

## Google Panda et Penguin : quelques conseils pour éviter les bestioles monochromes de Google...

[Retour au sommaire de la lettre](#)

<b>Domaine :</b>	Recherche	<b>Référencement</b>
<b>Niveau :</b>	<b>Pour tous</b>	Avancé

*Depuis la mise en place du filtre Penguin sur l'algorithme des résultats de Google, certains référenceurs et webmasters ne comprennent pas les raisons qui poussent le géant de la recherche à déclasser ou non certains sites. Dans le même temps, le moteur de recherche continue les mises à jour de son précédent filtre Panda, tout en brevetant un algorithme pour faire varier la position d'une page en fonction des modifications et optimisations qui y sont apportées : le filtre "Ranking Documents".*

Google souhaite améliorer de manière drastique la qualité de ses résultats, et il s'arrange pour laisser les référenceurs dans le flou sur ce qui fonctionne réellement lorsqu'on optimise ou que l'on crée des liens et des contenus. Le but est simple : pousser les référenceurs à créer des contenus pertinents sans sur-optimiser ou tromper son moteur.

Cet article vise donc à faire un point rapide sur ces différents filtres et sur une méthodologie pour créer ou analyser un contenu et des liens naturels.

### ***Penguins, Pandas et autres filtres***

Avant d'entrer dans le détail de ce qu'est un site Internet "naturel", il nous semble important de faire le point sur les récents changements de l'algorithme de Google. En moins de deux ans, 2 mises à jours majeures ont été mises en place dans le monde :

- Google Panda
- Google Penguin

En réalité, chaque version de ces modifications de l'algorithme ont subi de nombreuses mises à jour, auxquelles se rajoutent d'autres modifications mineures :

- i. Panda : 23 février 2011
- ii. Panda 2.0 : 11 Avril 2011
- iii. Panda 2.1 : 9 Mai 2011
- iv. Panda 2.2 : 21 Juin 2011
- v. Panda 2.3 : 23 juin 2011
- vi. Panda 3.2 : 18 Janvier 2012
- vii. Panda 3.3 : 27 Février 2012
- viii. Panda 3.4 : 23 Mars 2012
- ix. Panda 3.5 : 19 Avril 2012
- x. Penguin : 24 Avril 2012
- xi. Panda 3.6 : 27 Avril 2012
- xii. Penguin 1.1 : 25 Mai 2012
- xiii. Panda 3.7 : 8 Juin 2012
- xiv. Panda 3.8 : 25 Juin 2012
- xv. Avertissement sur les liens : 19 Juillet 2012 (dans le Centre Webmaster, Google informe le webmaster que certains liens entrants sont considérés comme douteux).
- xvi. Panda 3.9 : 24 Juillet 2012
- xvii. Panda 3.9.1 : 20 Août 2012

Source : <http://www.seomoz.org/google-algorithm-change>

Gardez précieusement cette liste : cet historique est utile pour savoir quelle version de l'algorithme a touché votre site. Vous pouvez ainsi très bien avoir été impacté par Google Panda 3 jours seulement après la sortie officielle de Google Penguin. Le fait d'avoir été touché

par l'un ou par l'autre va en effet changer les actions que vous devez mener pour rectifier le tir.

Voici maintenant une brève explication de ces différents algorithmes :

## Panda

L'algorithme Panda est un filtre qui vise à déclasser les sites ayant un contenu pauvre ou de mauvaise qualité. Cela inclut les sites avec du contenu dupliqué, du contenu en faible quantité sur chaque page ou avec beaucoup de publicité ou de spam.

Dans ces cas de figure, voici les conseils et actions que vous devez mener de manière générale :

- Supprimer les duplications de contenus (qu'elles soient internes à votre site ou présentes sur des sites externes).
- Réduire le nombre de publicités.
- Étoffer, améliorer ou créer de nouveaux contenus.

Google Panda est un algorithme en constante évolution, mais qui n'est appliqué qu'à date fixe. En d'autres termes, il inflige une pénalité lors de sa mise à jour, par exemple lors du passage de la version 3.8 à la version 3.9. Un site Internet ne peut donc être impacté que lorsque la version de Google Panda évolue, d'où l'utilité du calendrier donné précédemment.

Sources :

- [http://en.wikipedia.org/wiki/Google\\_Panda](http://en.wikipedia.org/wiki/Google_Panda)
- <http://www.abondance.com/actualites/20120705-11642-infographie-google-panda.html>

## Penguin

Google Penguin est un algorithme qui se concentre sur la sur-optimisation des sites au niveau du référencement naturel : cela inclut deux grands éléments :

- Le spam de Linking, c'est-à-dire les liens qui pointent vers votre site.
- Le spam de contenu, c'est à dire la sur-optimisation de vos contenus.

Google Penguin est une mise à jour qui va dégrader le positionnement obtenu avec :

- une sur-optimisation des liens entrants, par exemple des liens avec toujours la même ancre ou sans attribut NoFollow (aors qu'il s'agit d'un lien acheté).
- une sur-optimisation des contenus, par exemple avec du keyword stuffing (le fait de placer un même mot clé partout dans le site) ou encore avec du contenu caché.

