

L'audit SEO : que doit-il contenir ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Vous êtes certainement nombreux à vous demander si votre site est optimisé pour le référencement naturel et les moteurs de recherche. Voici une méthodologie simple et claire qui décrit les différentes étapes à suivre pour être sûr que vous n'avez pas manqué une étape et pour valider le fait que vos pages soient "Google friendly"...

Pour bien référencer un site Internet, il faut impérativement avoir réalisé un audit SEO complet de celui-ci. Cela permet de savoir à la fois ce qui a été fait, ce qui fonctionne bien, ce qui freine le référencement et ce qu'il reste à faire pour augmenter le trafic et les ventes. Ce guide vise donc à lister tout ce qu'un audit SEO doit contenir.

L'intérêt de l'audit

Comme expliqué dans l'introduction, un audit SEO va servir à plusieurs choses, à savoir :

- Que cherchent les internautes ? L'audit va permettre de dresser une liste de mots clés et d'expressions qui vont correspondre aux différents besoins des internautes sur un thème donné. Avec cela, on va savoir quels contenus et quelles offres doivent être mise en place.
- Qui sont mes concurrents et les leaders d'opinion ? Cette partie de l'audit va tout simplement vous permettre d'adapter les offres et les contenus du site pour se démarquer, tout en rendant possible la mise en place d'une veille concurrentielle et technologique. En d'autres termes, cela permet au site d'être bien plus réactif sur son secteur d'activité.
- Sur quoi suis-je déjà positionné et sur quoi dois-je me positionner ? C'est la partie la plus logique. Quel que soit le référenceur ou l'agence qui travaille sur un site, cette personne a besoin de savoir où en est le travail de positionnement afin de définir une stratégie de travail à court, moyen et long terme.
- Quels sont les freins de mon site et comment l'améliorer ? L'audit SEO va aussi s'attarder sur la manière dont est conçu le site. D'une part, cela permet de détecter tous les freins et bugs d'un site pour le référencement naturel, tout en mettant en avant tous les points qui peuvent être facilement optimisés.
- Quels contenus et quels liens dois-je créer ? Cette partie est très importante : elle va permettre au référenceur de savoir quels contenus il doit approfondir ou créer afin de se positionner sur tel et tel mot clé. Cela va aussi lui permettre de savoir quels liens il pourra créer, à la fois en interne et en externe.

Que doit contenir un audit SEO ?

L'audit de référencement naturel se décompose en plusieurs points.

Les mots clés

C'est la première étape, et elle est cruciale pour le bon déroulement de l'audit. Le but est de trouver tous les mots clés pertinents qu'il faudra travailler, que ce soit un mot clé généraliste ou une expression de longue traîne avec de nombreux mots clés différents. Pour bien travailler cet aspect, voici quelques conseils...

- Avant toute chose, pensez toujours que l'internaute sait ce qu'il veut, mais qu'il ne sait pas toujours le formuler correctement ni l'écrire de la bonne façon. Autrement dit, pensez à inclure dans votre audit les fautes de frappe ainsi que les fautes d'orthographe.

- Sur une feuille ou un document Word, commencez par écrire à chaud tous les mots clés qui vous viennent en tête par rapport au site à référencer. Il faut que cette liste soit la plus vaste possible. Elle doit inclure le nom de l'entreprise, ceux des responsables, la localisation géographique, le secteur d'activité, les produits et services proposés, les besoins relatifs à ces mêmes services et produits, ... Si nous prenons comme exemple le site d'Abondance, voici ce que donnerait le tout début de cette liste :

1. Abondance
2. Olivier Andrieu
3. Heiligenstein
4. référencement
5. Audit
6. Audite
7. SEO
8. newsletter seo
- 9...

Pour affiner cette première liste de mots clés, il existe un moyen très simple : discuter avec l'entreprise, les salariés et les clients. Ils seront presque à chaque fois la meilleure source d'information, bien plus que les autres outils que nous allons présenter ci-dessous. C'est encore plus vrai quand le site est situé sur un secteur technologique ou de niche (car difficilement compréhensible par chacun et par le référenceur).

