**Audit de site Internet complet**[**http://noviseo.fr/2013/04/audit-seo-prerequis/**](http://noviseo.fr/2013/04/audit-seo-prerequis/)

1ère partie : Audit SEO – Avant de commencer  
12 avril 2013

Vous l'attendiez tous, le voilà enfin : le guide ultime pour réaliser un audit SEO de A à Z. Un dossier très difficile à réaliser pour la simple et bonne raison que je n'effectue jamais un audit de la même façon. On va donc commencer par poser les bases avant d'entrer dans le vif du sujet.

Un audit se réalise **en fonction des objectifs du site**, et c'est la raison qui fait qu'un audit pour un site ne ressemblera pas à celui effectué pour un autre site. La première chose à prendre en compte est donc de clairement définir ce que veut le client pour savoir quelle direction vous devez prendre pour réaliser votre audit. Ceci étant dit, c'est le moment de commencer.

**Check-up :**Avant toute chose, essayez de récupérer le maximum d'informations possibles, dont les principales sont :

* l'accès à l'outil d'Analytics (Omniture, Google Analytics, etc.)
* l'accès à Google Webmaster Tools
* l'accès à Bing Webmaster Tools
* la récupération des logs
* l'accès au serveur, ou au minimum une copie du fichier .htaccess

Un outil d'analytics bien configuré vous donnera également le taux de conversion, les pages qui convertissent le plus, les objectifs du site, etc. Malheureusement, très souvent l'outil n'est pas configuré ni utilisé à 100 % et il faudra demander d'avoir ces statistiques d'une autre façon.

Il arrive également souvent qu'aucun compte Webmaster Tools n'existe, ça ne facilite pas forcément le travail mais comme rien n'est jamais trop tard, c'est le moment idéal pour les créer ;-)

Pourquoi ces informations sont-elles importantes ? Tout simplement parce qu'elles permettront :

* de connaitre l'activité des moteurs sur le site
* de connaitre le comportement des internautes sur le site
* de déjà repérer d'éventuels facteurs bloquants

**Les outils nécessaires :**De votre côté, il va falloir déployer un certain arsenal.

1. **Un crawler :**Si vous ne deviez utiliser qu'un seul outil, ce serait celui-ci. Au minimum, le crawler doit pouvoir mettre en exergue :

* l'URL
* le statut de l'URL (ou code de réponse, vous l'appelez comme vous voulez)
* le niveau de profondeur
* le type de fichier

Il s'agit là du minimum vital. Idéalement, le crawler pourra également ressortir des informations comme :

* les directives (meta robots, X-Robots-Tag)
* le Title des pages
* le poids des fichiers
* la présence ou non d'une canonical
* le nombre de liens internes et externes
* le contenu des h1 et h2

Bien sûr, en fonction des objectifs, d'autres informations peuvent être utiles/nécessaires. À vous de juger ce qu'il faut récupérer ou non ;-)

1. **Un analyseur de backlinks**Si les Webmaster Tools nous donnent déjà des informations sur les liens entrants, elles sont malheureusement loin d'être complètes. C'est donc pour ça qu'un outil d'analyse de liens nous sera utile. En effet, nous aurons certainement besoin de connaitre des informations comme :

* l'ancre des liens entrants
* le nombre total de liens et de domaines
* la qualité des domaines entrants

1. **Un outil de visibilité**Forcément imprécis mais pourtant indispensable, un outil de visibilité permettra d'avoir un aperçu de la présence du site sur les moteurs, de la présence des concurrents mais aussi d'évaluer le potentiel d'évolution du site en terme de visibilité, donc en terme de trafic, donc en terme de chiffre d'affaires.
2. **Un outil de positionnement de mots-clés**Je ne vous cache pas que celui-ci est vraiment peu important à mes yeux. Je vais quand même l'évoquer dans ce dossier et partager avec vous mon avis et ses utilités possibles.
3. **Un outil de performances**Pas forcément nécessaire puisqu'une analyse complète permet logiquement de définir sur quels leviers appuyer pour améliorer les performances du site, mais il faut avouer qu'un Page Speed Insight ou un Yslow facilite bien la vie ;-)
4. **Excel, Word et Powerpoint**Ok je vous l'accorde, ces trois-là coulent de source mais ils seront tout de même indispensables. Ceux-là ou leur équivalent, bien sûr ;-)

Vous avez maintenant sous la main tout ce qu'il vous faut pour vous lancer dans un bel audit SEO, il ne vous manque plus que la méthodologie. Une fois n'est pas coutume, je vous donne rendez-vous dès demain pour la suite de ce dossier. Petit teasing : on va commencer par l'analyse technique, mais je parie que vous vous en doutiez déjà. En attendant d'être à demain, vous pouvez jeter un coup d'oeil [à l'audit du site de l'Élysée](http://noviseo.fr/2013/01/audit-referencement-du-site-de-lelysee/)

|  |  |
| --- | --- |
| **Crawler :**  **Analyseurs backling** | Xenu  <http://www.screamingfrog.co.uk/seo-spider/> Integrity for Mac : <http://peacockmedia.co.uk/integrity/> |

2ème partie : Audit SEO – Technique  
13 avril 2013

On continue le dossier sur la réalisation d'un audit de référencement. Vous êtes maintenant équipé : vous avez les accès nécessaires et les outils requis. On va s'attaquer maintenant à la partie technique de l'audit. S'il est communément admis que la partie technique ne représente que 10 % des facteurs qui font qu'un référencement est réussi (les chiffres donnent 10 % technique, 40 % contenu et 50 % popularité, après on en fait ce qu'on veut hein...), c'est souvent ici qu'on pourra déjà repérer de nombreux facteurs bloquants ou ralentissants qui seront rapides à corriger et qui apporteront les meilleurs résultats à court terme.