Source : <https://sites.google.com/site/webmasterhelpforum/fr/fr-posts/une-autre-mesure-pour-recompenser-les-sites-de-haute-qualite>

D'ailleurs, la prochaine mise à jour de l'algorithme Penguin devrait faire pas mal de bruit si l'on en croit les dires de Matt Cutts (15 Août 2012) :

*"Because of that, expect that the next few Penguin updates will take longer, incorporate additional signals, and as a result will have more noticeable impact".*

Source : <http://www.abondance.com/actualites/20120817-11768-google-penguin-la-prochaine-version-devrait-faire-mal.html>

## Ranking Documents

Autre algorithme qui risque de faire parler de lui : le brevet *Ranking Documents*. Il s'inscrit dans la lignée des deux précédents dans la lutte contre le contenu créé et optimisé uniquement pour le référencement.

Le concept est simple : quand Google détecte une modification substantielle d'une URL (que ce soit au niveau du contenu de celle-ci ou des liens entrants), le moteur de recherche fera le Yo-yo avec la position de ce contenu. Sur une courte période, la page va ainsi gagner ou perdre jour après jour plusieurs positions, et sa position définitive se stabilisera plus tard.

Le but est simple : compliquer la tâche des référenceurs afin que ceux-ci ne puissent plus savoir aussi facilement ce qui marche ou non lorsque l'on travaille le référencement d'un site Internet. Le brevet indique notamment que les modifications suivantes peuvent déclencher l'application du filtre :

- Keyword stuffing ;
- Texte invisible ;
- Redirections de contenus ;
- Liens sur-optimisés.

Selon certaines sources, cet algorithme serait déjà en place depuis plusieurs mois voir plusieurs années, mais aucune confirmation officielle n'a été donnée sur ce sujet. Si vous subissez une pénalité ou que votre site est instable dans les résultats de Google, il est donc possible que votre site Internet soit déjà dans ce cas de figure. Il faut donc éviter de remodifier vos contenus pendant cette période transitoire car cela peut empirer le phénomène.

Cela peut aussi vous nuire dans le sens où la première modification pouvait être destinée à avoir un meilleur positionnement. Et dernière raison d'attendre, si vous remodifiez un contenu qui fait le Yo-yo dans les pages de résultats, cela indique clairement à Google que vous agissez et modifiez vos contenus en fonction du classement de ceux-ci.

Sources :

- <http://www.webrankinfo.com/dossiers/r-et-d/brevet-ranking-documents>
- <http://www.seobythesea.com/2012/08/google-rank-modifying-spammers-patent/>

## ***Les bases d'un référencement naturel***

Nous venons de voir que Google cherche avec plus ou moins de succès à améliorer la pertinence de ses résultats en réduisant l'impact des optimisations et des sur-optimisations que l'on peut apporter à son site Internet.

Voici quelques conseils pour créer ou optimiser un contenu pour le référencement naturel, sans pour autant le sur-optimiser.

### **Les optimisations et les contenus du site**

La première étape consiste à créer des contenus "naturels". Google cherche à déclasser les contenus ayant un contenu pauvre. Comme expliqué précédemment, cela inclut :

- le contenu dupliqué ;
- le contenu "faible", c'est à dire avec peu de contenu réel ;
- le contenu avec une surabondance de publicités ;
- le spam.

Voici un exemple concret de contenu considéré (à raison) comme Spam :



En d'autres termes, ne cherchez pas à placer vos mots clés dans chaque phrase ou chaque titre de vos différentes publications. En règle générale, si la lecture d'une publication vous semble bizarre au niveau des répétitions et de la redondance des mots clés et expressions, c'est que vous avez sur-optimisé votre contenu.

Il y a une règle d'or : le contenu doit être fait pour le visiteur. Si c'est le cas, il sera en général également optimisé pour le moteur de recherche :

- Une URL compréhensible ;
- Un titre explicite ;
- Des titres H1 à H6 pour segmenter le contenu et faciliter la lecture ;
- etc.

## Les liens

Les liens entrants représentent l'un des meilleurs leviers pour améliorer le positionnement d'un site. En fonction de la source, de la popularité et de l'ancre choisie, certains liens peuvent avoir un impact réel et mesurable.

A travers ses différentes mises à jour, Google cherche à donner moins de poids aux différents liens sur-optimisés. C'est souvent le cas quand un même site reçoit des liens toujours en *DoFollow* ou toujours avec la même ancre. Dans ce cas de figure, le site peut facilement tomber sous le coup de la pénalité de Google Penguin, surtout sur un site récent.

A l'heure actuelle, aucun référenceur ni site Internet n'a réussi à trouver la formule qui permet de savoir avec précision quand un site sera être impacté ou non par Google Penguin. La raison à cela est simple : d'autres facteurs entrent en jeu, comme nous l'avons dit précédemment (popularité, ancienneté, contenus, ...).

