

## Les pénalités de Google en termes de positionnement

[Retour au sommaire de la lettre](#)

*Google doit, depuis sa création, se défendre contre des webmasters peu scrupuleux qui tentent de détourner ses algorithmes de pertinence. Pour cela, il a mis en place une arsenal de pénalité sur lesquelles il communique très peu, et pour cause... Connues sous le nom de "Sandbox", "Minus 30 penalty", "Position 6 penalty" ou "Blacklist", voici une vision d'ensemble des risques que vous courez si vous tentez d'être plus fort que les équipes de "search quality" de Google... Et comment en sortir si vous vous faites prendre...*

Depuis que Google existe, ce moteur de recherche pénalise, comme la plupart de ses confrères, les sites web qui tentent de le "spamindexer", ou en d'autres termes, de contourner ses algorithmes de pertinence pour tenter, grâce à moults techniques prohibées, de mieux se positionner dans les pages de résultats.

Au fil du temps, Google a mis en place une panoplie assez complète de pénalités, parfois "softs", mais également parfois très dures (liste noire) selon la gravité estimée de la faute commise...

Nous avons essayé, dans cet article, de répertorier les différentes pénalités imaginées par les ingénieurs de Google, sachant qu'il est complexe d'en parler avec exactitude puisque Google n'a quasiment jamais communiqué sur ce point. La plupart des informations comprises dans cet article sont donc issues de tests empiriques, de discussions dans des forums, d'avis d'experts, etc.

Sachez cependant que cet article a été transmis à Google et lu par les équipes qui gèrent ces pénalités. Nous n'avons pas eu de retour détaillé de leur part sur son contenu (Google ne communique quasiment jamais sur ce sujet) mais nous prenons également cela comme le fait qu'il n'y a pas d'erreur monumentale dans cet article, car, dans ce cas, nos correspondants auraient certainement jugé utile de nous en faire part...

Il n'en reste pas moins vrai qu'il est intéressant de comprendre comment fonctionne la "Spam Brigade" de Google et de bien voir que la seule façon de ne jamais avoir à faire à eux est bien de concevoir un site web pour les internautes, optimisé pour les moteurs de recherche, mais sans chercher à aller trop loin. Toute velléité de "suroptimisation" est donc risquée... N'oubliez pas que, même si vos techniques un peu "border line" échappent aux ingénieurs de Google (qui ont beaucoup de choses à faire et de sites à vérifier), vos concurrents risquent, eux de s'en apercevoir et se feront un plaisir de vous dénoncer auprès du moteur sur son formulaire de *spam report* (<http://www.google.com/contact/spamreport.html>)... Un webmaster averti en vaut toujours deux...

### **Techniques à ne pas employer**

Il est tout d'abord important de comprendre quelles techniques d'optimisation sont à éviter pour ne pas avoir à subir les foudres des moteurs de recherche. En voici quelques-unes :

- Pages satellites (alias, fantômes, *doorway pages*, etc. : pages conçues spécialement pour les moteurs de recherche et contenant une redirection vers le site "réel").
- Cloaking (action de fournir des documents différents à un internaute et à un *spider* par détection automatique de ces derniers).
- *Keyword stuffing* (répétition non naturelle de mots clés à l'intérieur d'une page).
- Contenu textuel et/ou liens cachés au sein du code HTML d'une page à des fins de référencement.
- Ajout dans une page, d'une façon ou d'une autre, de contenu n'ayant pas de rapport direct avec celui qui apparaît de façon visible dans la page.
- Redirection frauduleuse.
- Etc.

En règle générale, n'essayez pas de jouer "aux gendarmes et aux voleurs" avec les moteurs de recherche et relisez bien cette page intitulée "Conseils aux webmasters " sur l'aide en ligne de Google :

<http://www.google.com/support/webmasters/bin/answer.py?answer=35769>

Le moteur de recherche leader vient également de mettre en ligne un guide intitulé *Making the most of Your Content - A publisher's guide to the Web*

(<http://books.google.com/googlebooks/pdf/webmastertools.pdf>) que vous pouvez lire avec intérêt (il est en anglais mais une version française est disponible sur le blog du site Abondance - <http://blog.abondance.com/>).

