

## Référencement : La règle des "3C" (2ème partie)

[Retour au sommaire de la lettre](#)

*Le référencement d'un site web est avant tout tributaire de la façon dont le site en question a été conçu. Dans ce cadre, une stratégie gagnante prendra en compte divers aspects de cette conception, dans le cadre d'un concept global intitulé "Règle des 3C". Le mois dernier, la Conception du site avait été étudiée. Ce mois-ci, nous examinons le deuxième "C", soit le Code HTML : comment créer un code qui prenne en compte toutes les particularités et contraintes des moteurs de recherche tout en étant parfaitement optimisé par rapport à leurs critères de pertinence ? Réponse sous la forme d'un "mémo" et d'un "pense-bête" qui vous permettra de ne pas rater une étape importante lors de vos réflexions...*

Pour optimiser au mieux le référencement d'un site, il est devenu important, voire primordial, de le penser dès le départ pour être compatible avec les différents moteurs de recherche qui vont venir le visiter, grâce à leurs *spiders*, robots qui viennent "aspirer" les pages web et suivre les liens qu'elles contiennent.

Pour qu'un site soit parfaitement "compris" et "analysé" par les moteurs de recherche, il faut donc qu'il ait été pensé pour être compatible avec les critères d'exploration et de pertinence de ces outils. Nous allons, dans cette série de trois articles, expliciter une règle qui nous est chère, et que nous avons pu expérimenter sur de nombreux sites, celle des "3C".

Le mois dernier, nous avons abordé le premier "C" soit celui de la **Conception initiale** du site. Ce mois-ci, nous traitons le second, le **Code HTML**. Comme d'habitude, nous allons donc étudier cette partie (votre code HTML est-il optimisé pour les moteurs de recherche ?) sous la forme d'un "momento" ou "pense-bête" afin de vous aider à ne rien oublier dans le cadre de la création - ou de la refonte - de votre site... Commençons par le commencement avec le "Header" (l'en-tête) des pages...

### Code HTML > Header

- Votre code commence-t-il par les zones suivantes ? :

```
<HTML>
```

```
<HEAD>
```

```
<TITLE>Titre de votre page</TITLE>
```

ce qui sera garant d'une bonne prise en compte de la balise TITLE par les différents moteurs de recherche...

- Les balises meta "Keywords" (peu utiles aujourd'hui pour le référencement) et "Description" (utiles pour mieux maîtriser la façon dont les moteurs affichent votre site dans leurs résultats) sont-elles prévues pour recevoir un contenu strictement en rapport avec le contenu de la page ? Si ce n'est pas le cas, il vaut mieux ne pas les proposer dans votre code car cela nuira à la bonne compréhension de votre contenu par les spiders... Et les moteurs.

- Votre Header propose-t-il une balise meta "language" sous cette forme ? :

```
<meta http-equiv="content-language" content="fr">
```

Cette balise est souvent lue par les moteurs, aussi il peut être intéressant de l'indiquer, notamment si vos pages sont dans une langue spécifique ou si l'ajout de mots dans une autre langue peut induire le moteur en erreur (cas d'un site en français qui vend du vin italien par exemple).

- Idem pour les balises Meta éventuelles :

```
<META NAME="keywords" lang="fr" CONTENT="contenu de la balise en français">
```

```
<META NAME="keywords" lang="en" CONTENT="contenu de la balise en anglais">
```

- Avez-vous prévu d'éventuelles balises Meta "Robots" pour indiquer certaines actions (indexation, suivi des liens) aux spiders des moteurs, notamment sur des pages de test ?

- Concernant les méta-données Dublin Core, il est conseillé de ne les utiliser que si elles ont une utilité côté serveur et webmaster. Elles n'ont en revanche pas d'intérêt pour le référencement...