1. **I'm the crawling king snake**Avec des phrases pareilles, Jim Morrison serait certainement devenu référenceur. Le crawl est le passage obligé de tout audit SEO. L'idée est donc de crawler le site pour récupérer toutes les informations qui peuvent nous être utiles. Mais bien sûr, il va falloir crawler intelligemment :

* un **premier crawl sans paramétrage particulier**
* un **second crawl en mettant en user-agent** celui du moteur de recherche qui vous intéresse (ok, on se fait passer pour Googlebot)
* un **troisième crawl en désactivant le javascript, voire le CSS**

Ça, c'est si on veut faire les choses vraiment bien. Dans les faits, la plupart du temps on ne fera qu'un seul crawl, qui mobilisera déjà beaucoup de ressources s'il s'agit d'un gros site. Comme indiqué dans l'article précédent, celui-ci va nous permettre de récupérer différentes informations :

* les statuts HTTP des pages
* le type de fichiers
* le niveau de profondeur des pages
* la construction des URL
* le poids des fichiers
* les directives robots
* la présence de canonical
* Etc.

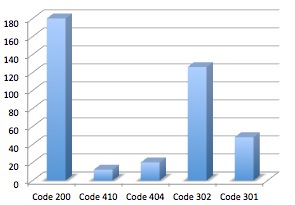
Nous allons donc commencer par nous attarder sur chacun de ces points.

1. **Les statuts http**Ce sont tout bêtement les statuts :

* 200
* 301
* 302
* 400
* 401
* 403
* 404
* 410
* 500
* etc. vous avez compris

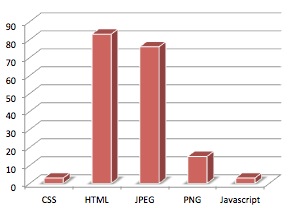
Cette première analyse va nous permettre d'avoir un petit aperçu de ce que peut penser Google du site analysé : y a-t-il beaucoup de codes d'erreurs ? Existe-t-il des 302 qui mériteraient d'être des 301 ?  Des 404 ne peuvent-elles pas être gérées autrement, en 410 ou 301 par exemple ? Etc.

Pour cette analyse là comme pour beaucoup d'autres, un graphique est souvent bien parlant.



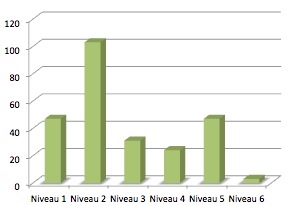
Google Webmaster Tools vous propose également un listing des erreurs dans la rubrique État de santé ==> Erreurs d'exploration.

1. **Le type de fichiers :**Information peu pertinente vous allez me dire, et bien pas tant que ça ! Nous allons voir pourquoi dès maintenant, mais aussi un peu plus tard ;-) En effet, nous allons pouvoir repérer s'il y a certaines pratiques à revoir, comme plusieurs fichiers CSS ou javascript qui mériteraient d'être regroupés en un seul fichier, des formats d'images peu adaptés, etc.

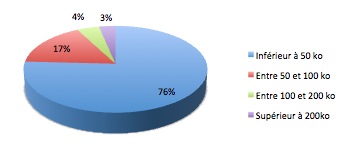


Si le site est 100 % Flash, vous aurez remarqué que vous n'avez qu'une seule URL qui est la page d'accueil. En effet le crawler ne saura pas suivre les liens, vous pouvez donc arrêter l'audit technique dès maintenant. Si seul le menu de navigation est en flash, vous pouvez continuer mais les résultats seront biaisés.

1. **Le niveau de profondeur des pages :**Comme vous le savez, plus les pages importantes sont en haut de l'arborescence, mieux c'est. Une analyse du niveau de profondeur permettra donc d'identifier des problèmes à ce niveau là : certaines pages sont-elles dans les tréfonds du site ? Combien y a-t-il de pages accessibles depuis la home ? Etc.



En plus d'avoir une idée du nombre de clics nécessaires pour arriver à la page la plus profonde, cette information nous sera très utile lorsque nous évaluerons l'arborescence du site.

1. **La construction des URL**Grâce au crawl vous avez la liste complète des URL sous la main, autant en profiter n'est-ce pas ? C'est le moment d'effectuer une petite analyse : sont-elles SEO friendly ? Contiennent-elles des paramètres ? Des ID de sessions ? Ont-elles une logique dans la construction ?
2. **Le poids des fichiers**Pour des raisons de performances, vous savez que plus les pages et fichiers sont légers, mieux c'est pour l'internaute et donc pour Google. On va donc avoir ici un aperçu des performances du site, même si le poids des fichiers ne fait pas tout. 
3. **Les directives Robots :**Comme vous avez un bon crawler, celui-ci peut interpréter le code des pages mais également les en-têtes HTTP (et oui, sinon vous n'auriez pas pu avoir les statuts HTTP des pages). C'est le moment de jeter un œil à ce que le site indique aux robots.  
     
   Globalement, ça se passe au niveau des balises **meta robots** (dans le code) et/ou des **X-robots-tag** (dans les en-têtes HTTP). On cherche ici à identifier des facteurs réellement bloquants comme une demande de non-indexation ou un non suivi des liens (typiquement les directives noindex et nofollow).  
     
   Mais les informations ne s'arrêtent pas là, on trouve par exemple des directives concernant la mise en cache des fichiers dans les en-têtes HTTP. C'est une vraie mine d'or, apprenez à les interpréter ;-)
4. **Les canonical**J'aurais pu les mettre dans le paragraphe précédent mais je me suis dit soyons fou, parlons-en à part. Vous savez tous que les balises canonical indiquent aux robots que la page sur laquelle ils sont actuellement est la copie d'une autre page accessible depuis une URL différente.  
     