Bref, il existe quelques "règles d'or" à suivre pour éviter tout problème de pénalité sur les moteurs de recherche. Les voici :

- 1. On ne cache rien (ce que l'internaute voit, le moteur le voit et vice et versa)**
- 2. Un site web est avant tout fait pour les internautes.**
- 3. Une optimisation de qualité pour les moteurs (Title, texte, liens, etc.) fournit rapidement une bonne visibilité à un contenu de qualité.**
- 4. Il est important de veiller à la bonne indexabilité de son site par les spiders (navigation, liens "spider friendly", fichier Sitemaps, etc.)**

Si vous suivez ces quelques conseils, vous ne devriez pas avoir de réels soucis avec les équipes de lutte contre le spam qui officient au sein des moteurs. Cependant, il se peut que, même sans le faire exprès, vous passiez, à un moment ou à un autre, de l'autre côté de la frontière. Voici à quelle sauce vous risquez alors d'être puni...

### ***Pénalité numéro 1 : Sandbox***

La "sandbox" ou "bac à sable" est une pénalité dont on a beaucoup parlé il y a quelques années de cela, notamment à partir de 2004. Il semblerait qu'elle frappe certains sites au moment où Google les découvre et estime que "quelque chose ne va pas" au niveau de l'analyse de ses liens entrants.

Par exemple, Google identifie, lors de sa première visite du site, que ce dernier a déjà obtenu de très nombreux liens entrants (*backlinks*). Ou de nombreux liens entrants depuis des sites distants, chaque fois avec le même "texte d'ancre" (notion de réputation). Etc. Google fait d'importants travaux sur la "courbe de vie d'un site" et, lorsqu'il découvre une nouvelle source d'information, il compare notamment sa structure et la situation constatée en termes de liens entrants par rapport à la moyenne de celle d'un site qui vient de sortir. Si les courbes et indices ne concordent pas, il peut y avoir manipulation.

Certains sites qui auraient trop "forcé la dose" dès leur lancement en termes de popularité et de réputation se seraient donc vues, par le passé, "mis en quarantaine" dans la Sandbox. Résultat : aucune possibilité de sortir en bonne position sur une quelconque requête pendant plusieurs semaines (*a priori* de 1 à 12 selon les chiffres le plus souvent constatés). Le site est bien là, il est bien "référéncé", mais jamais positionné. Puis, un jour, sa pénalité est terminée, il sort de la Sandbox, sa période de quarantaine est terminée, et il se classe tout de suite mieux... Il semblerait que la Sandbox touche, dans ce cas, toutes les pages d'un même site.

Il semblerait également que l'on entende beaucoup moins parler de ce phénomène depuis quelques temps. Google l'a-t-il abandonné ou remplacé par un autre système de pénalité ? Nul ne le sait... Mais il est sûr que ce phénomène a traumatisé plus d'un webmaster qui se voyaient parfois "sandboxé" tous les matins, créant une réelle psychose sur le Web...

**Pour en savoir plus sur la Sandbox...** Voici quelques articles qui devraient vous en dire plus, en français et en anglais sur le phénomène de "Sandbox" :

**Google et l'effet "sandbox"**

<http://www.webrankinfo.com/analyses/articles/sandbox.php>

**Analyse 2005 de la Sandbox**

<http://www.7-dragons.com/archives/sandbox.php>

**La Google SandBox**

<http://www.arkantos-consulting.com/articles-referencement/200605/google-sandbox.php>

**The new Google Patent and the new quality guidelines**

<http://www.seonewsblog.com/the-new-google-patent-and-the-new-quality-guidelines>

### **Pénalité numéro 2 : le déclassement**

Parfois, pour un site donné mais le plus souvent pour une requête donnée, un site web perd, du jour au lendemain, plusieurs (et parfois de très nombreuses) places dans les résultats du moteur. Ces pénalités sont connues sous le nom de "minus 30", "minus 60" ou "Position 6 penalty" pour l'une d'entre elles, apparue récemment (<http://blog.abondance.com/2008/01/position-6-penalty-mythe-ou-ralit.html>).