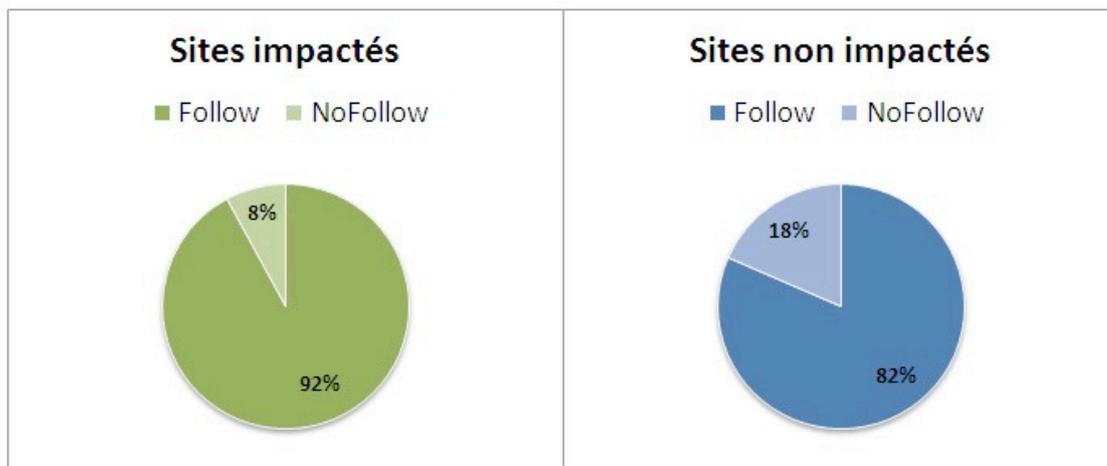
De manière générale, certains points communs sont cependant trouvés entre les sites impactés par une pénalité et ceux qui ne le sont pas.

## NoFollow

Parmi les référenceurs, tout le monde s'accorde à dire qu'un lien NoFollow ne sert à rien pour améliorer la popularité et le positionnement d'un site. Mais un site qui ne possède pas ou peu de liens avec l'attribut NoFollow est rare, pour la simple et bonne raison qu'un grand nombre de CMS ou de sites activent par défaut cet attribut sur les liens sortants.

Un site Internet doit donc avoir une partie de ses liens avec cet attribut afin de paraître plus naturel, même si cela ne vous fera pas gagner de positions.

Une étude de SeoMix sur ce sujet démontre que la part des liens NoFollow est beaucoup moins importante sur les sites impactés par Google Penguin que sur les autres (plus du double).



Source : <http://www.seomix.fr/google-penguin-liens/>

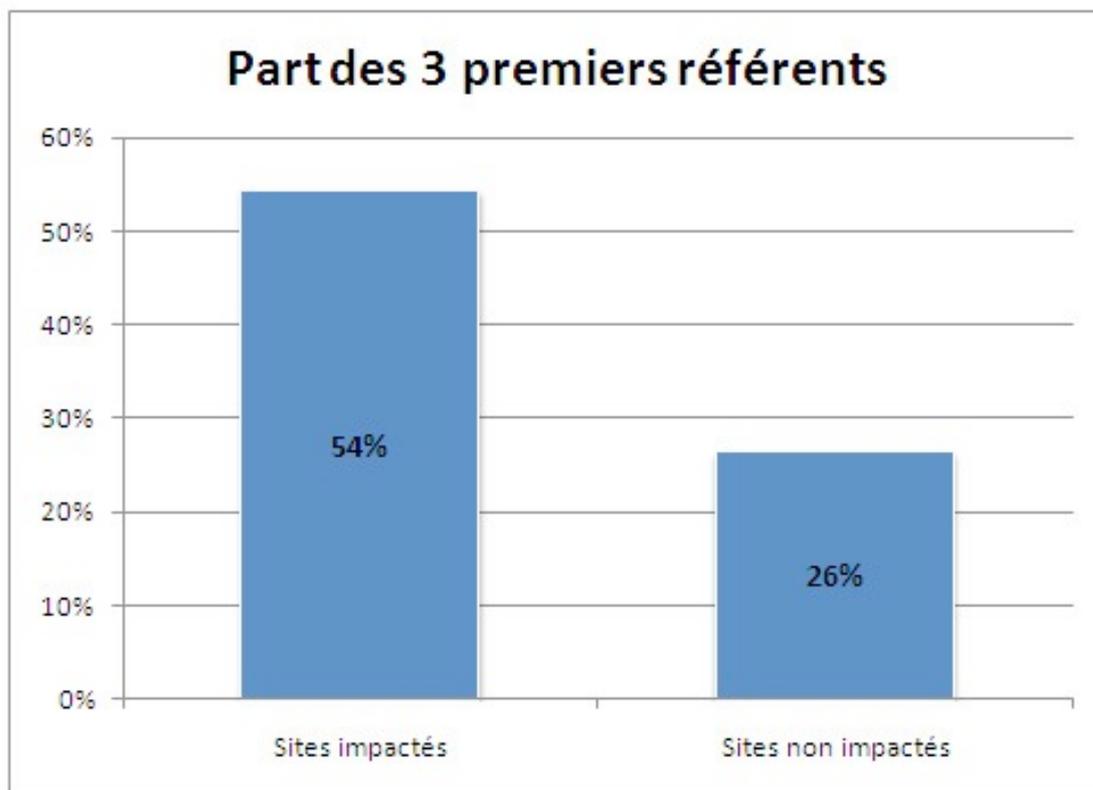
Vous devez donc inclure de temps à autre des liens en NoFollow : en d'autres termes, chaque lien ne doit pas forcément être optimisé.