   Leur but étant d'éviter le contenu dupliqué, ou tout du moins de signaler à Google quelle page il doit mettre en avant lorsque plusieurs pages sont identiques, il est nécessaire d'y prêter attention. Les pages dupliquées ont-elles une URL canonique ? Cette URL existe-t-elle ? Est-elle bien accessible ?
5. **L'arborescence du site**Vous avez déjà eu un aperçu en regardant la profondeur des pages, maintenant il ne reste plus qu'à vérifier que les pages importantes se retrouvent bien en haut de l'arborescence, qu'elles sont liées depuis d'autres pages importantes, que les pages qui convertissent (ou qui sont censées convertir) sont facilement accessibles, etc.  
     
   Si dans l'arborescence, les pages sont bien liées verticalement, le sont-elles également horizontalement ? Veillez à saisir toutes les opportunités concernant le maillage interne ;-)  
     
   D'ailleurs, qu'en est-il de la pagination ? Est-elle correctement mise en place ? Est-elle optimale ? Favorise-t-elle une réduction du niveau de profondeur de certaines pages ? Les balises rel="prev" et rel="next" sont-elles bien en place ?
6. **Le Robots.txt**On va très vite passer sur ce point, l'important est simplement de vérifier, si le fichier existe, qu'il n'y a pas de restriction sur des dossiers ou des pages importants Au besoin, vous pouvez relire [mon article sur le robots.txt](http://noviseo.fr/2013/01/larticle-ultime-sur-le-robots-txt/).  
     
   Après avoir relu cet article, vous allez me dire "oui mais dedans tu dis que le robots.txt n'empêche pas l'indexation, donc même si elles sont bloquées elles pourront apparaître sur Google, non ?" Oui, mais comme Google ne les aura pas crawlées, il ne connaitra pas son contenu (balise Title, description, textes, etc.) et ne les proposera pas dans les premiers résultats de recherches. C'est aussi simple que ça ;-)  
   Google Webmaster Tools vous permettra de tester le robots.txt
7. **Le Sitemap.xml**Je sais que beaucoup de référenceurs vous diront que le sitemap.xml n'est pas nécessaire. Pour ma part je suis un fervent défenseur de la présence de ce fichier, particulièrement lors d'un audit. Il permet notamment de répondre à ces deux questions :  
     
    - Une page existe-t-elle dans le sitemap et n'est pas présente dans mon crawl ? Si la réponse est oui, alors cette page est orpheline. C'est mal ;-)  
    - Une page est-elle présente dans mon crawl mais pas dans le sitemap ? Si la réponse est oui, ça signifie que que le sitemap n'est pas à jour.  
     
   Concernant la deuxième question, est-ce grave si le sitemap n'est pas à jour ? Foncièrement non, mais Google Webmaster Tools vous donnera des informations et des avertissements : existe-t-il une page listée dans le sitemap et qui est en même temps bloquée au crawl par le robots.txt ?
8. **Le fichier .htaccess :**S'il n'est pas nécessaire d'y avoir accès (et de toutes façons, il y a des risques pour que vous n'ayez accès que très rarement à ce fichier), le .htaccess peut tout de même se révéler d'une aide non négligeable. Celui-ci vous informera sur les directives données aux robots (comme vu plus haut, mais tout est sous les yeux), les informations sur la mise en cache (idem), sur la réécriture d'URL, etc.  
     
   Pour en savoir un peu plus sur le fichier htaccess, je vous recommande de lire [cet article](http://openweb.eu.org/articles/le-fichier-htaccess).
9. **Les performances du site :**Si le poids des fichiers et le .htaccess nous ont déjà renseignés sur ce point là, il peut être intéressant d'aller plus loin. Pour ça, plusieurs outils sont à notre disposition, et certains nous donnent même des conseils pour améliorer les performances :

* Vitesse du site dans Google Analytics
* [Pingdom Tools](http://tools.pingdom.com/)
* [Page Speed Insight](https://developers.google.com/speed/pagespeed/insights)
* [Gtmetrix](http://gtmetrix.com/)
* L'extension Yslow pour Firebug  
  Mis à part Google Analytics qui offre une vision générale, ces outils ne permettent une analyse que pour une page en particulier, ce n'est donc pas forcément viable pour un site entier, mais une fois que vous avez repéré des typologies de pages, il y a des chances pour que les autres pages du même type connaissent les mêmes problèmes que celle analysée. Vous me suivez ?
* <http://bertiaux.fr/2011/12/les-7-meilleurs-outils-pour-tester-la-rapidite-de-votre-site-web/>
* <http://www.revue-hebergement-web.com/articles/Webdesign_et_programmation/Comment_tester_la_vitesse_de_votre_site_gratuitement/>
* <http://www.webpagetest.org/result/140327_30_FX3/>

**Le reste :**Vous en avez déjà entendu parler ailleurs donc je ne vais pas en rajouter une couche pour garder l'article digeste, mais il est important (ou parfois juste intéressant) de prendre en compte ces choses là :

* l'âge du domaine (avec n'importe quel service de WhoIs)
* le voisinage du domaine (avec [YouGetSignal](http://www.yougetsignal.com/" \t "_blank)), sa localisation géographique
* le type d'hébergement (dédié ou mutualisé)
* le nombre de liens par pages (pas trop de liens internes ou externes, tout de même quelques liens externes)
* la présence de liens en nofollow (à éviter s'ils pointent vers des pages importantes, à méditer s'ils pointent vers des pages peu importantes)
* l'éventuelle duplication du domaine (le www est-il bien redirigé vers la version sans www, ou vice-versa ?)
* la présence d'éléments cachés en CSS (à éviter bien évidemment !)

