

Le placement dans les moteurs de recherche

Stephanie Booth, <http://climbtothestars.org>, juillet 2007

License Creative Commons by-nc-sa: <<http://creativecommons.org/licenses/by-nc-sa/2.5/ch/deed.fr>>

Introduction

On est souvent tenté de mesurer le succès d'un site internet par son placement dans les moteurs de recherche [1]. En effet, un bon placement dans les moteurs de recherche peut être un indicateur parmi d'autres du succès d'un site. Malheureusement, le fonctionnement des moteurs de recherche en général et du positionnement en particulier est difficile à comprendre et amène souvent à avoir des attentes peu réalistes en la matière, en particulier l'exigence des résultats précis pour certains mots-clés donnés.

Le positionnement dans un moteur de recherche ne peut pas être contrôlé [2]. On peut l'influencer positivement en soignant certains aspects techniques ou rédactionnels du site, mais le facteur principal dans un bon placement demeurera inévitablement la qualité du contenu.

En fait, le placement dans les moteurs de recherche est une conséquence de la popularité d'un site, et non sa cause. (Même si, bien entendu, un bon placement peut amener des visiteurs au site -- mais un bon placement seul ne suffit pas.)

1. Comment fonctionne un moteur de recherche ?

1.1. Fonctionnement général

La fonction principale d'un moteur de recherche est de fournir à son utilisateur la liste des pages Internet correspondant aux mots-clés que celui-ci aura spécifiés. Généralement [3], cela revient à fournir une liste de pages contenant ces mots-clés. Cela est en soi une opération relativement simple. Ce qui est très complexe, par contre, c'est le classement de tous ces résultats afin de mettre en tête de liste les plus pertinents. L'algorithme utilisé par Google à cet effet est secret, et comprend une bonne centaine de variables. On saisit donc aisément pourquoi il n'est pas possible de contrôler sa position dans un moteur de recherche. D'ailleurs, si ça l'était, ce positionnement perdrait tout son sens...

Le premier facteur qui va influencer le placement est bien entendu le contenu de la page : les mots-clés y apparaissent-ils ? Sont-ils proches les uns des autres, ou bien éloignés ? Apparaissent-ils une seule fois, ou bien plusieurs fois ? Apparaissent-ils dans le titre de la page, dans les divers titres qui organisent le contenu de celle-ci, dans le corps du texte, en gras, dans un lien, dans des citations, en haut de la page, en bas de la page... ?

Un autre facteur qui va avoir beaucoup d'influence sur "l'autorité" que Google accordera à un site Web est le nombre de liens entrants vers ce site. De plus, la provenance de ces liens a une importance : le "pagerank" (autorité) se calcule de façon circulaire. Logiquement, les liens provenant de sites dont l'autorité est déjà reconnue auront plus d'impact que des

liens provenant de sites moins cotés (mais un grand nombre de liens provenant de sites peu cotés comptera également...).

Le fait de définir des "mots-clés" [4] dans les différentes pages du site était très important pour le placement dans les moteurs de recherche à une certaine époque (pré-Google). Maintenant, l'importance de ces mots-clés est très limitée, voire même nulle selon certains.

1.2. Le choix des mots clés

Lorsque l'on cherche à mesurer la performance d'un site dans Google, par exemple, on introduit dans le moteur de recherche des mots clés qui semblent pertinents par rapport au contenu du site en question. Mais ces mots-clés sont-ils ceux qu'introduiraient les personnes faisant partie du public-cible du site?

Il est en fait très difficile de deviner quels mots-clés un internaute recherchant une information donnée va introduire dans un moteur de recherche. Parfois, les gens y tapent des questions entières. Parfois juste un mot. Parfois, les mots pour lesquels ils imaginent que le site web dont ils auraient besoin aura été optimisé. Et parfois encore, l'information qu'ils tirent du site et qui a de la valeur pour eux n'est pas celle qu'imaginait l'auteur du site.