Une fois cette liste bien conçue, passez chacun de vos mots clés dans un générateur Google Suggest pour trouver des suggestions plus ou moins fiable de ce que les internautes recherchent quand ils commencent à taper chacun de ces mots. L'outil conçu par SeoBlackOut est d'ailleurs très bien fait : <http://www.seoblackout.com/2010/01/27/google-suggest/>.

Si nous prenons l'exemple du mot clé "Audit Seo", on peut ainsi trouver :

- audit seo gratuit
- audit seo ligne (sous entendu « en ligne »)
- audit seo pdf

La liste devrait donc s'étoffer très vite. L'étape qui suit est de passer le tout dans le générateur de mots clés d'Adwords : <https://adwords.google.com/select/KeywordToolExternal>.

Copiez-y votre liste de mots clés pour trouver d'autres expressions, ainsi que pour analyser le nombre de clic moyen possible sur chacun d'entre eux. La case "Afficher uniquement les idées directement liées à mes termes de recherche" ne doit pas être cochée. D'ailleurs, conservez l'outil d'Adwords pour plus tard afin de faire une analyse du contenu des sites de vos concurrents.

Mot ou expression	Site Web	Catégorie
Abondance Olivier Andrieu Heiligenstein référencement Audit	www.google.com/page.html	Habillement

Afficher uniquement les idées directement liées à mes termes de recherche ?

Zones géographiques : Tous Langues : Français Appareils : ordinateurs de bureau et portables

La liste devrait être désormais assez longue (plusieurs centaines de mots clés en général). Faites le ménage dedans car il arrive régulièrement que des mots clés trop éloignés s'insèrent au passage. Repassez ensuite toutes les nouvelles expressions dans Google Suggest pour trouver encore d'autres mots clés.

Vous pouvez également utiliser un dictionnaire de synonymes pour que votre audit puisse faire appel à l'ensemble du champ sémantique d'un secteur d'activité (formulation soutenue, familière, technique, courante, ...). Vous pouvez par exemple utiliser ces deux sites :

<http://www.synonymes.com/>

<http://www.crisco.unicaen.fr/cgi-bin/cherches.cgi>

Enfin, vous pouvez aussi faire appel à un outil qui analyse la densité des mots clés sur votre propre site, sur celui d'un concurrent ou encore sur un site d'information proche de votre thématique. Cela permet parfois d'ajouter d'autres expressions à la liste. Vous pouvez par exemple utiliser celui-ci : <http://tools.seobook.com/general/keyword-density/>

Les concurrents

Une fois ce travail réalisé, la liste de vos adversaires va venir d'elle même. La liste des mots clés doit en effet contenir des expressions génériques et d'autres de longue traîne. Lors de l'analyse, les principaux concurrents doivent donc ressortir naturellement. Il existe en effet plusieurs moyens de les détecter :

- En parlant avec le responsable de l'entreprise, les clients ou les employés (attention, ils ne donneront pas forcément les mêmes concurrents).
- En tapant dans Google les principaux mots clés de votre liste.

Attention, il existe selon moi différents types de concurrents sur Internet :

- le concurrent direct, avec des produits et services équivalents.
- le concurrent indirect, qui répond aux mêmes besoins avec des produits différents.
- le faux concurrent, c'est-à-dire celui qui va créer un site d'information ou dédié aux passionnés et qui va donc prendre des positions sur Google.
- les *money sites* qui se font de l'argent sur un secteur d'activité, sans pour autant réellement travailler dedans (affiliés, comparateurs de prix, sites créés pour AdSense,...).