On arrive à la fin de la partie technique de l'audit SEO, ça avance bien non ? Mais ne croyez pas que c'est terminé, il nous reste encore pas mal de choses à voir ;-) Rendez-vous dès demain pour la suite de ce dossier avec l'audit de contenu. Vous avez hâte j'espère ?

**Commentaire :** Utilise une première fois Xenu pour récupérer ta liste d’URL que tu soumets ensuite à Screaming Frog pour plus de variables. [Voir la méthode](http://www.visibilite-referencement.fr/blog/excel-screaming-frog-seo).  
Comment utiliser Screaming Frog gratuitement :D <http://noviseo.fr/2012/12/utiliser-gratuitement-screaming-frog-seo-spider/>

3ème partie : Audit SEO – Contenu  
14 avril 2013

Ça y est, vous avez les bases en ce qui concerne la partie technique d'un audit de référencement, maintenant il va falloir s'attaquer à la partie contenu. Vous avez trouvé le travail sur l'aspect technique fastidieux ? Réjouissez-vous, ici ça va être encore pire !

Pire parce qu'il est quasi impossible d'automatiser une analyse de contenu car souvent les ressources vous feront défaut ; pire parce que cette même analyse est purement subjective, contrairement à tous les autres aspects de l'audit SEO ; pire parce qu'il va falloir presque tout faire à la main ; pire parce que plus il y a de pages, plus vous aller y passer du temps. Vous êtes quand même prêt à continuer ? C'est parti !

**Repérage du contenu dupliqué**

Si vous avez construit vous-même votre propre crawler et que vous avez les ressources nécessaires, alors ce travail sera facilité. Idéalement, votre crawler a repéré la zone où se trouve le contenu effectif des pages (ie le contenu sans l'en-tête, les sidebars et le footer), il ne vous restera donc plus qu'à filtrer les données récupérées et vérifier la présence de doublons.

Dans les faits, ceci n'arrive quasiment jamais, vous imaginez les ressources nécessaires pour un site de 100 000 pages ? Donc on va jouer un peu différemment. S'il n'est pas aussi puissant, notre crawler a tout de même récupéré les balises Title, Meta description, H1 et H2. Screaming Frog le fait, d'ailleurs ;-) On va pouvoir commencer par filtrer ces données pour repérer les doublons, au moins pour la balise Title. Si on trouve des Title doubles, c'est mal, et si après vérification on voit que les pages ont le même contenu, c'est encore pire !

Google Webmaster Tools nous indique les balises Title et Description dupliquées dans la catégrie Optimisations ==> Améliorations HTML.

Une vérification "à la scout" sera toujours intéressante : vous sélectionnez quelques pages, vous copiez une ligne du contenu au hasard et vous allez la coller dans Google sous la forme suivante :

"ligne que vous venez de copier" site:site-audité.fr

Vous saurez ainsi si Google connait plusieurs pages de votre site ayant le même contenu. Évidemment vous n'allez pas le faire sur toutes les pages du site mais simplement sur une sélection de différentes typologies de pages.

Concernant le contenu dupliqué sur d'autre sites, c'est un peu la même façon de faire, la partie "site:" en moins. Plus simplement, vous pouvez également demander au client comment sont gérés ses contenus : les fiches produits sont-elles des copies des fiches du fabricant ? Il saura vous répondre, et devra le faire. Hey ! Vous bossez **avec** lui, pas uniquement **pour** lui !

**La qualité du contenu**

Sujet brûlant tant il dépend de la stratégie et donc des objectifs, la qualité est tout de même importante. Même si globalement ça devrait aller assez vite sur ce point, on commence par ce qui va plaire à Mme Duchez (votre prof de français de troisième), ses dissertations et son Bescherelle :

* vérification de l'orthographe
* vérification de la grammaire
* vérification de la valeur ajoutée du texte

Une fois ces vérifications effectuées, il va falloir comprendre si le champ lexical est optimal : vise-t-il les meilleures expressions ? Les expressions visées sont-elles réalistes (ie : le positionnement sur ces expressions est-il réalisable, en prenant en compte le background du site) ? Sont-elles génératrices de trafic ? Et surtout, le plus important, sont-elles adaptées à la cible ?

**Keyword stuffing et "keyword missing"**

Le keyword stuffing est la répétition d'un mot clé, mais de façon bien trop importante. C'est une pratique que les moteurs repèrent et sanctionnent à plus ou moins forte échelle, c'est donc une pratique à éviter. Il est donc important de se poser la question si le site est sujet ou non au keyword stuffing :

* plusieurs occurrences de l'expression dans le Title
* utilisation de l'expression dans l'URL
* utilisation de l'expression dans le H1 et chacun des H2
* répétition trop importante de l'expression dans le texte
* mise en place de l'expression dans l'attribut alternatif des images
* etc.

À l'inverse on retrouve ce qu'on pourrait appeler le "keyword missing". Pour faire simple, c'est le fait de vouloir se placer sur une expression sans l'utiliser dans la page, ou en l'utilisant trop peu. Au final, on reprend à peu près les mêmes points que précédemment mais on recherche l'effet inverse :

* l'expression n'est pas utilisée dans le Title
* l'expression (ou une de ses variantes) n'est pas utilisée dans le H1
* l'expression n'est jamais citée dans les textes
* etc.

Mais comment repérer tout ça ? Quel est le taux idéal d'occurrences d'un mot clé dans une page ? Honnêtement, je n'en ai aucune idée, je m'en contrefous et je ne suis même pas sûr qu'il y ait un "taux idéal" quelque part dans l'algorithme de Google. La façon la plus simple de repérer un de ces deux extrêmes est de lire le texte. Si ce dernier utilise une expression si souvent qu'il en devient indigeste, alors c'est du keyword stuffing. Si vous n'arrivez pas à résumer le sujet d'une page en deux ou trois mots, ou que l'expression qui semblerait la plus pertinente n'apparait pas, alors c'est du keyword missing.