Ainsi, une page web sera "déclassée" pour une requête du type "*britney spears*" et disparaîtra en quelques heures des premières pages de résultats pour ces mots clés alors qu'elle reste toujours bien placée pour des requêtes utilisant la syntaxe "*allintext:birtney spears*" ou "*allintitle:britney spears*". Cela semble prouver qu'un site web, voire une page web, serait pénalisée par Google pour une requête bien particulière...

Selon la pénalité, la perte en termes de positions peut être minime (*Position 6 penalty* : le site passe de la 1ère à la 6ème place, ce qui le fait passer "en dessous de la ligne de flottaison", affectant quand même de façon forte le trafic généré) ou plus importante ("*minus 30*" : perte de 30 places ou plus...). Une pénalité "minus 950" a même été évoquée sur certains forums...

Là encore, un silence de cathédrale nous revient de la part de Google lorsque ces pénalités sont évoquées. Difficile donc de faire la part des choses entre mythe et réalité... Mais il semblerait bien que ces pénalité punissent une sur-optimisation des pages du site (*keyword stuffing*, texte et liens cachés, etc.) et ne touchent que certaines pages d'un site et pas les autres.

**Matt Cutts et les pénalités Google...** Matt Cutts est le "porte-parole référencement" chez Google et son blog personnel est très lu. Voici quelque sposts, sur ce blog, qui parlent des pénalités infligées par Google à certains sites. Ils sont hélas assez rares :

**Alerting site owners to problems** (20 septembre 2005)

<http://www.mattcutts.com/blog/webmaster-communication/>

**Confirming a penalty** (11 février 2006)

<http://www.mattcutts.com/blog/confirming-a-penalty/>

**Notifying webmasters of penalties** (26 avril 2006)

<http://www.mattcutts.com/blog/notifying-webmasters-of-penalties/>

### **Pénalité numéro 3 : la baisse de Pagerank**

Google a utilisé dernièrement, lors d'une campagne de communication contre la vente de liens (*paid linking*), une pénalité consistant à faire baisser, dans la barre d'outils qu'il propose (Googlebar), la valeur du Pagerank affiché, sans que cela affecte, *a priori*, le positionnement du site dans ses résultats sur les requêtes saisies par les internautes.

Cette pratique a fait couler beaucoup d'encre (virtuelle), aussi nous n'y reviendrons pas plus en détails, d'autant plus que les effets n'en sont pas réellement dévastateurs pour le trafic généré. Disons qu'il s'agit plus là d'un "avertissement" clairement destiné aux webmasters s'intéressant au référencement, avertissement qui permet de véhiculer une communication institutionnelle sur les blogs et les forums spécialisés...

### **Pénalité numéro 4 : la liste noire**

On rentre ici dans la pénalité la plus "dure" avec la "blacklist" ou "liste noire". Pour savoir si vous y avez plongé (malheur à vous !), le plus simple est d'utiliser une requête avec la syntaxe "site:www.votresite.com". Si Google ne renvoie aucun résultat alors qu'auparavant ce n'était pas le cas, il y a effectivement de fortes chances pour que votre site soit "blacklisté" et... que vous l'ayiez bien cherché...

### Que faire si vous êtes pénalisé ?

Nous avons déjà écrit un article sur la liste noire, dans la lettre R&R de novembre 2005 sous le titre "**Mon site est blacklisté, que faire ?**" (<http://abonnes.abondance.com/archives/2005-11/blacklist.pdf>). Nous ne reviendrons pas plus sur ce sujet, la plupart des informations disponibles dans cet article restant valables...