Cette liste de concurrents est utile à plusieurs niveaux pour le référencement. On va pouvoir :

- Connaître les offres et contenus des concurrents pour essayer de faire mieux ou d'innover.
- Pouvoir réagir rapidement à une modification du marché (nouveaux produits, nouvelles technologies, nouveaux concurrents...).
- Mettre en place une veille sur les contenus des différents sites traitant de la thématique pour s'en inspirer pour ses propres contenus.
- Savoir si un concurrent progresse en référencement et trouver quels leviers il utilise pour y parvenir.

Les contenus

Votre liste de mots clés correspond à une liste de besoins, que votre site doit combler pour vendre des services, des produits, récupérer un email ou pour faire fonctionner des espaces publicitaires. Autrement dit, chaque mot clé (ou groupe de mots clés quand ils sont proches) doit correspondre à une page de votre site. Si nous reprenons l'exemple d'Abondance, nous devrions avoir théoriquement une page sur Olivier Andrieu, une autre sur ce qu'est le référencement naturel, une troisième sur les prestations du site pour un audit SEO, ... Pour chaque mot clé, les contenus devront répondre à des questions simples :

- Quoi ?
- Qui, par qui et pour qui ?
- Quand ?
- Où ?
- Pourquoi ?
- Comment ?

Si vous appliquez ces questions basiques à toutes les expressions que vous avez listé, vous allez pouvoir créer un voir plusieurs contenus. En fait, chaque question peut faire elle-même l'objet d'un contenu unique. Toujours dans le même exemple, voilà ce que l'expression "Olivier Andrieu" pourrait donner :

- Qui est Olivier Andrieu ?
- Abondance, la société de référencement d'Olivier Andrieu
- Abondance, déjà X années pour l'entreprise d'Olivier Andrieu.
- Quand Olivier Andrieu a t-il créé la société Abondance ?
- Heiligenstein, le fief d'Olivier Andrieu.
- ...

D'ailleurs, Google Suggest vous aura sûrement donné des expressions de longue traîne qui sont elles-mêmes des questions comme "comment faire pour" ou encore "que veux dire". A partir de là, il ne vous reste plus qu'à rédiger vos différents contenus, et à faire des liens logiques entre ceux-ci.

Le code source

Une fois que nous avons nos mots clés, les contenus qui vont avec et la liste de nos concurrents, il est temps de s'attaquer au site en lui-même et à la manière dont il a été créé.

Title et balises métras :

C'est la base de la base. Toutes vos pages doivent posséder une balise Title et une balise méta "description". La première a un fort poids sur le positionnement de la page tandis que la seconde est le texte qui sera utilisé par Google pour l'affichage des résultats (ce qui peut donc augmenter ou baisser le taux de clic).

La balise Title doit faire moins de 100 caractères maximum, être unique et doit refléter le plus possible le contenu réel de la page.

```
<title>Mon super produit</title>
```

La balise méta description doit quant à elle faire 160 caractères maximum, contenir les mots clés ciblés par la page, être unique et doit être "vendeuse" pour inciter au clic.

```
<meta name="description" content="Ceci est la description complète de mon super produit" />
```

Pour savoir si vous avez ces deux balises métras sur l'ensemble de votre site ou si celles-ci sont dupliquées, il existe trois outils vraiment utiles :

- Le logiciel gratuit Xenu Link Sleuth : <http://www.seomix.fr/referencement/naturel/xenu/>
- Le logiciel payant SeeUrank Falcon : <http://www.seomix.fr/referencement/naturel/seeurank-falcon/>
- Le centre pour Webmasters de Google : <http://www.google.fr/webmasters/>

Description Meta	Pages
Balises META description en double	2
Balises META description longues	0
Balises META description courtes	0

Balise title	Pages
Balises title manquantes	0
Balises title en double	36
Balises title longues	0
Balises title courtes	0
Balises title non informatives	0

Contenu non indexable	Pages
Nous n'avons détecté aucun problème de contenu non indexable sur votre site.	

