

Comment optimiser une page web pour les moteurs de recherche ? (2ème partie)

Deuxième article d'une série qui, nous l'espérons, vous permettra de mieux optimiser vos pages web afin de les rendre réactives aux critères de pertinence des moteurs de recherche. Après le titre, le mois dernier, nous étudions ce mois-ci les différentes façons de mieux optimiser les textes de vos pages web.

Nous vous proposons, dans une série d'articles qui seront consacrés à ce thème au fil des mois, de vous indiquer les meilleures pratiques pour optimiser une page web pour les moteurs de recherche et donc pour fournir à votre contenu (de qualité, s'entend) une meilleure visibilité au travers des outils de recherche. Au sein de ces articles successifs, plusieurs "zones" des pages web seront prises en compte chaque mois : le titre, le texte visible, les balises Meta, les noms de domaine, les urls, etc.

Ce mois-ci, nous allons tenter de faire le tour de la question au sujet du texte "visible" des pages web : comment l'optimiser au mieux pour que ces textes soient les plus réactifs possibles en regard des critères de pertinence des moteurs de recherche ? Suivez le guide...

La notion de "texte visible"

Tout d'abord, définissons clairement ce que nous entendons par "texte visible" :

- "**Texte**" car ce dont nous allons parler dans cet article est le contenu **textuel** de vos pages. En d'autres termes, tout le contenu que vous pouvez sélectionner avec votre souris, copier puis coller dans un traitement de texte comme Word. Dans ce sens, un texte inclu dans une image n'est pas considéré comme étant au format textuel. Idem pour un texte inséré dans une animation Flash, etc.

- "**Visible**" car nous ne parlerons ici que du texte proposé "loyalement" sur les pages web, sans traiter des cas de spam datant du paléolithique inférieur du Web consistant à insérer, par exemple, du texte en blanc sur fond blanc, invisible pour l'internaute mais soit-disant pris en compte par le moteur. S'il est vrai que ce type de spam marche parfois sur certains moteurs, et pas des moindres (et notamment ceux qui disent combattre ce type de pratique :-)), il s'agit clairement de spam et nous n'en parlerons donc pas. De plus, lorsqu'un internaute découvre le pot-aux-roses (et ce n'est pas vraiment très compliqué), la perte de crédibilité envers le site coupable est telle que cela devrait décourager tout webmaster de commettre ce type de "méfait"...

La notion de "texte visible" signifie également que nous ne traiterons pas ici, par exemple, de l'option "ALT" des balises images, qui sera étudiée dans un prochain article. Même si ce texte est "visible" lorsqu'on passe la souris sur l'image (sur certains navigateurs), il ne l'est pas sans action précise sur la page. Il n'en sera donc pas fait mention dans cet article. Idem pour les commentaires et tout contenu du code HTML non visualisé à l'écran du navigateur.

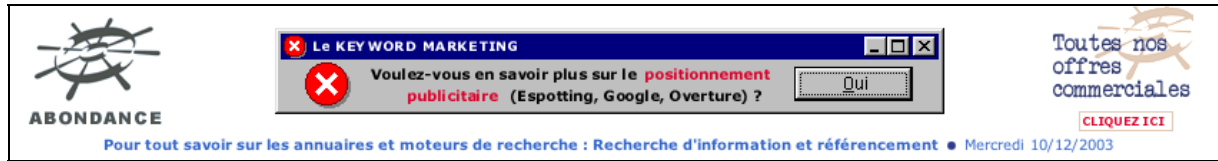
Ceci étant dit, nous allons donc commencer notre "revue d'effectif" des différentes possibilités d'optimisation du "texte visible" des pages...

Localisation du texte

L'un des critères importants pris en compte par les moteurs lors du calcul de pertinence d'une page par rapport à un mot donné est la présence de ce terme au début du document plutôt qu'à la fin. Plus le mot en question sera placé haut dans la page (le plus proche possible de la balise <BODY> pour ceux qui connaissent le langage HTML), plus sa présence sera jugée pertinente.

Globalement, on a tendance à dire, pour résumer, que le **premier paragraphe** des pages est le plus important. Une page optimisée doit contenir **au minimum 100 mots**, si possible plus, mais les 30 à 50 premiers termes, les deux à trois premières phrases, seront primordiaux. Vous devez donc insérer, dans ce premier paragraphe, le nom de votre entreprise/association/entité et les mots clés importants pour votre activité. Bref, reprendre ou développer le titre de votre page, point étudié le mois dernier.

Exemple sur le site Abondance :



Le logo, le bandeau publicitaire et l'accroche commerciale (à droite) sont des images. Elles ne sont donc pas prises en compte au format texte.