Deux points sont cependant importants dans le cadre de toutes ces pénalités infligées par les moteurs de recherche en général et Google en particulier :

- La plupart des pénalités sont infligées par des êtres humains. Si des "alertes" sont envoyées par des outils automatisés ou *via* le formulaire de *spam report*, les équipes de "search quality" de Google effectuent la plupart des pénalisations à la main, après vérification de la fraude éventuelle.

- Dans 9 cas sur 10, lorsqu'un site est déclassé, mis en sandbox ou blacklisté, son webmaster sait pourquoi sans que Google ait à le lui expliquer... Reste un cas sur 10 qui est parfois assez incompréhensible...

Sachez que, depuis que l'interface "Webmaster Tools" existe, Google améliore souvent celle-ci et cela a été le cas en juillet 2007 avec la mise en place d'un centre de correspondance privée avec les webmasters (<http://actu.abondance.com/2007-29/google-message-center.php>).

Ainsi, dans l'interface de ses outils pour webmasters (<http://www.google.com/webmasters/tools/>), le moteur de recherche propose une zone de contact avec ses équipes :

The screenshot shows the Google Webmaster Tools interface. At the top, it says "Google Webmaster Tools" and "andrieu@gmail.com | My Account | Help | Sign Out". Below this is a "Dashboard" section. On the left, there's a "Sites" table with columns for "Manage", "Sitemap", and "Verified?". The table lists three sites: www.abondance.com, www.boutique-abondance.com, and www.googlefight.com. On the right, there's a "Message Center" section with a "No messages" link and a "Tools" section with links for "Download data for all sites", "Report spam in our index", "Report paid links", and "Request reconsideration".

Manage	Sitemap	Verified?
<a href="#">www.abondance.com</a>	1	✓
<a href="#">www.boutique-abondance.com</a>	1	✓
<a href="#">www.googlefight.com</a>	Add	✓

This is a close-up of the "Message Center" and "Tools" sections. The "Message Center" section has a "No messages" link. The "Tools" section has four links: "Download data for all sites", "Report spam in our index", "Report paid links", and "Request reconsideration".

C'est par ce biais, ce "Message Center", que, si votre site est déclassé ou pénalisé, vous risquez de recevoir un message vous signalant ce fait. Le blog officiel de Google pour les webmaster Tools l'a clairement indiqué à l'époque (<http://googlewebmastercentral.blogspot.com/2007/07/message-center-let-us-communicate-with.html>).

Pour cela :

- Vous devez être inscrit sur les outils pour webmasters de Google (vérité de La Palice...).
- Et vous devez avoir lancé une procédure de vérification de vos sites (colonne "Verified?" du "Dashboard", voir illustration ci-dessus).

C'est également *via* cette interface que vous pourrez vous expliquer et demander le retrait de la pénalité après avoir corrigé votre site. La situation devrait alors s'améliorer rapidement. Tout du moins, nous l'espérons pour vous...

PS : nous n'avons pas parlé, dans cet article, de filtres utilisés par Google, notamment dans le domaine du "duplicate content", pour harmoniser le contenu de son index. Cela fera cependant l'objet d'un projet article. Mais il ne s'agit pas là vraiment de pénalités infligées suite à des tentatives de spam...

**Pour en savoir plus sur les pénalités de Google...** Voici quelques articles qui devraient vous en dire plus, en français et en anglais sur les différentes pénalités infligées par Google dans le cadre de sa "recherche qualité" :

**Les pénalités infligées par Google**

<http://www.annuaire-info.com/google-penalites.html>

**Sandbox et Blacklist de Google**

<http://www.go-referencement.org/google/sandbox-et-blacklist-de-google.html>

**Coffee Talk with Senior Google Engineer : Matt Cutts**

<http://www.seroundtable.com/archives/002809.html>

**Google's Most Common Penalty**

<http://www.threadwatch.org/node/11502>

**Google Ranking #6 Penalty / Filter**

<http://www.seobook.com/google-ranking-6-penalty-filter>

*Olivier Andrieu*

**Réagissez à cet article sur le blog des abonnés d'Abondance :**

<http://abonnes.abondance.com/blogpro/2008/02/les-pnalits-de-google-en-termes-de.html>