Se focaliser sur certains mots-clés n'est donc pas une stratégie très porteuse. Il vaut mieux se concentrer sur le référencement général du site et la qualité de son contenu, et laisser au moteur de recherche le soin de déterminer quels sont les mots-clés les plus pertinents pour lui.

1.3. Langue et géographie

Un moteur de recherche comme Google donne la possibilité de limiter les langues dans lesquelles la recherche est effectuée, ou de circonscrire la recherche à un pays donné. Il faut cependant prendre garde au fait que la détection de la langue ou du pays de provenance d'une page peut facilement être erronée. La langue est détectée automatiquement en analysant le contenu de la page (ce qui crée vite des problèmes lorsqu'une page contient plus d'une langue), et pour ce qui est du pays, il semblerait que la localisation physique du serveur hébergeant le site web soit déterminante.

Notons également qu'une recherche "en français" donne des résultats légèrement différents selon qu'on utilise google.fr ou google.ch.

Il ne faut donc pas trop compter sur ces distinctions qui fluctuent facilement -- mais elles ont le mérite de rappeler qu'un site visant un public local est "en compétition", pour ce qui est de la pertinence des mots-clés, avec tous les autres sites publiés dans la même langue. Les seules véritables frontières sur internet sont linguistiques.

2. Que conclure d'un placement insatisfaisant?

C'est souvent le constat d'un placement peu satisfaisant dans les moteurs de recherche qui amène à se demander ce que l'on peut faire pour l'améliorer. Il est important, tout d'abord, de ne pas perdre de vue l'élément de subjectivité que comporte un tel constat. Le placement n'est pas quelque chose d'absolu: une page va être classée en fonction de sa pertinence par rapport aux mots-clés qui ont été donnés comme critères de recherche. La sélection des mots-clés doit donc se faire très soigneusement s'ils doivent servir à évaluer d'une façon ou d'une autre si le placement est "bon".

Les causes d'un placement insatisfaisant peuvent être multiples. Pour y remédier, il est important des les identifier correctement. Tout d'abord, il faut distinguer parmi les facteurs agissant sur le placement ceux qui sont "internes", propres au site placé, et donc sur lesquels on peut agir, de ceux qui sont "externes", sur lesquels on n'a pas prise.

2.1. Facteurs externes

Ils sont au nombre de deux:

- les mots-clés choisis
- les sites "concurrents"

Chaque site est placé dans la liste des résultats pour les mots-clés choisis par rapport aux autres sites ayant un contenu similaire (potentiellement pertinent pour les mots-clés recherchés). Obtenir un "bon placement" dans un tel cas de figure sera donc une toute autre histoire s'il s'agit d'un domaine dans lequel il existe quantité de sites ayant déjà une présence forte, ou si le domaine en question est relativement peu documenté sur internet.

2.2. Facteurs internes

Les facteurs internes influençant le positionnement d'un site sont les suivants:

- qualité du contenu
- mise en valeur rédactionnelle du contenu
- mise en valeur technique du contenu

La *qualité du contenu* influence de deux manières le placement. Premièrement, d'une façon tout à fait mécanique, comme expliqué plus haut: la fréquence et la place des mots-clés dans le texte de la page est analysé et interprété par le moteur de recherche. Celui-ci, par contre, n'a aucun moyen d'évaluer la "qualité" d'un texte au sens où un être humain pourrait l'entendre. Cette "qualité" est en fait mesurée au nombre de liens entrants établis par d'autres créateurs de site. Ce n'est donc pas directement le moteur de recherche qui est influencé ici, mais les êtres humains qui jugeront le contenu digne ou non d'une mention sur leur site.

La "*mise en valeur rédactionnelle*" consiste à organiser le texte d'une façon qui tienne compte de la manière dont le moteur de recherche indexera le contenu: que va-t-on

regrouper sur une même page? Quelles expressions choisir pour les titres? comment utiliser gras, liens, sous-titres de façon adéquate?

