

## Compatibilité W3C et référencement : quel impact ?

[Retour au sommaire de la lettre](#)

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

*On entend souvent dire que, pour être "bien référencée", une page doit avoir son code HTML en conformité avec les standards et doit donc passer sans heurts le contrôle du parfois tant redouté valideur du W3C. Mais quel est le réel impact de cette validité sur le référencement ? Les moteurs de recherche classent-ils réellement mieux les pages qui sont "W3C friendly" ? Jusqu'où aller dans cette analyse ? Réponse...*

Le W3C (Word Wide Web Consortium) est un organisme créé en 1994 pour standardiser et uniformiser le langage utilisé dans les technologies web (HTML mais aussi XML, CSS, XHTML, RDF...). Son rôle est de définir la grammaire et l'orthographe à utiliser dans les pages web, de façon à ce que celles-ci s'affichent correctement sur les ordinateurs utilisés par les internautes du monde entier.

Une page web respectant les normes W3C est donc avant tout une page bien écrite d'un point de vue technique. Ceci a un impact significatif sur l'affichage dans le navigateur et favorise l'accessibilité des internautes. Mais est-ce que cela joue sur le référencement et la façon dont les moteurs de recherche perçoivent la page ?

### Les moteurs et le W3C

En mars 2008, Google proposait un article sur Accessible Search (<http://googlewebmastercentral.blogspot.com/2008/03/tips-for-making-information-universally.html>), un moteur proposant uniquement des résultats utilisables par les personnes malvoyantes, et basé en grande partie sur le respect des normes W3C. Dans un autre article datant de septembre 2008, intitulé *Workin' it on all browsers* (<http://googlewebmastercentral.blogspot.com/2008/09/workin-it-on-all-browsers.html>), Google rappelait l'importance du respect des normes W3C pour l'affichage d'un site sur tous les navigateurs.

Tout ceci semble indiquer que le respect des normes W3C est pris en compte par Google, au moins pour l'aspect qualitatif des sites web.

Si l'on consulte le guide en ligne Google destiné aux webmasters, on trouve dans la section, "Assurez-vous que votre site s'affiche normalement sur différents navigateurs" (<http://www.google.com/support/webmasters/bin/answer.py?answer=100782>) des conseils tels que :

*Votre site peut apparaître correctement dans certains navigateurs même si votre code HTML est incorrect. Par contre, il n'est pas garanti qu'il s'affichera correctement dans tous les navigateurs, ni dans les versions à venir de tout navigateur. La meilleure façon de vous assurer que votre page présente le même aspect dans tous les navigateurs est de la créer en utilisant des codes HTML et CSS corrects, puis de la tester dans le plus grand nombre de navigateurs possible. Un code HTML correct et clair équivaut à une bonne police d'assurance. L'utilisation du code CSS permet de distinguer la présentation du contenu et d'obtenir un meilleur rendu et un chargement des pages plus rapide. Les outils de validation, tels que les valideurs HTML et CSS fournis par la société W3 Consortium, vous permettent de vérifier votre site.*

La validation W3C est donc utile avant tout pour l'internaute car elle apporte une meilleure qualité d'affichage des pages web.

Il existe un service gratuit, accessible à l'adresse <http://validator.w3.org/> et permettant d'analyser en ligne le code source de ses pages et d'en relever les problèmes. Un véritable

correcteur orthographique pour site web, qui permet d'obtenir un logo « valide W3C » à placer sur site.

Est- il utile pour le référencement de passer des heures à peaufiner le code source pour obtenir le fameux logo ?

Le message relayé Matt Cutts, porte parole officiel de Google, a toujours été sibyllin à ce sujet. Tout récemment, au mois de septembre, Matt Cutts répondait en vidéo (<http://www.youtube.com/watch?v=FPBACTS-tyg>) à une question d'un internaute lui demandant pourquoi la homepage Google n'était pas valide W3C. La réponse était que Google préférerait travailler l'efficacité et la compatibilité de son site plutôt que le rendre parfaitement valide W3C. « Faites ce que l'on dit, pas ce que l'on fait » pourrait être la morale de l'histoire...