Juste en dessous, la phrase suivante est indiquée : "Pour tout savoir sur les annuaires et moteurs de recherche : Recherche d'information et référencement". Cette phrase n'est pas là par hasard. Elle contient de nombreux termes importants pour décrire le site : "annuaires", "moteurs de recherche", "référencement", "recherche d'information".

Puis, si vous regardez à l'intérieur du code HTML de la page, cette phrase est suivie par les termes situés à la gauche de l'écran et qui décrivent les grandes rubriques du site : "Actualité", "Dossiers / Articles", "Outils de recherche", "Outils / Audit de site", "Forums / Chat", etc.

Bref, la page a été conçue pour que tout le texte situé "en haut du code" contienne les mots clés importants pour décrire son contenu.

Donc, en résumé, n'oubliez pas de privilégier le début du texte de vos pages pour y insérer vos mots clés importants !

Une conséquence directe et importante de cela est que les sites en Flash ou comportant beaucoup d'images seront difficilement référencables et connaîtront des problèmes à être bien positionnés. Si le moteur n'y trouve pas, au format texte, vos mots clés favoris, il sera difficile pour vous d'être bien positionné dessus ! Un bon titre seul ne suffit pas toujours pour être bien positionné. La présence d'un contenu textuel qui allie la qualité à la quantité est important.

La mise en exergue du texte

Les moteurs de recherche privilégient également les mots qui sont mis en exergue dans les pages web.

Premier exemple : la **mise en gras**. N'hésitez pas à placer vos mots importants en gras, comme ici :

Nous vendons des **amortisseurs** et toutes les **pièces détachées** pour votre voiture.

Ce qui donnera l'affichage suivant :

Nous vendons des **amortisseurs** et toutes les **pièces détachées pour votre voiture**.

Une page qui contient le mot "amortisseurs" en gras sera donc mieux positionnée, toutes choses égales par ailleurs, qu'une page contenant ce même mot en roman.

N'en profitez pas non plus pour mettre tout votre texte en gras, ce qui donnerait un résultat assez horrible à la lecture. Seuls les mots clés importants doivent être ainsi mis en exergue.

Autre point important : la **taille des caractères**. Il semblerait que plusieurs moteurs prennent en compte cette donnée. N'hésitez donc pas à proposer (notamment pour les titres des paragraphes) les phrases contenant vos principaux mots clés en taille importante. Exemple :

Ceci est le texte d'un titre de paragraphe

Troisième point pour mettre en avant votre texte : **les liens**. Pour tous les moteurs aujourd'hui, le fait qu'un mot soit "cliquable" est important. Il est à la fois important par rapport au poids du mot en question dans la page mais également pour le positionnement de la page cible. Explications... Si vous insérez le code suivant dans votre page :

```
Voici des informations sur l'<a  
href="http://www.votresite.com/assurance.html">assurance-vie</a>
```

Le résultat sera :

Voici des informations sur l'[assurance-vie](#).

Ceci aura deux effets :

- La page contenant ce code sera mieux référencée pour l'expression "assurance-vie" car ce mot y est "cliquable".

- Mais la page "pointée par le lien" (ici "assurance.html") le sera aussi pour cette même expression (nous y reviendrons dans un prochain article consacré aux liens hypertexte).

Vous faites donc d'une pierre deux coups. Google, notamment, est très sensible au texte des liens pointant vers les pages pour classer ces pages de destination. Tenez en compte.

Voilà, par exemple, un exemple de mauvais lien :

Pour avoir des informations sur l'assurance-vie, [cliquez ici](#).

A priori, l'expression "cliquez ici" n'est pas vraiment importante pour votre activité et il y a peu de chances pour qu'un internaute la saisisse sur un moteur de recherche :-), aussi il n'est pas vraiment nécessaire de la mettre en valeur dans vos pages en le rendant cliquable...

En revanche, le texte suivant sera très bien optimisé :

Voici des informations sur l'[assurance-vie](#).

Les mots importants sont "cliquables" et en gras. Bravo !

Pendant longtemps, les moteurs ont également pris en compte les **balises <Hn>** (<H1> à <H6>) pour donner un poids aux pages web sur une requête donnée. Si un mot était compris entre des balises <H1> et </H1> (plus grande taille de titre en HTML), cela avait un poids important pour le classement du document pour ce terme. Les balises <Hn> étant tombées en désuétude, il semblerait que ce critère soit moins important aujourd'hui, même si Google semble encore bien les aimer... Cependant, certains webmasters utilisent la balise <H1> en corrélation avec les feuilles de style pour arriver à des résultats intéressants. Comme cela présente certains risques de spam, nous ne nous étendrons pas sur la question ;-)

Les moteurs prennent-ils en compte les feuilles de style ?