**La taille des textes**

Là encore on est vraiment dans le "très subjectif" et ici aussi, je serais tout simplement incapable de définir un volume de texte optimal. Lorsque vous auditez les textes d'un site, le plus pertinent est de se mettre dans la peau d'un internaute : se sent-il lésé concernant l'information partagée ou au contraire se sent-il étouffé par une quantité de textes si énorme qu'elle en devient imbuvable ? Dans ce deuxième cas, le texte ne mérite-t-il pas d'être divisé en plusieurs pages ?

Google veut proposer quelque chose qui plait à l'internaute. Si à la première vision un texte ne vous plait pas, il en sera autant pour tous les autres internautes.

**Les autres types de contenu**

Si le texte est le seul type de contenu que Google sait lire, c'est loin d'être le seul contenu disponible sur internet. Les autres contenus sont pourtant bien susceptibles d'être présents sur le site audité. Quid des images ? L'attribut alternatif est-il renseigné ? Le nom du fichier est-il descriptif ? Un texte explicatif les accompagne-t-il ? Même chose pour les vidéos, les sons, les présentations, les formulaires, etc. : sont-ils accompagnés d'un environnement susceptible de les définir ?



Nous venons de terminer un bon tour du propriétaire après avoir réalisé la partie technique et la partie contenu de l'audit de référencement. Vous pensez en avoir fini ? Non, bien sûr que non ! Ça tombe alors très bien puisque vous pourrez revenir demain pour vous attaquer à la partie statistique de l'audit ;-)

Commentaires :

1. Je rajouterai qu’il faut vérifier qu’il n’y ai pas trop de publicité sur les pages web du site.  
   De plus, si on découvre du contenu dupliqué, est ce que le webmaster est au courant de ce contenu duppliqué en mettant en place une canonical par exemple.

Comme tu dis, il existe d’autres type de contenus que le contenu textuel. Mais est ce qu’un site sans images, avec uniquement du contenu textuel est une bonne chose ? Ne vaut-il pas mieux inserer quelques images au sein d’un site web ? Tout ça dans le but de paraitre plus naturel pour Google. Après je ne pense pas qu’on puisse le savoir, mais bon…

1. [**Hamza**](http://www.tousgeeks.com/author-rank-ameliorer-referencement/) *says:*

[14 avril 2013 at 12:16](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18591)

Bonjour Sylvain et merci pour ce dossier très complet, ça m’aide tellement dans mon expérience de seo/blogueur

Je me posais depuis un bon la question : Sur un site, la page d’accueil manque de contenu, est-ce que c’est le référenceur qui s’en charge ou bien le client ? Car des fois, y a un site avec un sujet que je connais pas, qui manque de contenu…

1. [**RaphSEo**](http://www.visibilite-referencement.fr/audit-referencement.html) *says:*

[14 avril 2013 at 13:30](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18592)

Hello Sylvain,très sympa ta série de billets et conseils sur la technique et méthodo d’un audit.  
Sans vouloir trop en dire jettez un oeil à cette page <http://www.visibilite-referencement.fr/blog/bookmarklet-seo> (oui je sais elle est de moi mais ce n’est pas pour cette raison. )  
Regardez le bookmarklet sur la duplication de title. Ce n’est qu’un exemple, creusez vous un peu la tête ;)  
++  
Raph

1. [**Lionel**](http://www.les-options.com/avis-topoption/) *says:*

[14 avril 2013 at 14:36](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18594)

Je ne pense pas qu’il y ait de densité idéale, et d’ailleurs Google l’a affirmé (même si ça ne prouve rien). Ce qui est sur c’est que si le mot n’est pas présent au moins une fois dans le texte c’est impossible à ranker, même avec des BL contenant le mot clé en question.  
Une bonne solution pour le keyword stuffing est de remplacer par des synonymes.

1. [**lionel**](http://www.agence-web-vietnam.com/audit-referencement.php) *says:*

[14 avril 2013 at 16:45](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18595)

Bonjour Sylvain,

Google constitue un robot qui juge un texte suivant des critères assez basiques tels que : la duplication de contenu, la répétition des mots-clés, la répétition des occurrences, le contexte, la structure, la titraille…

Ceci étant, je pense aussi qu’un auditeur doit être en mesure de conseiller sur un contenu textuel optimisé à la fois pour le référencement mais également sur sa capacité à remplir un objectif de conversion… un équilibre à gérer et pas toujours évident !

1. [**rocketSeo**](http://www.aurele-gala.com/) *says:*

[14 avril 2013 at 17:12](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18596)

Super intéressant, personnellement je ne connaissais pas la notion de **keyword missing**, enfin je veux dire la façon de nommer ce constat. Par contre, il me semble qu il manque tout de même quelque chose d assez fondamentale : tout ce qui concerne le balisage (**headings**). en effet, on en parle pas dans le technique et dans le contenu…Du coup, je me demandais quand cela sera abordé. mais je suis peut être un peu impatient… ;)

J attends avec impatience votre vision de l’audit netlinking… il y a surement de bonnes idées à reprendre. En tout cas, le fait le couper vos articles en différents feuilletons est une bonne idée. C est à a la fois agréable et addictif pour l internaute.

1. **GreenArrowFrance** *says:*

[14 avril 2013 at 19:21](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18597)

Merci beaucoup pour ces conseils. Je suis très intéressé par Screaming Frog maintenant que tu as décrit ses fonctionnalités. Je vais le télécharger pour tester, il m’a l’air d’être un outil parfait pour optimiser le contenu du site en particulier pour le duplicate content.