Obtenir de l'aide :
Erreurs d'exploration
Codes d'état HTTP
Création de pages d'erreur 404 utiles
Bloquer ou supprimer des pages à l'aide d'un fichier robots.txt
Erreurs de type "soft 404"
Centre d'aide
Rechercher dans

Dernière mise à jour le 10 oct. 2011

La duplication de contenus :

C'est la bête noire du référencier lors de l'audit, car beaucoup de sites sont mal conçus. Dans ces cas de figure, le problème de la duplication de contenu est important et nuisible.

Avant tout chose, il faut comprendre qu'un contenu dupliqué interne est moins grave qu'un contenu dupliqué externe, pour la simple et bonne raison qu'un site ne peut-être pénalisé par le premier, tandis qu'il peut subir un malus si le même contenu est présent sur d'autres sites.

En revanche, le duplicate content interne est nuisible pour plusieurs raisons :

- Il freine l'indexation. En trouvant plusieurs fois un même contenu, Google va avoir tendance à trouver moins intéressant de continuer l'indexation du site.
- Il provoque une cannibalisation des mots clés. Sur certaines requêtes, le moteur de recherche ne saura quel contenu choisir, et il va donc parfois choisir une page moins pertinente.
- Les outils donnés dans le point précédent vont permettre de détecter cette duplication, soit directement dans le centre webmaster de Google, soit en analysant les doublons de balises title, de balises descriptions ou de balises h1 dans les deux autres logiciels.

D'ailleurs, vérifiez aussi la présence de la balise "Canonical" qui peut aider à réduire ce problème en indiquant aux moteurs de recherche l'URL réelle d'une page pour réduire ces problèmes de duplication. La balise canonical a la forme suivante :

`<link rel='canonical' href='URL réelle du contenu' />`

Lisez cet article pour plus d'informations sur l'implantation d'une telle balise :
<http://googlewebmastercentral.blogspot.com/2009/02/specify-your-canonical.html>

La page plan de site :

C'est un élément qui n'est pas obligatoire pour le référencement naturel. Mais c'est un avantage réel pour tout ce qui concerne l'indexation et le maillage interne, et qui fait donc partie de l'audit.

C'est une page accessible dès l'accueil du site et qui va regrouper l'intégralité des contenus d'un site. Le concept est simple : donner la possibilité à Google d'atteindre n'importe quelle page en moins de 2 clics, tout en lui permettant de n'oublier aucun contenu.

C'est un excellent complément au fichier Sitemap XML pour l'indexation des pages, et il faudra donc vérifier sa présence sur l'ensemble des pages du site lors de votre audit SEO.

Le fichier sitemap :

Le fichier sitemap est complémentaire de la page Plan de site. Ce fichier regroupe l'intégralité des contenus du site avec leur importance et la dernière date de publication, ce qui va là aussi permettre l'indexation de tous les contenus. Quand on analyse le code source d'un site, il faut donc :

- vérifier la présence de ce fichier.
- vérifier si celui-ci est soumis dans le centre webmaster de Google.
- vérifier si le format de celui-ci est conforme.

Votre fichier sitemap doit en effet ressembler à ce code :

```
<url>
<loc>http://www.seomix.fr/referencement/naturel/seo-contenu-pour-se-referencer/</loc>
<lastmod>2010-11-18T10:21:06+00:00</lastmod>
<changefreq>weekly</changefreq>
<priority>0.6</priority>
</url>
```

Voici quelques explications :

- **loc** : l'url du contenu.
- **lastmod** : la date, au format suivant année-mois-jour-heure.
- **changefreq** : la fréquence de mise à jour du contenu.
- **priority** : l'importance du contenu, de 0.1 à 1.0.

Le fichier robots.txt :

Lors de l'audit SEO, la non-présence de ce fichier situé à la racine du site n'est pas forcément préjudiciable, mais peut parfois avoir son importance.

