommes toujours heureux de recevoir toute information ou suggestion qui nous permettraient d'augmenter la qualité de nos résultats et de réduire le spam dans notre index. Nous proposons, à cet effet, une page spécifique (<http://www.google.fr/contact/spamreport.html>) et une adresse e-mail (spamreport@google.com). Nous utilisons ces données à la fois pour améliorer l'algorithme automatique du moteur mais aussi pour traiter certains cas « à la main ».

- Comment s'effectue une opération de blacklistage chez Google ? Est-elle automatique ou manuelle ? Existe-t-il des "Brigades anti-spam" ?

Nous préférons développer des solutions automatisées et graduelles à ce type de problème, nous tentons donc de minimiser les interventions manuelles contre le spam. Les rapports que nous recevons (NDLR : voir réponse à la question précédente) nous permettent de créer des algorithmes qui reconnaissent et bloquent de futures tentatives de spam. Mais, dans certains cas, notamment de tentatives répétitives, nous pouvons être amenés à exclure de façon globale un site de l'index (NDLR : cela semble signifier que, sinon, seule la page incriminée peut être exclue, pas le site entier)

- Le blacklistage ou l'étude de cas particuliers est-il essentiellement basé sur la délation des internautes ou sur des actions automatiques ?

Des premières vérifications sont effectuées sur les sites suggérés par nos algorithmes, mais nous prenons clairement en compte les remarques de nos utilisateurs lorsqu'il s'agit d'établir les priorités sur les sites à étudier...

- Les remous provoqués par cette action n'incitent-ils pas Google à clarifier sa politique vis-à-vis du spam ?

La « politique qualité » de Google (<http://www.google.fr/webmasters/guidelines.html#quality>) est en ligne depuis des années et les « guidelines » que nous fournissons en ligne me semblent clairs. Comme je le disais précédemment, des actions comme le texte ou les liens cachés ou les redirections « sournoises » sont sans ambiguïté en contradiction avec notre ligne directrice. Plusieurs autres pages, sur notre site, sont disponibles, et notamment une section complète (<http://www.google.fr/webmasters/seo.html>) traitant des risques encourus dans le cadre de l'optimisation de pages web. Dans de nombreux cas, le fait de lire ces pages répondra à de nombreuses questions pour savoir si une technique est ou non dangereuse. Nous travaillons actuellement à proposer bientôt de nouvelles informations dans ce sens. Nous ne pouvons pas, en revanche, décrire nos algorithmes de détection et expliquer comment nous faisons pour identifier le spam, car ces informations pourraient immédiatement être mises à profit par des fraudeurs...

- Qu'en est-il des pages satellites, pourtant souvent utilisées par les référenceurs ?

Si une page est construite uniquement pour les moteurs de recherche et est "remplie" de mots clés, Google considérera qu'il s'agit de spam. Les pages satellites sont le plus souvent créées uniquement pour les moteurs de recherche, ce qui n'est pas réellement utile pour les internautes. Google considère que toute tactique employée avec l'intention de tromper l'internaute par le biais d'artifices est inappropriée.

- Beaucoup disent que ces blacklistages ont pour but, de la part de Google, d'attirer les clients éditeurs de sites vers l'achat d'AdWords plutôt que vers les prestations de SEOs. Quelle est la part de vérité dans ces affirmations ?

Non, Les Adwords n'avaient rien à voir dans cette action. Google prend toute tentative de dégradation de son index très au sérieux et doit mettre en place des actions plus ou moins strictes afin que la qualité de ses résultats reste la meilleure possible. Nous enlevons de notre index les sites qui tentent d'aller contre ce credo.

- La nomination de médiateurs, qui joueraient un rôle d'interface entre Google et les SEOs n'est-elle pas envisageable, voire indispensable, à l'avenir, notamment pour définir de façon claire des politiques de référencement / indexation de sites techniquement complexes ?

Il existe certainement des exemples positifs de sociétés de référencement qui fournissent des avis constructifs et aident les sites web à améliorer leur architecture pour faciliter le « crawling » par nos robots. Dans le même temps, Google s'efforce de faire en sorte que l'indexation d'un site soit la plus simple et efficace possible. Nous tentons également de proposer le plus d'informations possible sur notre site à cet effet. Si le propriétaire d'un site se sent submergé par les propositions des entreprises de référencement, nous fournissons un certain nombre de points à vérifier sur notre site (NDLR : voir url ci-dessus) avant de choisir son prestataire.

- Google serait-il ouvert, actuellement, à un dialogue avec les SEOs pour clarifier la situation ? Finalement, une plus grande ouverture des moteurs de recherche vers les SEOs ne serait-elle pas la solution au problème du spam ? Ne serait-il pas possible d'imaginer un "label qualité" décerné par Google à certaines sociétés qui utiliseraient des techniques "approuvées" par Google ?

Nous essayons par plusieurs voies de fournir de l'information à ce sujet : nos pages d'aides, sur notre site, notre weblog (<http://www.google.com/googleblog/>) et également le fait de répondre à certaines questions sur des forums américains comme Webmasterworld (<http://www.webmasterworld.com/>). Nous sommes toujours ravis de recevoir des suggestions, sur les forums, sur le Web ou par mail sur la façon d'améliorer la qualité des informations que nous fournissons à nos visiteurs. L'un des problèmes d'une procédure de labellisation est d'être certain que le référenceur suit à la lettre notre ligne directrice. Il est complexe d'être sûr qu'une compagnie est « propre » et qu'elle continue à l'être dans le temps. Mais Google est toujours ouvert à toute suggestion à ce sujet...

Merci Matt Cutts pour avoir répondu à nos questions. « Special thanks » à Myriam et Debbie pour leur aide.