Justement, voici une question qui revient souvent sur les forums. Une question complexe car il semble y avoir peu de littérature sur ce point précis, que ce soit sur les web anglophone ou francophone.

Après avoir exploré de nombreux forums et lu de nombreux sites web, il semblerait que la plupart des moteurs de recherche ne prennent pas en compte les feuilles de style (CSS). Notez bien que nous n'avons aucune certitude à ce sujet et qu'il s'agit plutôt d'une supposition. Donc, si un texte est paramétré en gras dans la feuille de style correspondante, il y a de fortes chances qu'il ne soit pas considéré comme tel par le moteur.

Un "ruse" pour détourner ce fait consiste à créer une feuille de style pour le texte en question et d'y indiquer un style "roman" (font-weight:normal), puis de proposer la mise en gras dans la page elle-même.

Exemple de feuille de style :

```
.exemple {font-family:Verdana,Helvetica;font-size:12px;color:#3b3b3b;font-style:normal;font-weight:normal;text-decoration:none;}
```

Exemple de texte dans la page :

```
<p class="exemple">Ceci est un texte en roman. <B>Ceci est un texte en gras</B>.
```

Ce qui donnera dans la page :

Ceci est un texte en roman. **Ceci est un texte en gras.**

On perd, bien sûr, le fait d'indiquer la mise en gras directement dans la feuille de style, ce qui est la fonction première des CSS. En revanche, on est sûr que cette mise en exergue sera prise en compte par les moteurs de recherche. En revanche, dans cet exemple, on perd la notion de taille des caractères, importante également pour les moteurs, si elle est indiquée dans la feuille de style. Mais il est possible de la prendre également en compte au moyen de la balise . A vous de voir car, à ce moment-là, on peut quand même se demander à quoi serviront les feuilles de style si elles ne servent qu'à définir la couleur des caractères...

Nombre d'occurrences des mots

Pendant longtemps, le nombre d'occurrences d'un mot dans la page a été très important pour un moteur de recherche. Même si cette notion revêt encore une certaine importance, elle semble moins critique aujourd'hui.

En effet, de nombreux moteurs ne prennent plus en compte que les **X premières occurrences** du mot demandé dans la page.

Exemple : vous tapez le mot clé "référencement" sur un moteur. Celui-ci va aller chercher dans son index les pages qui contiennent au moins une fois ce terme. Ensuite, dans son algorithme de classement, il va partir du début du code HTML pour prendre en compte les X premières occurrences rencontrées. Imaginons que X=3. Dans ce cas, le moteur va prendre en compte, par exemple, la première occurrence du mot "référencement" dans le titre, puis dans la balise meta "Description" et enfin, la première fois où le mot apparaît dans le texte visible. Toutes les occurrences suivantes sont alors ignorées.

Dans ce cas, une page peut donc contenir 3, 5, 10 ou 100 fois le même mot, il y a de fortes chances pour que, au travers de ce système, le résultat soit le même en termes de positionnement. Les valeurs du paramètre X qui sont parvenues à nos oreilles oscillent, selon les moteurs, dans une fourchette allant de 3 à 6. En revanche, pour l'instant, impossible de savoir si Google prend en compte une telle contrainte... Dans l'absolu, ne "truffez" donc pas vos pages d'innombrables répétitions de mots, cela n'aura *a priori* pas d'influence sur les moteurs et surtout, cela déplaira énormément à vos visiteurs !

L'indice de densité des mots

Une notion longtemps adoptée par les moteurs a été l'indice de densité de mot (IDM) des termes les plus importants. En clair, pour un mot donné, **le nombre d'occurrences du mot dans la page divisé par le nombre total de mots du document**. Même si ce type de critère est à relativiser en raison du nombre d'occurrences limité d'un mot pris en compte par les moteurs (voir paragraphe précédent), l'IDM peut être pris en compte pour mieux "calibrer" le contenu de vos pages. En tout état de cause, essayez de ne jamais dépasser un indice de **10%** pour vos mots clés importants. N'oubliez pas que le site Outiref (<http://www.outiref.com/>) vous propose gratuitement un utilitaire de calcul d'IDM. Intéressant pour calculer celui des pages les mieux positionnées sur vos mots clés favoris... ;-))

Pour cela, deux solutions se présentent à vous : soit vous proposez assez de texte autour d'un mot donné si la page est courte, soit vous répétez le mot plusieurs fois si la page est longue.

Dans le premier cas, vous pouvez proposer des pages courtes, mais très "denses" (tout en tenant compte du fait qu'une "bonne" page propose au moins 100 mots). Dans le second cas, faites attention à ne pas être taxé de spam : si un mot a plusieurs occurrences, celles-ci doivent se situer en différents endroits de la page, bien espacées les unes des autres.