La "*mise en valeur technique*", quant à elle, consiste à s'assurer qu'il n'y a pas d'obstacles techniques à la bonne indexation du site et à la création de liens vers les différentes pages qu'il comporte. Par exemple, on évitera les cadres ("frames"), les liens en javascript, le balisage non sémantique, etc.

Lorsque le référencement n'est pas satisfaisant, il convient donc de repérer dans quel domaine les améliorations doivent être faites. Pour ce faire, on regardera dans l'ordre:

- les aspects techniques
- les aspects rédactionnels
- le contenu proprement dit.

En effet, s'il y a des obstacles techniques majeurs à une bonne indexation, il ne sert à rien de s'acharner à améliorer la qualité ou l'intérêt du contenu. Celui-ci est peut-être parfaitement adéquat, mais il n'est simplement pas mis correctement en valeur. Une fois que la qualité technique du site est assurée, on peut s'intéresser à la technique de rédaction, par exemple en améliorant la façon dont les titres sont choisis, ou en optimisant la façon dont le contenu est organisé sur les diverses pages du site.

C'est seulement une fois ces questions techniques réglées (du moins pour ce qui est des grandes lignes) que l'on peut réellement évaluer si la qualité proprement dite du contenu est en cause. Et à ce moment-là, il est important d'inclure dans l'évaluation les facteurs extérieurs comme la place occupée par les autres sites consacrés à des sujets similaires.

3. Que faire pour améliorer le placement général ?

Encore une fois, la chose principale à faire pour avoir un bon placement est d'avoir un contenu de qualité. En effet, les algorithmes des moteurs de recherche visent à mettre en avant les pages qui sont utiles aux gens faisant les recherches -- le contenu est donc primordial.

Les créateurs de sites peuvent ensuite aider les moteurs de recherche à indexer correctement leur contenu en prenant garde à certains points:

Aspects techniques :

- utiliser un balisage HTML propre et sémantique
- prévoir une architecture de site qui encourage les autres propriétaires de sites à faire des liens vers différentes pages de celui-ci
- créer suffisamment de liens entre les différentes parties du site pour permettre une navigation facile (particulièrement dans le corps du texte lorsque cela est pertinent)
- utiliser un schéma d'URL "sympathique" pour les moteurs de recherche
- éviter les frames (cadres), les images représentant du texte et les liens en javascript.

Aspects rédactionnels :

- structurer les pages avec titres et sous-titres
- choisir des titres qui contiennent des mots-clés en adéquation avec le contenu de la page.

On notera qu'il y a deux sortes de "bon placement": l'autorité que Google confère à une page de façon générale, ou "PageRank" [5], et la position dans laquelle cette page se trouvera pour la recherche de mots-clés précis.

Une page peut avoir un très bon PageRank mais être mal placée pour une recherche selon certains mots-clés. Cela peut être le cas parce que ces mots-clés n'occupent pas une place assez prédominante (au sens de Google) dans la page, voire n'y figurent pas (!). Mais aussi, d'autres sites existent peut-être dont le contenu relatif à ces mots-clés est plus pertinent. Cette pertinence apparaîtra (aux yeux du moteur de recherche) à travers la place qu'occupent ces mots-clés dans le contenu de la page en question et des pages proches, et surtout à travers les liens entrants vers cette page: le texte de ces liens utilise-t-il régulièrement ces mots-clés, ou non?

Par exemple, on trouvera normal qu'en tapant les mots-clés "prévention+sida" le premier résultat apparaissant soit <<http://www.preventionsida.org/>>, un site consacré uniquement à la prévention Sida, plutôt que <<http://www.ciao.ch/f/sexualite/infos/10.1sida-virus-et-maladie>>, la page à ce sujet sur ciao.ch, un site qui comporte des informations sur des sujets très variés.

4. Qu'apporte un bon placement ?

Il peut être utile de regarder d'un peu plus près les raisons qui poussent à vouloir un bon placement dans un moteur de recherche. Effectivement, un site bien placé, qui se situe dans les premiers résultats pour les mots-clés que recherchent les gens qui seraient intéressés par le contenu du site, va attirer plus de visiteurs par ce biais.