1. **Sylvain** *says:*

[14 avril 2013 at 21:52](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18599)

@Hamza : tout dépend du prix de la presta, de ce qui y est inclus, etc. Après, le rôle du SEO est quand même de se rencarder sur l’activité de son client ;-) Si jamais tu ne te sens pas chaud, tu peux demander au client de simplement lister des points qui lui paraissent importants et que tu te charges de mettre en forme.

@Raph : dis donc, c’est fini de citer tes articles comme source d’infos supplémentaires dans mes commentaires, oui ? Un lecteur a déjà cité ton article sur Screaming Frog dans l’article sur la technique :D Mais oui en effet, c’est également une possibilité supplémentaire ;-)

@rocketseo : ah oui tiens, c’est un oubli de ma part, ça devait effectivement aller dans cet article. Je suis bon pour le corriger maintenant ;-) Pour la notion de keyword missing, c’est pas illogique que tu n’en aies pas entendu parler avant puisque l’expression m’est venue en rédigeant !

1. [**James**](http://www.jeux2taxi.net/wild-wild-taxi.html) *says:*

[15 avril 2013 at 05:21](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18600)

La bonne taille d’un texte, c’est une question que je me suis toujours posé, mais c’est vrai que je n’avais pas abordé le sujet de cette manière, en me mettant à la place de l’internaute, ce qui en plus très logique. Donc désormais je ferai en sorte que les textes soient informatif juste ce qu’il le faut avec une densité moyenne de mots clés, sans forcément me dire qu’il n’est pas assez long ou parfois trop.

1. **RaphSEO** *says:*

[15 avril 2013 at 07:42](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18601)

Sylvain, loin de moi l’idée de vouloir te spammer :)  
Mais c’est de ta faute, tu n’as qu’à écrire des choses nulles et stupides et je ne viendrais pas chez toi :D  
Plus sérieusement c’est ta commande « ligne que vous venez de copier » site:site-audité.fr qui m’y a fait penser.  
Vu que je l’ai semi-automatisée avec le bookmarklet je pensais que cela pourrait intéresser tes lecteurs.

1. **Sylvain** *says:*

[15 avril 2013 at 09:18](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18602)

@Raph : pas de soucis, tu sais ce que j’en pense : si c’est bon pour le lecteur, alors c’est bon pour moi ;-)

1. **Véronique Duong** *says:*

[15 avril 2013 at 09:37](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18603)

Merci pour la suite de ce dossier « Audit SEO » ! C’est vrai que pour l’audit contenu sémantique, il vaut mieux le retravailler le plus humainement possible pour s’assurer de la qualité.

1. **gab** *says:*

[15 avril 2013 at 14:56](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18609)

Je pense qu’il nous est impossible de connaitre la densité de mots clés acceptée par Google.

[Par contre il est facilement possible de reconnaitre les spammeurs compulsifs]

1. [**Jerry**](http://bloganime.fr/post/2011/08/10/Sp%C3%A9cialiste-du-diagnostic-thermique-:-Ad%C3%A9othermie) *says:*

[16 avril 2013 at 18:04](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18634)

Super article, il y manque juste le linkbaiting non ? @rocketSeo : le keyword missing est une tactique principalement utilisée par les blogs et/ou sites soupçonnés voire pénalisés d’avoir abusé du SEO. Ca leur permet de se faire comprendre de l’internaute en travaillant sur des keywords alternatifs mais connexes à la thématique de leur moneysite.

1. [**Jerome**](http://www.lije-creative.com/) *says:*

[17 avril 2013 at 20:21](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18662)

Quand tu cherches le contenu dupliqué, ta façon de faire permet de trouver le DUST (dupliqué interne), mais pas le dupliqué externe. Le dupliqué interne est moins grave que le dupliqué externe. Mais tu as raison de dire que dans les 2 cas, c’est un problème à résoudre.

1. [**Monica**](http://www.monicamedias.com/) *says:*

[17 avril 2013 at 22:15](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18667)

Bonne technique que de feuilletonner les bons plans du SEO: une technique journalistique à la Médiapart sauf qu’il n’y a pas de scandale politico-financier dans tes billets, mais cependant une grande tendance à l’humilité. A l’heure où ou GG multiplie les variables, les filtres, les stratégies, les outils, cela devient en effet irrationnel d’affirmer une théorie parfaite.  
Ton ode au TRES SUBJECTIF est le filigrane de ce billet: ben oui, taille de texte, densité de mots-clef, qualité de contenu, tout cela a sombré dans l’abstraction en même temps me que GG s’attaque au continent immergé de la sémantique.  
Par contre dans un registre plus concret, tes armes anti-duplicate ne font pas dans la dentelle. J’ajouterai que stratégiquement se renforcer au niveau de l’Author-rank permet de passer outre le problème, mais est-ce bien rationnel là aussi?

1. **Sylvain** *says:*

[17 avril 2013 at 22:30](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18668)

@Monica : Et bien, très beau commentaire ! Je rebondis sur l’Author Rank : je ne saurais dire si c’est rationnel ou non, tout dépend de la stratégie de l’entreprise. Quid d’une entreprise qui fait rédiger ses textes ailleurs ? Je ne pense pas qu’il soit rationnel de lier un compte auteur à ces contenus, pas toujours de bonne qualité. Mais dans ce cas, pourquoi externaliser son contenu ? C’est toute une stratégie à revoir du coup =)

1. **Hug0b** *says:*

[19 avril 2013 at 09:38](http://noviseo.fr/2013/04/audit-seo-contenu/#comment-18686)

Le contenu, un des paramètres essentiel d’un site web « content is king ». Après concernant la densité des KW, c’est toujours autant le flou autour de cette donnée, y’a des pour, y’a des contres, on sait plus trop quoi penser …

**4ème partie : Audit SEO – Statistiques**

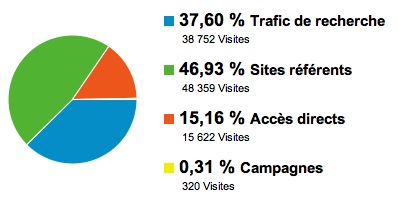
15 avril 2013

L'audit technique et l'audit de contenu sont effectués, il va falloir passer maintenant à la partie statistiques. Nous n'allons pas ici faire un rapport détaillé des statistiques du site mais simplement faire une analyse des points SEO importants.