### Les référents

Autre élément issu de la même étude : la répartition des liens sur l'ensemble des référents. Deux éléments ressortent ainsi :

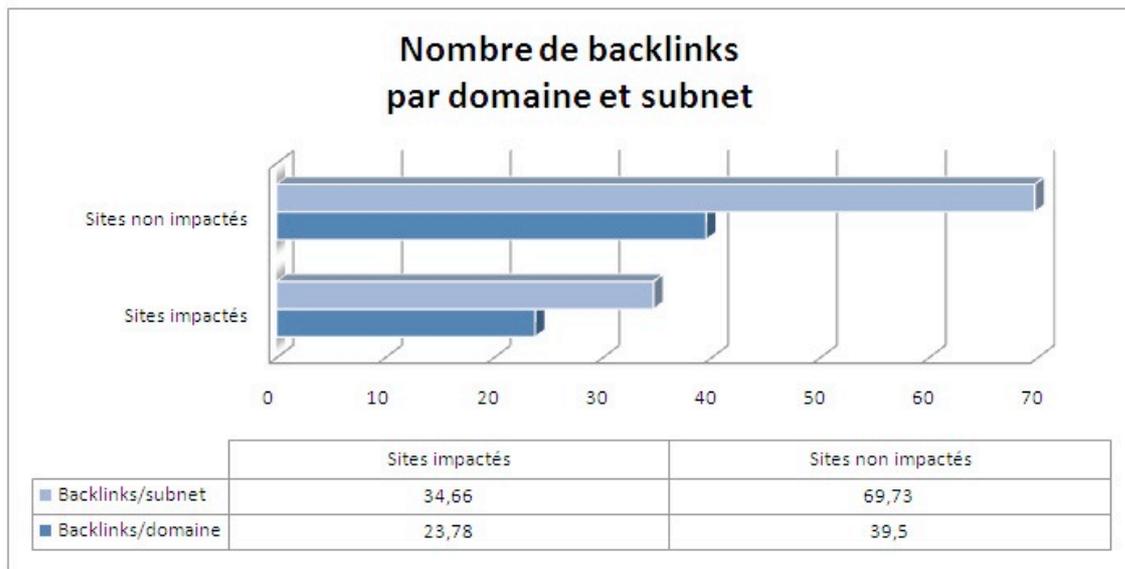
Un site impacté par Google Penguin possède de très nombreux liens sur un faible nombre de noms de domaine. En termes de nombres de liens, les 3 premiers sites représentent ainsi près de 50% du total des liens, contre 25% environ pour les sites non impactés.

Un analyse brève de ces chiffres peut prouver à Google que vous faites de l'échange de liens pour augmenter artificiellement votre positionnement (d'où un très grand nombre de liens sur un faible nombre de domaines). A l'inverse, les sites non impactés répartissent mieux la totalité des liens sur l'ensemble des domaines référents. Dans ce cas de figure, la part des 3 plus gros sites référents diminue de moitié.



Source : <http://www.seomix.fr/google-penguin-liens/>

Un autre aspect de l'analyse a montré que les sites qui n'ont pas été touchés par Penguin avaient généralement plus de liens sur chaque domaine référent. Contrairement au graphique précédent, cela tend à dire que chaque site fait généralement plus de liens vers votre site, sans pour autant que cela soit disproportionné par rapport aux 3 premiers sites référents.



Source : <http://www.seomix.fr/google-penguin-liens/>

*Définition pour l'image ci-dessus : un subnet est un début d'adresse IP identique. Deux noms de domaines ayant le même subnet sont situés sur le même serveur, et il est donc probable qu'ils appartiennent au même propriétaire. Par exemple, ces deux adresses IP de serveur ont le même subnet : 108.222.54.123 et 108.222.54.180*

Faisons une comparaison en mixant les données des deux derniers graphiques afin de mieux comprendre comment la répartition des liens par nom de domaine peut avoir un impact.

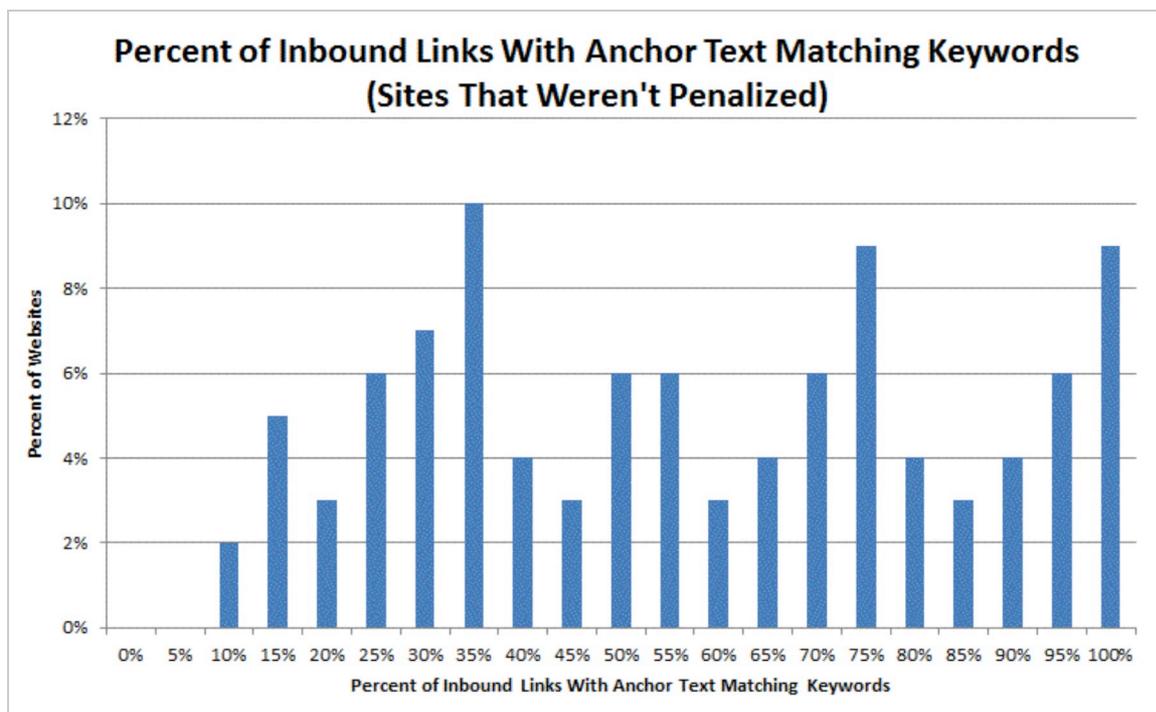
- Le profil d'un site non impacté
  - o Les 3 premiers sites représentent 25% du total des liens
  - o Chaque site fait en moyenne 40 liens
- Le profil d'un site impacté
  - o Les 3 premiers sites représentent 50% du total des liens
  - o Chaque site fait en moyenne 25 liens