De toutes manières, rappelait Matt Cutts, il existe très peu de sites valides W3C et Google ne tient donc pas compte du respect des normes pour favoriser tel ou tel site dans les résultats de recherche car cela serait pénalisant pour le plus grand nombre.

La même chose apparaît dans le guide en ligne Google précédemment cité : « Bien que nous recommandions l'utilisation d'un code HTML correct, cela n'a normalement aucune incidence sur la façon dont Google explore et indexe votre site. » (Notez bien le « normalement »).

Pour s'en convaincre, on peut par exemple consulter le résultat d'une expérience menée en 2007 : [http://cass-hacks.com/bits\\_n\\_bytes/validation\\_not\\_important/](http://cass-hacks.com/bits_n_bytes/validation_not_important/). Lors de cette expérience, une page web volontairement bourrée d'erreurs a été mise en ligne, ce qui ne l'a pas empêchée d'être indexée par Google et même d'acquérir des positions intéressantes.

En ce qui concerne les autres moteurs, on observe plus ou moins le même comportement. Microsoft Bing propose plusieurs conseils aux webmasters dans son blog dédié et dans un article intitulé Common errors that can tank a site (<http://www.bing.com/community/blogs/webmaster/archive/2009/07/31/common-errors-that-can-tank-a-site-sem-101.aspx>), on trouve ceci :

*« Si votre code HTML est mauvais, vous risquez d'avoir des problèmes d'exploration. Mais vous risquez de ne pas savoir que les problèmes existent si votre test a seulement consisté à « De quoi ça a l'air dans mon navigateur ? ». Les navigateurs de bureau modernes sont tout à fait capables de deviner ce que vous vouliez probablement dire et d'en faire une présentation exploitable à l'écran. Ils peuvent souvent se débrouiller avec du code qui est libre comme l'air vis à vis du respect des standards. Mais les robots des moteurs de recherche ne sont pas aussi flexibles que les navigateurs de bureau, et les problèmes de code peuvent souvent les induire en erreur et faire stopper l'exploration de votre site. En complément de cela, les navigateurs des interfaces mobiles ne sont pas aussi accommodants avec le code source pauvrement écrit que les navigateurs de bureau. Tout ce que vous pouvez faire pour rendre votre code solide et conforme aux standards est bon, à la fois pour les utilisateurs et les moteurs. »*

Selon Microsoft, les standards W3C ont une utilité certaine pour améliorer la qualité d'un site web et cela peut jouer sur l'exploration d'un site par les moteurs. Mais cela peut-il jouer sur le positionnement ? La question n'est pas tranchée.

## **Le W3C : pour ou contre**

Faut-il se soucier ou non du respect du W3C si cela n'influe pas sur le référencement ? La réponse n'est pas simple... Certaines erreurs W3C sont des peccadilles qui n'ont aucun impact sur l'affichage dans les navigateurs et sur la lisibilité des pages. D'autres erreurs sont plus gênantes car elles touchent à la structure profonde d'un site et à son fonctionnement. Une chose est sûre : quand une page web contient des centaines d'erreurs W3C, on peut se poser des questions sur son fonctionnement et sur sa prise en compte par les moteurs de recherche.

W3C ou pas W3C, le débat divise profondément les webmasters, entre ceux qui pensent qu'il est inutile de se casser la tête à rédiger son code source et d'autres qui pensent au contraire, qu'il faut respecter la grammaire et l'orthographe web (voir par exemple les avis publiés sur <http://www.arkantos-consulting.com/articles-marketing-internet/webdesign/20070803-le-mythe-de-la-conformite-w3c.php> et <http://blog.axe-net.fr/site-conforme-w3c-xhtml-mon-oeil/>).

La validation W3C peut servir ces différents objectifs :

- Assurer l'accessibilité de tous les internautes.

A l'exception des responsables de portails associatifs ou institutionnels, peu de webmasters se préoccupent de savoir si un internaute handicapé peut facilement utiliser un site web. Néanmoins, un robot de moteur de recherche peut être considéré comme un internaute malvoyant et handicapé (il ne peut pas voir les images par exemple) : optimiser un site en vue de l'accessibilité c'est aussi optimiser pour les moteurs.