- Le codage des lettres accentuées a-t-il été prise en compte ? Si les pages de votre site sont codées en ISO-8859-1 (balise meta : <meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />) ou si vos pages sont en UTF-8 (balise meta : <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />), les lettres accentuées doivent être codées de façon cohérente avec le codage choisi au préalable ainsi que pour le serveur web ou le serveur mysql éventuellement utilisé. Si vous n'êtes pas sûr de vous, codez ces lettres en HTML ("&eacute;" pour le "é" par exemple)...

Voici un exemple de "Header" HTML compatible avec les moteurs de recherche :

```
<html>
<head>
<title>R&eacute;f&eacute;rencement et moteur de recherche : toute l'info et
l'actu sur le référencement avec Abondance</title>
<meta name="description" content="Abondance d'infos sur le
r&eacute;f&eacute;rencement et les moteurs de recherche : description des
moteurs, actualit&eacute;, fags, outils d'audit, m&eacute;thodologie de
r&eacute;f&eacute;rencement, articles, offres d'emploi, bibliographie,
etc.">
<meta name="keywords" content="référencement, abondance, moteur de
recherche, recherche d'information, r&eacute;f&eacute;rencement">
<meta http-equiv="content-language" content="fr">
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
<link title="styles abondance" type="text/css" rel="stylesheet"
href="styles-homepage.css">
<link rel="shortcut icon" href="http://www.abondance.com/Bin/favicon.ico">
<SCRIPT LANGUAGE="javascript"
SRC="http://www.abondance.com/js/scripts.js"></SCRIPT>
</head>
```

- Externalisation des éléments CSS et JavaScript : plus la taille de la page est limitée à la portion de contenu textuel, plus la page est réactive. Donc CSS et JS doivent être appelés à l'aide de fichiers externes à la page. Exemples :

```
<link title="styles abondance" type="text/css" rel="stylesheet"
href="styles-homepage.css">
<SCRIPT LANGUAGE="javascript"
SRC="http://www.abondance.com/js/scripts.js"></SCRIPT>
```

- Si vous utilisez des frames, avez-vous fait tout ce qu'il fallait pour qu'elles soient bien prises en compte par les moteurs de recherche (voir nos nombreux articles précédents à ce sujet dans les archives de cette lettre) ? Idem pour le Flash...

### **Code > Contenu Textuel**

- Vos pages sont-elles prévues pour proposer au moins 200 mots en texte visible ?

- Avez-vous la possibilité (ou donnez-vous la possibilité aux rédacteurs) de créer /modifier des liens et le texte des liens affichés dans les pages ?

- Vos balises Meta "Description" sont-elles prévues pour recevoir au plus 200 caractères, soit une extension du contenu de la balise TITLE ?

- Votre code HTML est-il prévu pour ne pas être surchargé en commentaires (non lus par les moteurs), ce qui peut en revanche alourdir le poids (en Ko) des documents ? Pas de balises commentaires (<!-- -->) donc, autres que les informations nécessaires au balisage utile au développement...

- Si vous utilisez du code JavaScript, avez-vous vérifié qu'il était compatible avec le fonctionnement des spiders des moteurs de recherche (voir les nombreux articles à ce sujet dans les archives de cette lettre) ?

- Avez-vous passé vos pages en test sur le site Spider Simulator (<http://www.spider-simulator.com/>) ou les avez-vous regardé dans la version "Cache" de Google avec l'option "uniquement le texte mis en cache" afin de visualiser la façon dont le moteur comprend vos contenus ?
- Avez-vous vérifié, notamment grâce à l'outil Spider Simulator, que la mise en page de vos pages n'empêche pas le texte le plus important d'être proposé en premier (le "plus haut" possible) aux moteurs de recherche ?
- Utilisez-vous les div CSS, qui s'attachent à l'essentiel (le contenu) plutôt que les tableaux qui alourdissent la taille d'une page ?