## Les ancres

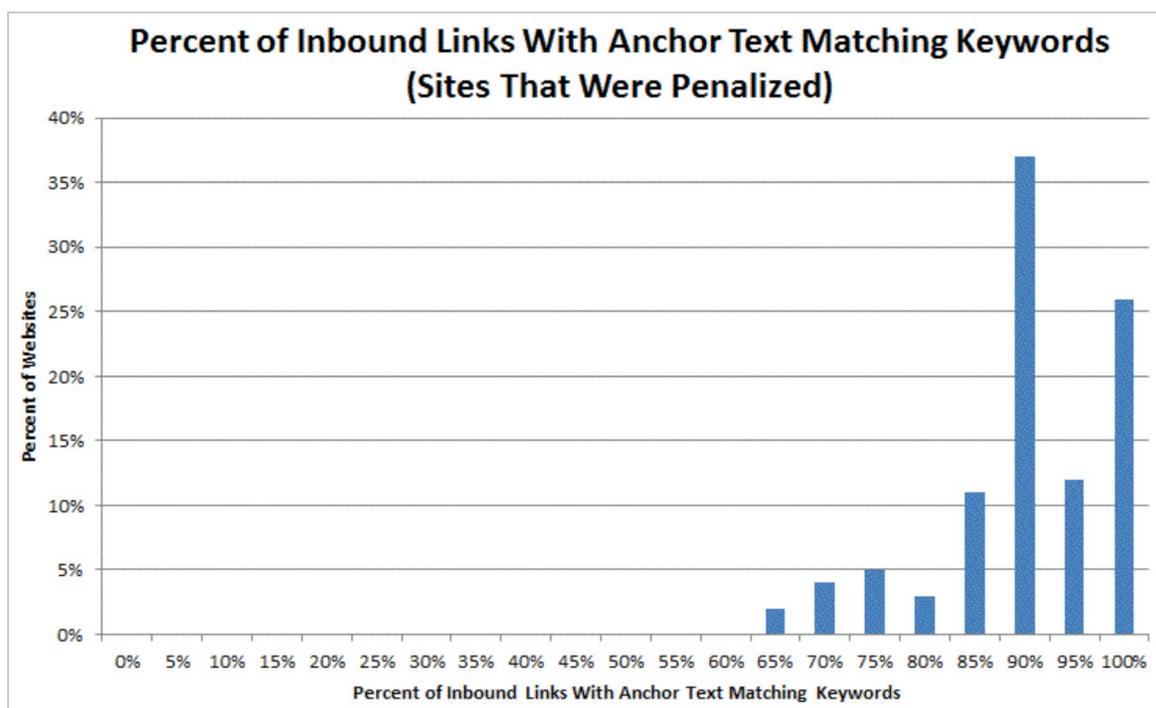
Un des points les plus flagrants concerne les ancres utilisées dans les liens. La plupart des sites Internet pénalisés par Penguin ont sur-optimisé l'ancre utilisée. On se retrouve ainsi parfois avec près de 80% ou 90% des liens qui possèdent une ancre exacte par rapport à l'expression que l'on veut cibler. En d'autres termes, le profil de liens qui en découle n'est pas naturel.

Source pour les deux images suivantes : <http://www.micrositemasters.com/blog/penguin-analysis-seo-isnt-dead-but-you-need-to-act-smarter-and-5-easy-ways-to-do-so/>

Ci-dessous, voici une étude sur les sites non pénalisés et la répartition des liens avec une ancre exacte :



Ci-dessous, le même graphique pour les sites touchés par Google Penguin :



Une ancre optimisée facilite le référencement d'un site Internet. Mais pour que cela n'ait pas d'impact négatif, il faut varier les ancres. De temps à autre, il faut donc créer des liens non-optimisés, par exemple avec :

- un lien avec uniquement l'URL en dur : <http://www.abondance.com>
- un lien avec une ancre "Prénom" : [Olivier](#)
- un lien avec une ancre sans signification : [Cliquez-ici](#)
- Etc.

Il faut donc éviter les textes où tous les liens sont entièrement optimisés comme dans cet exemple :

## Get Fit Using These Simple And Easy Methods.

Posted on [April 6, 2012](#) by [mary](#)

If you agree you are too active to get time and energy to exercise, you'll be amazed to find out available a fantastic exercise routine a lot sooner than you imagine. This post includes numerous ideas that could show you to improve your workout in a short amount of time, which enable it to [pay day loan](#) you stay healthy and keep the kitchen connoisseur.