Le fichier robots.txt indique en effet aux moteurs de recherche les sections qu'ils ne doivent pas indexer. Si c'est le cas du site que vous auditez, il faut que ce fichier soit présent. C'est notamment le cas des sections privées (pour les abonnés par exemple) ou encore pour les menus d'administrations.

Voici à quoi doit ressembler ce fichier :

```
User-agent: *
Disallow: /url-a-ne-pas-indexer
User-agent: Googlebot-Image
Allow: /
# BEGIN XML-SITEMAP-PLUGIN
Sitemap: url-du-sitemap
# END XML-SITEMAP-PLUGIN
```

Voici également quelques explications :

- **User-agent** : cette ligne définit à quel robot les règles qui suivent vont s'appliquer. L'étoile signifie "pour tous".
- **Allow** ou **Disallow** : cette ligne indique quelles URL il faut indexer ou ne pas indexer.
- **Sitemap** : cette ligne permet d'indiquer dans le fichier robots.txt l'endroit où se trouve le fichier sitemap.

Le site suivant explique en détail le contenu de ce site :

<http://www.robotstxt.org/robotstxt.html>

La hiérarchisation de l'information :

L'analyse d'un site doit également regarder si les pages sont bien conçues au niveau de la hiérarchisation de l'information. Nous entendons par là la décomposition d'une page en titres de niveaux H1 à H6. Ces balises vont hiérarchiser l'information pour le moteur de recherche (ce qui facilite l'indexation) et également pour le visiteur puisque ces éléments vont découper visuellement le contenu.

Il faut que ces balises soient présentes sur toutes les pages du site, mais surtout qu'elles soient placées logiquement. Il faut en effet que le site audité respecte plusieurs principes de base :

- Le H1 est le titre de la page.
- Sur la page d'accueil, c'est le nom du site.
- Sur une fiche produit, c'est le nom du produit.
- Sur un article, c'est le titre de ce dernier.
- ...

Les balises h4 à H6 sont en général inutiles, et la balise H3 à une importance faible. Une balise de niveau ne peut être seule, excepté pour le titre H1. Autrement dit, on ne peut avoir une balise H2 seule ou une balise H3 seule, car cela n'a pas de sens sémantique. Soit on a deux balises H2, soit on n'en a pas.

Une balise de niveau doit toujours dépendre du niveau supérieur. Autrement dit, on ne peut sauter un niveau, comme par exemple une page ayant un H1 et plusieurs H3 mais aucun H2. Cette hiérarchisation est valable sur tous les types de pages, que ce soit les contenus, l'accueil, les catégories, la page de contact, ... En général, l'audit pointe presque à chaque fois cet aspect de la conception d'un site, car certaines pages n'ont aucun titre H1, tandis que d'autres ont des H2 sur des éléments redondants, par exemple un h2 sur un titre « Laisser un commentaire », qui est pourtant un élément présent sur toutes les pages du site.

Les attributs alt :

Parlons des images. Un site bien conçu doit avoir renseigné tous les attributs alternatifs des images qui s'affichent (la fameuse balise Alt). C'est elle qui va permettre à Google de savoir de quoi parle un visuel afin de le référencer et de donner plus de poids au reste de la page. L'outil Xenu Link Sleuth vous permettra d'analyser ce point. Quand votre analyse sera faite, vous pourrez exporter sous Excel les données, puis vous pourrez filtrer les urls par type (uniquement les images) afin de trouver toutes les images n'ayant pas de description associée.

Attention, vous devrez aussi mettre en avant les fausses images qui sont utilisées pour la charte graphique du site. Nous parlons cette fois-ci des images de mise en page comme les puces, les pictogrammes ou certains boutons. Ceux-ci n'ont pas lieu d'être indexés et référencés par les différents moteurs. L'audit devra donc les mettre en avant pour une meilleure intégration en CSS et non pas sous forme de balises images.