#### **Code > Mise en exergue des mots importants de la page**

- Vos pages sont-elles prévues pour recevoir un titre (balise TITLE) (Eh oui, on voit de tout sur le Web...) ? Pourrez-vous le modifier au quotidien quelle que soit la page du site ?
- La mise en page utilise-t-elle des balises "Hn" (H1 à H6) pour ses différents niveaux de titres éditoriaux ?
- Le code HTML de votre page est-il prévu pour mettre en exergue les mots importants qui la décrivent dans les différentes "zones chaudes" : balise TITLE, balise meta "Description", titre éditorial mis dans une balise H1, début du texte, etc. ?
- Les balises H1, H2, etc. sont-elles bien redéfinies dans les feuilles de style pour apparaître de façon "agréable" dans un navigateur ?
- Vos images contiennent-elles des options "ALT" sur les balises "IMG" ? Ces options textuelles contiennent-elles du contenu intelligible et descriptif pour les moteurs de recherche ? Comment les rédacteurs de contenu auront-ils accès à cette zone ?
- Les mots importants décrivant la page sont-ils indiqués en gras ? Le gras est-il présent dans le code HTML (sous forme de balises <B> ou <STRONG>), ce qui permettra aux moteurs de les prendre en compte comme tel, ou dans les feuilles de style (que ne lisent pas les moteurs) ?
- Avez-vous prévu des intitulés de liens intelligibles, en évitant les "Cliquez ici" ou les "Pour en savoir plus", afin de soigner au mieux la notion de "réputation" de votre site ?
- Avez-vous vérifié que vous ne disposez aucun contenu caché dans vos pages, que l'internaute voit tout ce que le spider voit et vice-versa ?

#### **Code > Homogénéisation des différentes "zones chaudes"**

- Avez-vous pensé l'organisation de votre code HTML pour que les "zones chaudes" suivantes soient toutes cohérentes au niveau de leur contenu, ce qui est essentiel pour un bon référencement :
  - \* Balise Title ;
  - \* Balise Meta "Description" ;
  - \* Balise Meta "Keywords" (si vous l'affichez dans votre code) ;
  - \* Titre éditorial (dans des balises H1 ou H2) ;
  - \* Début du texte visible de votre page (le premier paragraphe, les 2 ou 3 premières phrases).
- Si vous utilisez un CMS (*Content Management System*), l'avez-vous configuré pour que cette homogénéité se fasse de façon quasi automatique ? Sans oublier la date de dernière modification à passer en *header* (*Last-modified*) qui peut toujours être utile pour indiquer au moteur que le changement de date rapide. Les pages HTML ont ce *header* par défaut, ce n'est pas le cas des pages dynamiques (SPIP et Joomla gèrent cela par défaut).

#### **Code > Compatibilité W3C**

- Avez-vous passé les pages les plus importantes de votre site web, et notamment sa page d'accueil, au validateur de code du W3C (<http://validator.w3.org/>) ? Ont-elle passé ce test sans problème ? Avez-vous corrigé les éventuelles erreurs signalées ?
- Avez-vous testé le site une fois conçu avec au moins Firefox et Internet Explorer pour vérifier que tout s'affiche ? Si des éléments ne s'affichent pas sous Firefox notamment, il y a de grandes chances que Google ne les voie pas non plus.
- Idéalement, votre code présente-t-il tous les atouts et critères d'accessibilité (<http://www.accessiweb.org/>) ?

Bref, votre code est-il :

- parfaitement compatible avec toutes les contraintes induites par les moteurs ?
- optimisé par rapport aux critères de pertinence de ces mêmes moteurs ?

Si oui, et si vous avez déjà appliqué la première règle édictée dans notre précédent article, vous êtes sur la piste d'un excellent référencement. Il ne vous reste donc plus qu'à suivre nos conseils pour le troisième "C", que nous étudierons le mois prochain... D'ici là, bon Code !!

*Olivier Andrieu*

**Réagissez à cet article sur le blog des abonnés d'Abondance :**  
<http://abonnes.abondance.com/blogpro/2007/09/rfrencement-la-rgle-des-3c-2me-partie.html>