To assist you to recover [loan](#) coming from a tricky exercise routine, try out offering the muscle groups exercise the next day. You want to do this softly, about 20 on the weight that one could elevate on one occasion. Try to do 25 repetitions in 2 packages. Choosing this, you'll have additional blood and nutrients sent to the muscle groups for quicker fix.

Climbing is a terrific way to stay fit while not having to expend every day [fast cash loans](#) a health club. Circumstances car park is a superb destination for a walk, since a lot of them have effectively groomed, predesignated hiking trails. You won't just obtain a cardiovascular exercise routine, there is however a high probability additionally, you will take in some stunning views.

Source : <https://sites.google.com/site/webmasterhelpforum/fr/fr-posts/une-autre-mesure-pour-recompenser-les-sites-de-haute-qualite>

## Ce qu'il ne faut pas faire

Voici maintenant quelques conseils si vous avez été touché par un changement d'algorithme.

### Quelle mise à jour ?

Tout d'abord, consultez le calendrier en début d'article pour déterminer quelle mise à jour a eu un impact sur votre site Internet.

En fonction de ce que vous trouverez, vous devriez pouvoir savoir si ce sont vos liens, vos contenus ou les deux qui posent problèmes.

### Ne supprimez pas vos liens

Deuxième conseil, ne cherchez pas à supprimer les liens déjà obtenus, même s'ils sont sur-optimisés. Si vous faites cela, vous allez perdre du temps, vous allez prouver à Google que vous connaissiez cette sur-optimisation et vous allez perdre le travail déjà accompli.

Le mieux est de :

- modifier les ancres de certains liens sur-optimisés (pas sur tous mais sur ceux que vous pouvez modifier facilement)..
- continuer la création de liens en alternant liens optimisés et liens non-optimisés (NoFollow, ancres, ...).
- continuer la création de nouveaux contenus.

### Ne modifiez pas en continu votre site

Autre conseil : ne modifiez pas en continu votre site, car vous ne saurez pas ce qui marche ou ne marche pas.

Attention, cela ne veut pas dire que vous ne devez pas continuer à créer ou modifier vos contenus. Cela veut dire que si vous faites une modification importante d'un contenu ou de l'ensemble de votre site, il est fortement conseillé d'attendre quelques jours voire une à deux semaines pour en connaître l'impact définitif.

## Quelques outils

Nous venons de voir quels types de contenus peuvent impacter négativement un site, ce qui peut être détecté avec une simple visite et lecture des contenus de votre site.

Nous venons également de voir quels types de liens peuvent potentiellement vous nuire. Mais dans ce cas de figure, il est plus difficile de détecter lesquels vont vous nuire. Pour vous aider à mieux analyser un profil de liens, certains outils peuvent réellement vous aider. en voici quelques un ainsi que les différentes façons de les utiliser :

## Google Webmaster Tools

Depuis cet été, le centre webmaster de Google peut vous aider à analyser vos liens. Il permet cette action à deux niveaux : la première est la mise en place d'alertes pour le webmaster. Google va en effet vous prévenir en cas de détection de liens "frauduleux". A vous de les supprimer (si vous n'avez pas le choix) ou de modifier les ancres de ceux-ci.

☆July 19, 2012

▲**Outils Google pour les webmasters : détection de liens factices sur le site**

---

Cher propriétaire ou webmaster du site [http://](#)

Nous avons détecté que certaines des pages de votre site utilisent des techniques contraires aux [Consignes Google aux webmasters](#).

Nous vous invitons plus particulièrement à rechercher des liens potentiellement artificiels ou factices qui redirigeraient l'internaute vers votre site dans le but de manipuler le classement PageRank. Ces liens factices peuvent être ajoutés, par exemple, dans le cadre de la [vente de liens pour améliorer le classement PageRank](#) ou de [processus d'échange de liens](#).

Nous vous encourageons à apporter des modifications à votre site afin de le mettre en conformité avec nos consignes relatives à la qualité. Une fois ces changements effectués, veuillez envoyer une [demande de réexamen](#) pour que votre site soit inclus dans les résultats de recherche de Google.

Si vous trouvez des liens factices qui redirigent vers votre site et que vous êtes incapable de contrôler ou de supprimer, veuillez nous fournir des informations détaillées dans votre demande de réexamen.

Pour toute question relative à la résolution de ce problème, veuillez consulter notre [Forum d'aide pour les webmasters](#).

Cordialement,  
L'équipe Google chargée de la qualité des recherches



Source : <http://www.abondance.com/actualites/20120723-11710-des-messages-dans-les-google-webmaster-tools-ement-la-panique.html>

La seconde est la possibilité de pouvoir télécharger la liste des liens dans un fichier csv. Vous pourrez ainsi l'importer dans Excel pour pouvoir mieux les analyser, notamment pour comprendre la répartition de vos liens sur différents noms de domaine ou encore pour étudier les ancres utilisées dans ceux-ci.

**Vue d'ensemble » Toutes les pages référencées par le biais de liens**  
Vos 1 000 pages principales référencées sur d'autres domaines par le biais de liens

Télécharger ce tableau    Télécharger d'autres exemples de liens    Télécharger les derniers liens

Afficher 25 lignes    1 à 25 sur 1 000    < >

Vos pages	Liens ▲	Domaines sources
<a href="http://www.abondance.com/">http://www.abondance.com/</a>	522 684	2 520

Source : <http://www.abondance.com/actualites/20120719-11700-les-google-webmaster-tools-proposent-les-nouveaux-liens-classes-par-date.html>

L'outil : <http://www.google.fr/webmasters/>

## Open Site Explorer

Open Site Explorer est un outil gratuit qui vous permet d'analyser les liens de votre site Internet. Comme tout outil gratuit, il dispose de sa propre base de données. Les chiffres d'Open Site Explorer sont donc différents du centre webmaster de Google ou des deux autres outils que nous présenterons ensuite. Cependant, cela vous permet rapidement d'analyser le profil de liens d'un site.