Gardons cependant à l'esprit que le placement est un moyen d'attirer des visiteurs, et non une fin en soi. Ce n'est d'ailleurs pas le seul moyen d'attirer des visiteurs sur son site: les gens peuvent aussi y arriver parce qu'ils ont cliqué sur un lien se trouvant sur un autre site, parce qu'on leur en a parlé, ou qu'on leur a envoyé l'adresse par e-mail, chat, ou messagerie instantanée.

Ces derniers moyens font intervenir quelque chose qui est de l'ordre du bouche à oreille: une recommandation "de personne à personne", qui est généralement beaucoup plus efficace qu'une entrée dans un répertoire automatisé (le moteur de recherche), et qui prend de plus en plus d'importance dans l'internet d'aujourd'hui. Un lien sur un site web est une recommandation, et cela d'autant plus si ce site est un site personnel, comme un blog, par exemple.

Bien entendu, les résultats des moteurs de recherche vont refléter la part "internet" de ces recommandations en bouche à oreille. Une adresse web transmise dans une liste de

discussion par e-mail va se retrouver indexée via les archives en ligne de celle-ci. Les blogs sont des sites internet, et "comptent" évidemment pour le moteur de recherche -- d'autant plus que bon nombre d'entre eux sont en fait très bien "cotés" (PageRank).

Si beaucoup de propriétaires de sites jugent que le contenu d'un site vaut la peine qu'on en parle, le placement de ce site dans les moteurs de recherche va s'en trouver amélioré, mais ce placement est la conséquence de cet intérêt et non sa cause. Tenter d'agir directement sur le placement est donc insuffisant: c'est le contenu du site qui va jouer un rôle prépondérant.

Conclusion

Le placement d'un site dans un moteur de recherche est un indicateur de qualité parmi d'autres. Ce placement ne peut pas être contrôlé, et ne dépend pas uniquement du site concerné -- il est toujours relatif au reste du matériel existant sur internet.

Il est bien plus efficace d'agir sur le contenu et la qualité formelle d'un site afin de l'améliorer plutôt que de chercher à obtenir directement un résultat en termes de placement par rapport à certains mots-clés, objectif qui est souvent peu réaliste.

Notes

1. La plupart des moteurs de recherche fonctionnent sur le modèle de Google. Si ce document parle des moteurs de recherche en général, c'est plus particulièrement de Google et des moteurs de recherche fonctionnant sur des principes similaires dont il s'agit ici.

2. Les organisations qui garantissent un placement dans les x premiers résultats pour tel ou tel mot-clé utilisent des techniques peu éthiques qui reviennent en somme à tricher avec les algorithmes de classement des moteurs de recherche

[<<http://www.google.com/support/webmasters/bin/answer.py?answer=35291>>]. D'un moyen pour mesurer quelque chose (la qualité), le placement devient un but en soi.

3. Google tient en effet également compte des mots utilisés pour créer des liens vers une page donnée. Le "googlebombling" abuse de ce fait pour imposer une page spécifique en première position lorsqu'on l'on recherche certains mots-clés, en obtenant qu'un grand nombre de sites Web utilisent ceux-ci dans le texte de liens pointant vers cette page. Un exemple célèbre plaçait la biographie officielle de George Bush en première position des résultats lorsque l'on cherchait l'expression "miserable failure", même si la page ne contenait bien sûr pas cette expression. Voir aussi

<http://fr.wikipedia.org/wiki/Bombardement_Google>.

4. Ces fameux mots-clés, importants pour les moteurs de recherche de la génération précédant Google, sont ceux que l'on place dans la balise HTML <meta name="keywords" content="..." />.

5. <<http://en.wikipedia.org/wiki/PageRank>> Le PageRank est une valeur allant de 0 à 10 et qui est basée sur le nombre de liens entrants vers une page, pondérée par la valeur de PageRank de la page sur laquelle se trouve ce lien.