L'objectif de cette analyse est de mettre en exergue les perspectives d'évolutions d'un point de vue SEO mais également les freins éventuels au sein du site concernant le parcours utilisateur et la conversion. Il est donc extrêmement important de se procurer les accès du site audité pour pouvoir pointer ces éléments.

**La ventilation du trafic**

Finalement au moment de l'audit le trafic total nous intéresse peu, excepté si on connait bien la thématique. Ce qui est plus intéressant par contre, c'est la ventilation de ce trafic : quels sont les points d'entrées sur le site ?



Ce qui nous intéresse ici est de connaitre la part de trafic SEO. Est-elle forte ? Est-elle faible ? Est-il possible d'augmenter cette part ? Bien sûr, si idéalement un site récupère autant de trafic SEO que de trafic direct et de trafic référent, le référenceur va toujours chercher à rendre la part de trafic en provenance des moteurs plus importante. Pourquoi ? Parce que de cette façon il montre qu'il est indispensable, donc qu'il faut le garder s'il est in-house ou qu'il faut signer à nouveau si c'est une agence ou un freelance.

Revenons à l'audit et donnons un ordre de grandeur : si la part de trafic SEO est inférieure à 25 % (sur une période globale et non à un instant T, on n'est jamais à l'abri d'un buzz), alors c'est que le site a de gros problèmes de ce point de vue là.

**Les sites référents**

Si le trafic SEO nous intéresse avant tout, il ne faut pas pour autant mettre de côté le trafic référent. Certains sites peuvent rapporter énormément de trafic, il est utile de les connaitre pour la suite (type de site, éventuel spot de linking, etc.) Ce n'est pas la chose la plus importante de l'audit, loin de là, mais il ne faut pas non plus l'oublier ;-)

**L'évolution du trafic SEO**

Cette donnée est particulièrement importante, bien que toute seule elle ne veule pas dire grand chose. Si le site est sujet à la saisonnalité, une comparaison avec l'année précédente à la même période est cruciale : le site fait-il plus, moins ou autant de trafic que l'année passée ? Les creux et pics se situent-ils au mêmes moments ou sont-ils totalement distincts ? L'évolution se reflète-t-elle sur toute la période ou simplement à un moment particulier ? Tout ceci peut permettre de mettre le doigt sur un problème épineux s'il y a de fortes disparités.



**Les landing pages**

C'est selon moi l'un des points clés de l'audit SEO : où atterrissent les internautes ? Pourquoi est-ce si important ? Tout simplement parce que ça permettra d'identifier d'éventuelles coquilles dans la priorisation des pages sur le site mais aussi de trouver des "spots internes" pour travailler le maillage. Voici les questions à se poser lorsqu'on effectue cette analyse :

* Lorsque les internautes arrivent sur le site depuis les moteurs, sur quelles pages arrivent-ils ?
* Ces pages sont-elles pertinentes si on se place du point de vue de la conversion ?
* Et pour l'utilisateur, sont-elles également pertinentes ?
* Est-ce que des pages peu importantes cannibalisent d'autres pages plus importantes en termes de visites SEO ?
* Est-il possible de placer des liens vers d'autres pages, plus sujettes à convertir ?

Jeter un oeil au taux de rebond de ces pages et au temps passé permet de mieux définir si elles sont pertinentes pour l'internaute.

**La conversion**

Un régal si tout est bien paramétré, un calvaire si ce n'est pas le cas ! Si des objectifs et un entonnoir de conversion ont été mis en place, tout est beaucoup plus simple :

* Quel chemin est utilisé pour atteindre les objectifs ?
* Quel pourcentage d'utilisateurs ayant entamé l'entonnoir de conversion va jusqu'au bout ?
* Quels sont les pages de sortie de l'entonnoir ?
* D'où proviennent les visites qui convertissent ? À quel taux ?

Le fait de répondre à toutes ces questions répondra plus à un travail d'ergonomie, mais après tout ça fait quand même partie du job ;-)

Si vous êtes plus à l'aise avec les schémas, pensez à la visualisation par flux sur Analytics.

**Recherche sur site**

Ici aussi, si le taggage de la recherche sur site a été paramétré, c'est une véritable mine d'or qui est présentée par Analytics. Celle-ci permettra de pointer du doigt ce qui intéresse les internautes et qu'ils n'ont pas trouvé via la navigation, voire même ce qu'ils n'ont pas trouvé du tout. Dans ce cas, c'est une indication précise des manquements sur le site en termes de contenu.

Les principales informations SEO fournies par les analyses ont été abordées, l'audit commence quand même à ressembler à quelque chose si on y ajoute les parties technique et contenu que nous avons vues dans les articles précédents. Mais c'est encore loin d'être terminé, il y a encore pas mal à faire. À demain !