La partie gratuite vous donne tout d'abord accès à quelques informations générales sur les liens :



Chaque onglet vous donne ensuite accès à des statistiques plus précises sur :

- Les liens entrants ;
- Les pages les mieux référencées ;
- Les domaines qui font des liens vers vous ;
- Les ancres utilisées.

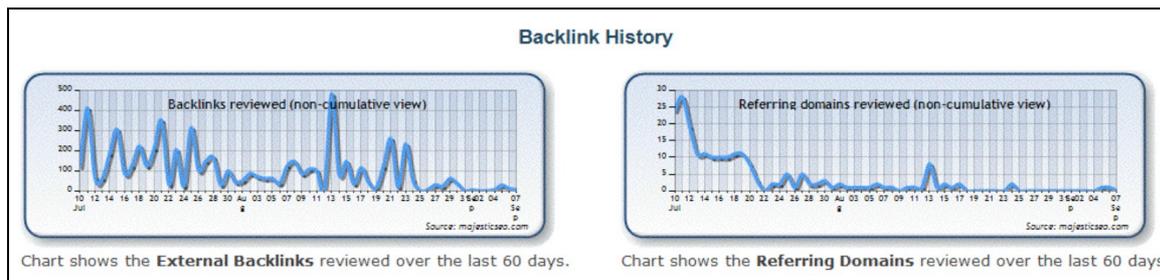
Il vous faudra cependant payer pour accéder à l'intégralité des données, mais la partie gratuite peut déjà vous aider dans votre travail.

L'outil : <http://www.opensiteexplorer.org/>

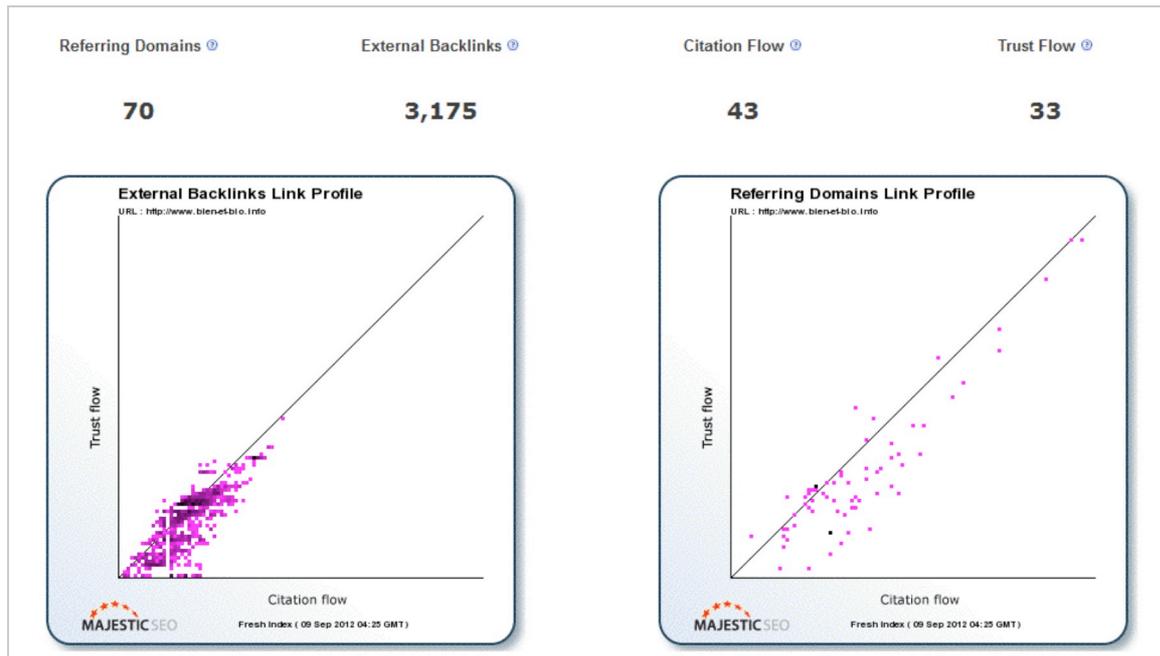
## Majestic SEO

Majestic est un autre outil intéressant pour analyser un profil de liens. Comme Open Site Explorer, vous aurez accès à quelques données en version gratuite :

- Une courbe d'évolution des domaines référents et des liens.
- Les meilleurs backlinks par ordre de "Citation Flow" (un critère de classement propre à Majestic SEO).
- Les domaines référents.
- Les meilleures pages.



Le dernier onglet permet d'avoir une répartition visuelle de vos liens selon les deux critères propres à l'outil. Plus un lien est haut, plus il est jugé crédible et populaire. Plus il est à droite, plus un même site va faire de liens en même temps vers le vôtre.



L'outil : <https://www.majesticseo.com/>

## Ahrefs

Ahrefs est sans doute l'un des plus puissants outils gratuits d'analyse de backlinks, bien que plus récent que les deux présentés d'avant.