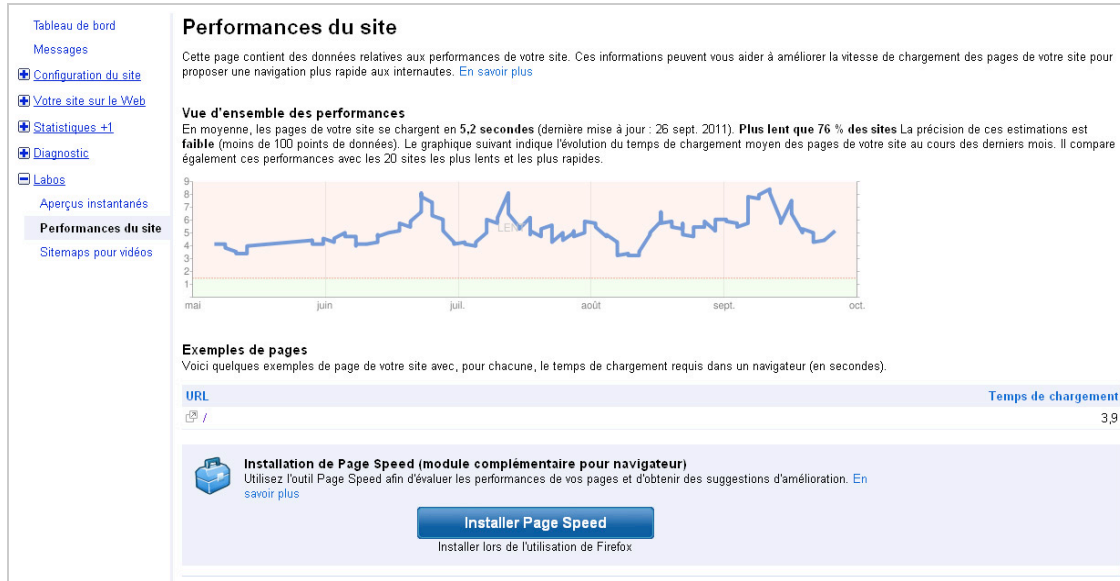
La vitesse :

Le temps de chargement d'un site est primordial pour deux raisons lors de l'audit :

- la vitesse a de plus en plus de poids sur le référencement naturel (même s'il est encore faible).
- la vitesse a un impact décisif sur les visiteurs : cela ne sert à rien d'attirer du trafic s'il repart aussi vite qu'il est venu.

Cet aspect de l'audit pourrait faire l'objet lui-même d'un article complet tellement il existe de points différents sur lesquels on peut jouer pour augmenter la vitesse et réduire le temps de chargement (voir l'article d'acril 2010 : <http://recherche-referencement.abondance.com/2010/04/comment-ameliorer-le-temps-de.html>). Le mieux, c'est encore de faire appel à des outils qui vont lister les problèmes et les solutions correspondantes :

- Le centre webmaster de Google (une fois de plus).



- Google Analytics avec certains scripts :

<http://www.seomix.fr/webmarketing/webanalytics/temps-chargement-analytics/>

- GtMetrix, un excellent outil pour optimiser son site : <http://gtmetrix.com/>

- WebpageTest, un autre outil excellent : <http://www.webpagetest.org/>

- Les outils Google sur la vitesse : <http://code.google.com/intl/fr/speed/tools.html>

Les freins à l'indexation :

Ce dernier élément à prendre en compte est également le plus vaste et le plus préjudiciable : il s'agit de tous les obstacles au référencement.

Ils sont malheureusement trop nombreux et parfois trop spécifiques pour être tous cités, mais certains éléments doivent facilement vous alerter :

- Le format flash, qui n'est pas compris par les moteurs de recherche.
- Les liens nofollow, qui deviennent de plus en plus inutiles car ils font « disparaître » de la popularité de vos autres pages. Autant ne pas mettre le lien ou le mettre sous une autre forme comme un bouton javascript.
- Le cloaking : c'est-à-dire le fait de présenter un autre contenu en fonction du visiteur et/ou du moteur de recherche.
- Une page sans liens internes qui pointent vers elle.