Les différentes formes, l'éloignement et l'ordre des mots

Si vous en avez la possibilité, n'oubliez pas d'insérer dans vos pages les **féminins**, les **pluriels** de vos mots clés importants, ainsi que certains termes qui auraient la **même racine** (poisson, poissons, poissonnerie, poissonneries, poissonnier, poissonniers, poissonnière, poissonnières, etc.). Rappelez-vous qu'une page bien positionnée sur "chien" ne le sera pas obligatoirement sur "chiens", notamment sur Google. Pensez donc à indiquer dans vos pages les différentes occurrences des termes susceptibles d'être saisis sur un moteur de recherche. N'oubliez pas que chaque page de votre site peut être optimisée en fonction de certains mots clés. Ne tablez pas que sur votre page d'accueil pour être bien positionné ! On entend souvent dire, dans le domaine du référencement, qu'il est complexe de voir une page très réactive (donc bien positionnée) sur plus de 2 ou 3 mots clés ou expressions. Ce n'est certainement pas faux.

Privilégiez également les couples de mots et les **expressions** ("chaussures de tennis", "société de service", etc.), c'est ce qui fera peut-être la différence sur des recherches plus précises émanant d'un internaute. N'oubliez pas les termes connexes : dans "restaurant à Paris", il y a "restaurant" - qui est un terme important - mais également "Paris". Essayez de faire en sorte d'être également bien classé sur ce couple de mots.

D'autre part, si vous désirez être positionné, par exemple, sur l'expression "paris dakar", faites en sorte que les deux mots soient présents dans la page l'un à côté de l'autre et non **pas éloigné**. En d'autres termes, une page contenant "paris dakar" l'un à côté de l'autre sera plus réactive qu'une page contenant "paris" au début et "dakar" à la fin.

L'**ordre** est également important, notamment sur Google. Une page contenant "dakar paris" sera moins réactive sur l'expression "paris dakar"... Tenez-en compte !

La taille des pages

Privilégiez les pages plutôt **courtes** qui parlent d'un **thème unique** plutôt que de longs documents qui traitent de nombreux concepts très différents. Nous en avons parlé le mois précédent. Certains moteurs tentent de faire ressortir d'un document l'idée principale d'un document et en tiennent compte dans leurs classements. Facilitez-leur la tâche...

Proposez sur votre site plutôt de nombreuses petites pages mono-thème (mais contenant quand même au minimum 100 mots en texte) que de longs documents indigestes pour les moteurs... et l'internaute, d'ailleurs !

Langue du texte

Evitez également les pages **bilingues** ou **trilingues**, comme nous l'avons déjà vu dans notre article du mois dernier. Comme pour les titres, les moteurs auront du mal à bien traiter une page contenant des termes dans plusieurs langues différentes. **Privilégiez les pages monolingues !**

Dernier point : lorsque les pages de votre site seront indexées par les moteurs, vous aurez certainement besoin — à un moment ou à un autre — de connaître combien d'entre elles sont présentes dans l'index de ces outils. La plupart des moteurs vous permettent de le faire simplement à l'aide d'une syntaxe telle que host: ou domain:. Mais d'autres ne proposent aucune fonction avancée pour ce type de recherches. Mais, même si la syntaxe de recherche avancée existe, vous aurez peut-être du mal à identifier telle ou telle page de façon précise.

Vous pouvez insérer au préalable dans vos pages quelques mots très spécifiques ou inconnus (par exemple un code précis donné à chaque fichier, de type ABX20-08-4), de façon transparente ou

non (dans les balises <META> ou les attributs ALT des balises IMG, par exemple) pour retrouver facilement vos documents. Une simple recherche sur ABX20-08*, par exemple, retrouvera tous les documents qui contiennent un terme qui débute par ces lettres, donc certainement les vôtres. Cela vous permettra également de répertorier ceux qui ne sont pas présents dans la base.

Pour résumer...

Voici quelques conseils pour bien optimiser le texte visible de vos pages :

- Soignez au maximum le contenu du **premier paragraphe** de vos pages (les 2 ou 3 premières phrases).
- Créez des pages **courtes, monothèmes, monolingues**, mais contenant quand même au moins **100 mots**.
- **Mettez en exergue** les mots clés importants pour votre activité : gras, taille, texte des liens, etc.
- Prenez en compte les **différentes formes** des mots : féminin/masculin, singulier/pluriel, etc.
- Prenez en compte la **proximité** des mots entre eux ainsi que leur **ordre**.
- Ne recherchez pas obligatoirement à "truffer" vos pages de mots clés identiques, cela n'aura que peu d'effet sur le "poids" de vos pages.
- N'oubliez pas que **chaque page** de votre site peut être optimisée pour un mot clé ou une expression. Ne misez pas tout sur la seule page d'accueil.