Comme pour Majestic SEO ou Open Site Explorer, il permet d'analyser les backlinks par date, par page, par ancre ou encore par nom de domaine. L'avantage majeur est qu'il est bien plus visuel et donc pratique à utiliser. Il permet de savoir rapidement ce qui ne va pas dans le profil de lien du site analysé.

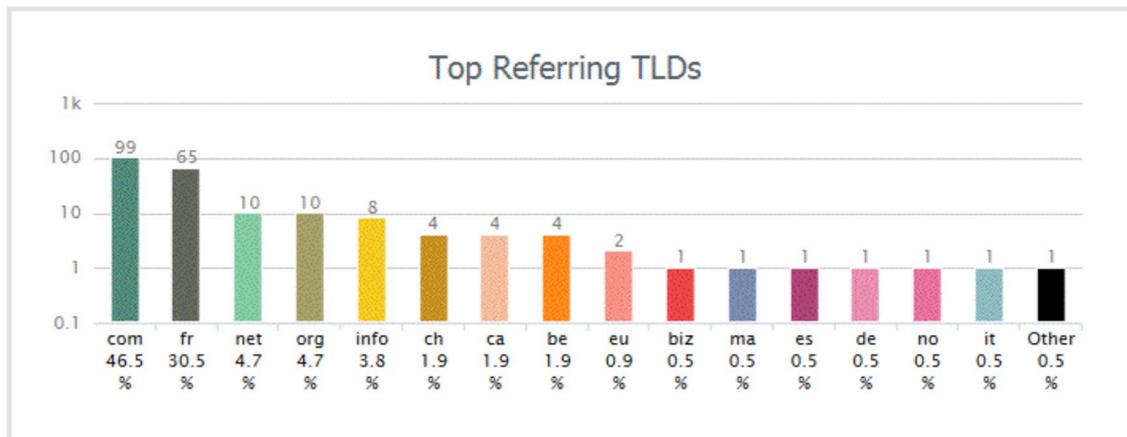
Dès le départ, il nous donne la répartition des liens par :

- type de liens ;
- NoFollow / DoFollow ;
- présence du lien sur tout le site ou sur une partie.

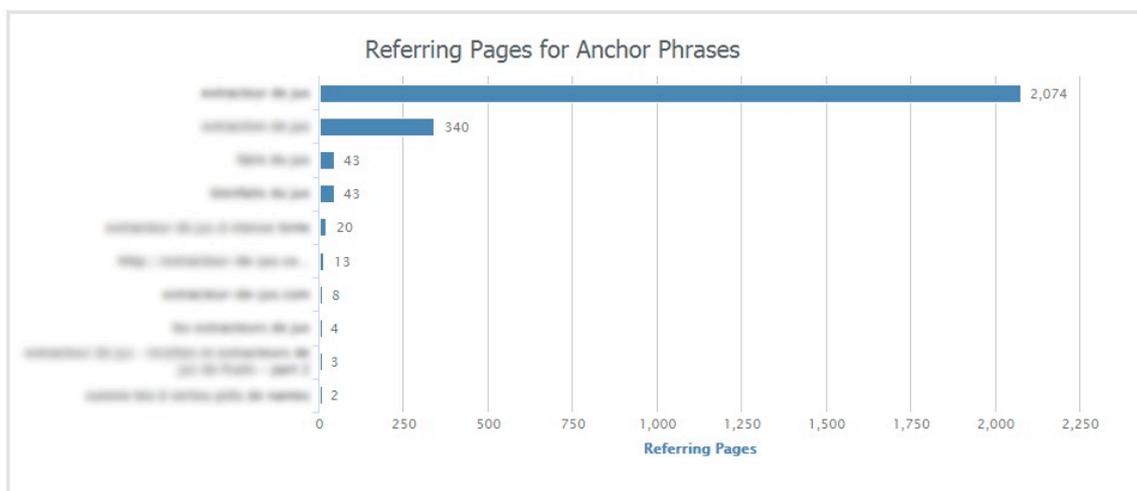


On peut également analyser la répartition des liens par nom de domaine (TLD). Si la dernière colonne *Other* est aussi faible que dans l'exemple ci-dessous, cela signifie que peu de domaines différents font des liens vers votre site. Nous pouvons en tirer deux conclusions différentes:

- Le profil des liens est récent ;
- Le profil des liens n'est pas naturel.



Et voici, pour terminer, un exemple de mauvaise répartition des ancres utilisées dans les liens pointant vers un site, et qui ont provoqué une pénalité suite à la mise en place de la mise à jour de Google Penguin :



L'outil : <http://ahrefs.com/>

## Conclusion

Internet est exactement comme notre planète : il est varié, multiple et hétérogène. Vos liens et vos contenus doivent être optimisés, mais pas de façon systématique. Pensez toujours à vos visiteurs en premier et vous devriez éviter toute sur-optimisation inutile.

Et si vous avez déjà subi une pénalité, il est tout à fait possible d'en sortir en variant vos liens et en réduisant les optimisations de votre contenu. Dans ce cas-là, utilisez les outils précédents pour savoir quels sont les points les plus négatifs (NoFollow, variation des ancres, diversité des noms de domaine, ...).

Bref, soyez naturel. Et tout devrait bien se passer. C'est du moins tout le mal que nous souhaitons à votre site web...

**Daniel Roch**, *Consultant WordPress, Référencement et Webmarketing chez SeoMix*  
(<http://www.seomix.fr/>)