- Assurer la compatibilité d'un site web avec différents outils de navigation

On parle souvent de la compatibilité entre différents navigateurs (Firefox *versus* Internet Explorer) mais on peut parler aussi de la percée importante du web mobile. Les téléphones de dernière génération peuvent lire directement les sites web, à condition que ceux-ci respectent la bonne grammaire. Voilà un point non négligeable en faveur de la validation W3C.

- Utiliser un langage universel

Un site web conforme aux standards sera utilisable dans tous les pays du monde et son code source sera compréhensible par tous les webmasters pouvant être amenés à travailler dessus. Respecter le W3C c'est donc penser à la globalisation du web et pas seulement aux internautes français qui peuvent accéder au contenu.

Comme on le voit, le respect des normes W3C tient de la philosophie de vie : c'est l'idée de rendre un contenu universel et accessible à tous, y compris aux moteurs de recherche et aux internautes les plus éloignés de la cible privilégiée par un site web. On ne travaille pas son code source pour asseoir sa réputation et son classement auprès des moteurs, mais plutôt pour proposer un site web compatible et efficace avec toutes les technologies.

Une chose pourrait néanmoins faire bientôt pencher la balance du côté des pro-W3C : la prise en compte des balises sémantiques par les moteurs de recherche.

Dans un article daté du 26 octobre, Help us make the web better : an update on rich snippets (<http://googlewebmastercentral.blogspot.com/2009/10/help-us-make-web-better-update-on-rich.html>), Google présentait de nouvelles avancées sur la prise en compte des balises sémantiques pour obtenir des résultats Google enrichis (NDLR : voir également notre article sur les "rich snippets" dans la lettre R&R de septembre 2009). Cette innovation paraît prometteuse pour les webmasters qui souhaitent se démarquer de leurs concurrents et Google n'est pas le seul à faire des démarches en ce sens (voir notamment toutes les expériences

menées par Yahoo dans son programme Search Monkey :  
<http://www.ysearchblog.com/2009/08/28/see-more-searchmonkey/>).

Il est clair que l'utilisation de ces nouvelles balises devra se faire de façon normalisée et le consortium W3C pourrait revenir sur le devant de la scène avec ses recommandations sur le web sémantique (<http://www.w3.org/standards/semanticweb/>).

## **Conclusion**

Le respect des standards W3C est utile pour les internautes, ainsi que pour les moteurs de recherche. Tout d'abord parce qu'un site normalisé est plus facilement explorable et indexable, mais aussi parce qu'un site *User-Friendly* peut bénéficier d'une prime de classement et de notoriété. En effet un site bien conçu et bien construit sera plus souvent visité et obtiendra facilement des liens et de bonnes critiques dans les médias sociaux, ce qui le poussera dans le classement des moteurs.

Faut-il pour autant corriger 100% des erreurs de son code source et atteindre la perfection, quitte à recommencer le travail à chaque mise à jour des pages web ? La réponse est non, dans la mesure où il s'agit là d'un travail laborieux qui est actuellement peu valorisé par les moteurs.

L'idée est plutôt de corriger les grosses erreurs W3C, celles qui peuvent nuire à l'affichage et à l'indexation d'un site. Parmi elles on peut citer l'absence de document de déclaration (doctype), les problèmes d'imbrication de balises, les confusions entre apostrophes et double cotes, les balises non fermées... L'outil de validation W3C (<http://validator.w3.org/>) est donc indispensable pour tout bon webmaster.

Respecter la grammaire et l'orthographe du web n'est pas toujours facile mais cela peut rapporter gros !

**Guillaume Thavaud, cellule de veille ([veille@brioude-internet.fr](mailto:veille@brioude-internet.fr)) de la société Brioude-Internet Référencement (<http://www.referencement-2000.com/>)**

**Réagissez à cet article sur le blog des abonnés d'Abondance :**  
<http://abonnes.abondance.com/blogpro/2009/11/compatibilite-w3c-et-referencement-quel.html>