Xenu Link Sleuth vous indiquera là aussi un grand nombre de ces freins, comme par exemple les erreurs 404, les redirections, ...

URL et maillage interne

Le dernier point de l'audit SEO va quant à lui se consacrer sur la présence ou non d'une optimisation pour le SEO.

La mise en avant des contenus :

Avec l'analyse initiale des mots clés, le référenceur sait quels contenus sont prioritaires (soit en termes de chiffre d'affaires, soit en termes de volume de trafic). Il doit donc savoir quels pages il doit mettre en avant.

Lors de l'audit, il faut donc vérifier que ces contenus sont mis en avant à plusieurs endroits :

- Le slider, s'il y en a un.
- Le footer (à condition de ne pas y placer trop de liens).
- Dans d'éventuels blocs "notre sélection de XXX", "nos coups de cœur" ou "le meilleur XXX".
- Dans les autres contenus traitant du même thème, et qui doivent donc faire un lien vers la page "principale" liée au mot clé.

Les articles et produits relatifs :

Ces modules permettent d'associer de manière automatique ou manuelle des contenus similaires à la page que l'internaute visite. Cela permet donc de lier facilement des contenus ayant une thématique proche, leur donnant ainsi plus de poids pour les moteurs de recherche et en permettant aux internautes de continuer plus facilement et logiquement leur visite.

La présence d'un tel module est obligatoire pour avoir un effet de synergie entre les contenus, surtout quand le site publie régulièrement de nouveaux articles. De manière manuelle, on s'assure d'avoir un maillage le plus pertinent possible. De manière automatique, on s'assure que les anciens contenus restent suffisamment mis en avant à chaque nouvelle publication. L'audit doit donc impérativement vérifier qu'une telle fonctionnalité existe !

Si vous utilisez WordPress, le meilleur module d'articles relatifs est YARPP :
<http://www.seomix.fr/wordpress/plugins/articles-relatifs-yarpp/>

Les mots clés :

Si le site que vous analysez contient un module de mots clés ou "tags", il faut en analyser la pertinence. C'est à la fois un excellent outil et à la fois un « excellent » frein pour le référencement.

Sur le principe, cela permet d'associer des contenus proches en leur donnant le même mot clé. Cela a pour finalité de créer une page dédiée où l'on retrouve tous les contenus ayant cette expression associée (comme le ferait une catégorie). C'est donc utile pour se référencer tout comme cela peut provoquer une duplication de contenu et un mauvais maillage interne.

Un bon système de mots clés doit prendre en compte différentes règles :

- Chaque contenu ne doit pas avoir trop de mot clés associés : 2 ou 3 maximums par article.
- Chaque mot clé ne doit pas exister en tant que catégorie. Inutile donc d'associer un mot clé "référencement naturel" si la catégorie "référencement naturel" existe déjà.
- Chaque mot clé a un sens réel s'il est sorti de son contexte. Il vaut mieux un mot clé "Conseil en référencement" que "Conseil".
- Chaque mot clé a de préférence au moins 2 ou 3 articles associés.

Les pages de mots clés ont de préférence un affichage différent de celui des catégories. Par exemple, les listes d'articles dans une catégorie pourraient utiliser la balise méta description tandis que les listes d'articles pour un mot clé pourraient reprendre les 200 premiers caractères de chaque publication.

Si vos mots clés respectent ces règles, c'est que l'outil est déjà bien implanté.

Votre audit est terminé. Il ne reste plus qu'à appliquer point par point les recommandations qui en découlent pour optimiser le site, créer des contenus et ajouter des liens. Bon courage !

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

Réagissez à cet article sur le blog des abonnés d'Abondance :
<http://blog-abonnes.abondance.com/2011/10/audit-seo-que-doit-il-contenir.html>