Commentaires :

1. J’aime beaucoup discuter avec les clients du camembert de répartition de trafic. Avoir une grosse part du trafic en provenance des moteurs les réjouit la plupart du temps mais c’est aussi une grosse faiblesse : si Google décide de ne plus envoyer autant de trafic, la chute en terme de fréquentation est très visible. Toujours privilégier des camemberts équilibrés pour minimiser la dépendance (et les sueurs froides lorsqu’un Panda par exemple s’invite dans le jeu).
2. [**lereferenceur**](http://www.doyoubuzz.com/marc-san-emeterio) *says:*

[15 avril 2013 at 12:55](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18606)

Il est souvent conseillé d’avoir le minimum de % venant du trafic SEO afin de ne pas tout perdre si une pénalité pointait le bout de son nez.  
Le travail le plus difficile est donc d’augmenter considérablement le nombre de visites venant des moteurs de recherches tout en faisant baisser le % de visites provenant de ce dernier. Et c’est là que ce n’est pas facile.

Pour le moment votre site marche très bien, vous avez 90% de votre trafic provenant de Google, tout va bien, vous ne faites donc rien ? Imaginez si une pénalité se pointe et que vous perdez 90% de votre CA ? Dur dur…

1. **Véronique Duong** *says:*

[15 avril 2013 at 14:55](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18608)

Merci !! impatiente de lire la suite de ce dossier ;)

1. [**Emmanuelle**](http://www.numastickwebfactory.fr/Star-Ducks-SEO/25/Site-malveillant-suspecte-que-faire-pour-en-sortir.html) *says:*

[15 avril 2013 at 20:59](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18613)

« le référenceur va toujours chercher à rendre la part de trafic en provenance des moteurs plus importante »  
Ben heu… et du coup le rendre de plus en plus sujet aux variations de big GG ?  
Chais pas, mais je l’sens moyen, quand même, non ?

1. [**rocketSeo**](http://www.aurele-gala.com/) *says:*

[15 avril 2013 at 22:04](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18614)

Je suis pas tout fait d accord avec les commentaires précédents. Alors il me semble tout à fait logique de tenter de minimiser la dépendance d un site vis à vis des moteurs de recherche ( enfin Google). La diversité des sources de trafic est la clé bien sur.

La question, c est pourquoi avoir peur des représailles de Google si on est 100% White hat et que l on base sa prestation sur l ensemble des leviers Seo qualitatifs. Je pense que ceux qui craignent le plus les mises à jour sont ceux qui sont borderline. Dans ce cas, il y a en effet des raisons de se faire du soucis, parfois ça passe et un jour ça ne passe plus. ;)

1. **Sylvain** *says:*

[15 avril 2013 at 23:34](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18616)

@Emmanuelle : oui et non, mine de rien il faut bien que tu prouves que ton travail est utile. Alors certes on cherche à augmenter le trafic mais pas forcément sa ventilation, mais dans les faits c’est ce qui va se passer : si on place de bons liens, le trafic référent va augmenter, mais pas à la même vitesse que le trafic SEO. Quant au trafic direct, la notoriété acquise grâce aux liens et au SEO le fera certainement augmenter, mais pas dans les mêmes proportions. Donc j’ai dans l’idée que la part du trafic SEO, même indirectement, on cherche forcément à l’accroître ;-)

1. [**Thierry**](http://marredesefairebaladersurlenet.com/post/2010/12/25/Expert-en-performance-%C3%A9nerg%C3%A9tique-en-Loire-Atlantique) *says:*

[16 avril 2013 at 17:52](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18633)

J’aurais bien mis le paragraphe « conversion » avant le passage sur les landing pages ;-) car on retrouve ces pages dans la majorité des outils de « call to action » en webmarketing, c’est souvent là que tout se joue…

Et d’ailleurs voici un modèle gratuit à télécharger sur ThemeForest :<http://themeforest.net/item/hosted-landing-page/504873?WT.ac=free_file&WT.seg_1=free_file> (il suffit de se créer un compte gratuit pour pouvoir le télécharger)

1. [**Lionel**](http://www.plamour.com/) *says:*

[16 avril 2013 at 23:16](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18639)

Pour la ventilation du trafic, je pense vraiment que cela dépend du type de site. Pour un nouveau site, qui n’a pas d’image de marque, le but est à mon avis d’avoir une très grande part du trafic provenant des requêtes de recherches (plutôt longues traîne du coup car plus faciles au départ).  
Les points sur les landing pages et la conversion sont intéressants, je vais essayer d’améliorer cela !

1. [**Emmanuelle**](http://www.numastickwebfactory.fr/) *says:*

[17 avril 2013 at 08:44](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18645)

Suis pas hyper convaincue… dans le sens où l’idée est de faire du boulot pérenne quand même pour un money site, non ?  
Mais bon, j’ai peut-être mal compris ce que tu voulais dire…  
On en reparle à l’occase autour d’une bière (et p’têt alors que je t’expliquerais le coup du « briquet pendant ») ?  
seeya

1. [**Fred**](http://www.assurancedecennale.fr/) *says:*

[17 avril 2013 at 22:08](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18665)

Article bien rédigé. je trouve que tu as bien segmenté tes différentes parties. Tres accessible pour les référenceurs débutant. Je pense que ta liste d’article sur l’audit SEO devrait être lu par tout les référenceurs qui ont perdu les bases du référencement pour des inscriptions bourrins dans les annuaires ou les techniques Black Seo. Encore une tres bonne partie, j’ai hate de lire la suite de ton article.

1. [**Jerome**](http://www.la-revanche-des-sites.fr/) *says:*

[24 avril 2013 at 11:20](http://noviseo.fr/2013/04/audit-seo-statistiques/#comment-18740)

Je suis complètement d’accord avec l’analyse des landing pages mais surtout parce qu’il faut commencer oublier la donnée « mot clé » certes elle est toujours présente mais de moins en moins alors il faut plutôt penser « optimisation du contenu » par l’analyse des landing pages.