

## Audit de site Internet complet

<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 connaître l'activité des moteurs sur le site
- de connaître 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'appellez 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 😊

## 2. 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 connaître des informations comme :

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

## 3. 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.

## 4. 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.

## 5. 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 😊

## 6. 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](#)

<b>Crawler :</b>	Xenu <a href="http://www.screamingfrog.co.uk/seo-spider/">http://www.screamingfrog.co.uk/seo-spider/</a> Integrity for Mac : <a href="http://peacockmedia.co.uk/integrity/">http://peacockmedia.co.uk/integrity/</a>
<b>Analyseurs backling</b>	

## 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.

### 2. Les statuts http

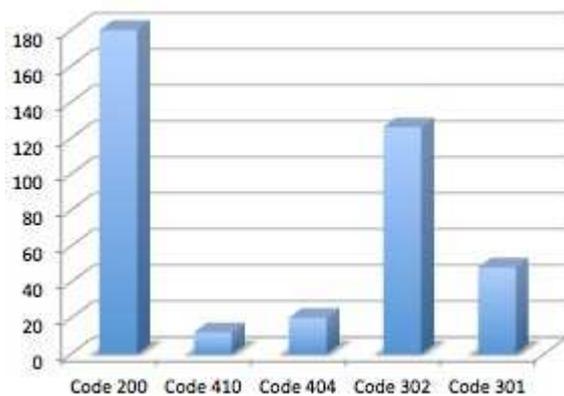
Ce sont tout bêtement les statuts :

- 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.

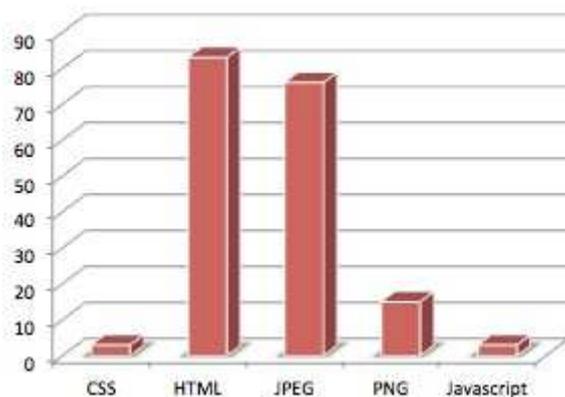
Pour cette analyse là comme pour beaucoup d'autres, un graphique est souvent bien parlant.



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

### 3. 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.

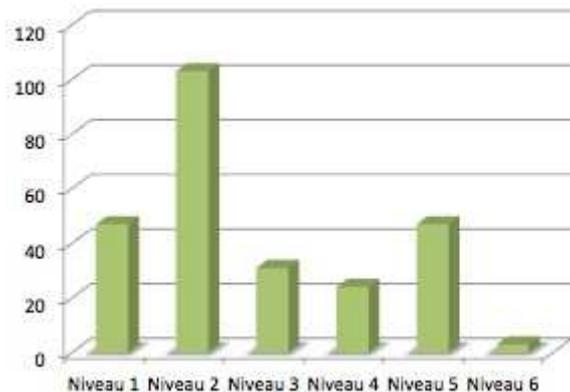


Si le site est 100 % Flash, vous aurez remarqué que vous n'avez qu'une seule URL qui est la page d'accueil. En effet le crawler ne saura pas suivre les liens, vous pouvez donc

arrêter l'audit technique dès maintenant. Si seul le menu de navigation est en flash, vous pouvez continuer mais les résultats seront biaisés.

#### 4. 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.



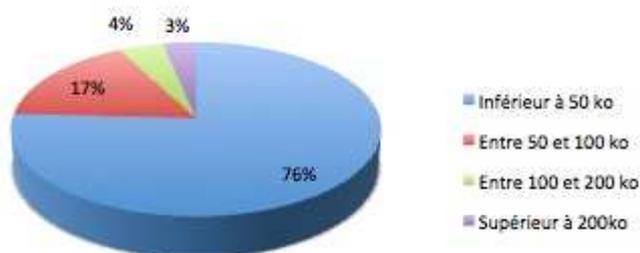
En plus d'avoir une idée du nombre de clics nécessaires pour arriver à la page la plus profonde, cette information nous sera très utile lorsque nous évaluerons l'arborescence du site.

#### 5. 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 ?

#### 6. 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.



#### 7. 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 😊

## 8. 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 ?

## 9. 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 ?

## 10. 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](#).

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 connaîtra 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

## 11. 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 le sitemap n'est pas à jour.

Concernant la deuxième question, est-ce grave si le sitemap n'est pas à jour ? Fondièremment 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 ?

## 12. 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](#).

## 13. 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](#)
- [Page Speed Insight](#)
- [Gtmetrix](#)
- 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 ?

## 14. Le reste :

Vous en avez déjà entendu parlé 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](#)), 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 soumetts ensuite à Screaming Frog pour plus de variables. [Voir la méthode.](#)

Comment utiliser Screaming Frog gratuitement 😊 <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 allez aller y passer du temps. Vous êtes quand même prêt à continuer ? C'est parti !



### 1. 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 catégorie 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-audit.fr Vous saurez ainsi si Google connaît 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'autres 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 !

### 2. 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 ?

### 3. 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'apparaît pas, alors c'est du keyword missing.

### 4. 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.

### 5. 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 😊

## 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 statistique. 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.

### 1. La ventilation du trafic :

Finalement au moment de l'audit le trafic total nous intéresse peu, excepté si on connaît 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 connaître 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 référencement 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.

### 2. 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 connaître 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 😊

### 3. 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 aux 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.



### 4. 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 œil au taux de rebond de ces pages et au temps passé permet de mieux définir si elles sont pertinentes pour l'internaute.

### 5. 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.

### 6. 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

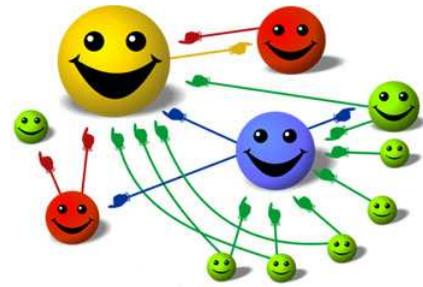
manquements sur le site en termes de contenu.

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 !

## 5ème partie : Audit SEO – Visibilité

16 avril 2013

On continue d'avancer dans cet audit SEO. Après la technique, le contenu et les statistiques, on va maintenant s'attaquer à la visibilité. Pour ne rien vous cacher, c'est une des analyses que j'aurais tendance à réaliser en premier avec l'analyse statistique qu'on a vu juste avant, pour la simple et bonne raison que ce sont ces deux points qui permettront de mettre le doigt sur les éventuels problèmes existants.



Tout le monde est prêt à reprendre cet audit ? C'est parti !

### 1. Le site est-il crawlé par Google ?

J'aurais pu mettre cet aspect dans la section technique de l'audit mais j'ai pris le parti de le placer dans la section visibilité, après tout pour qu'il soit visible il faut bien que Google le crawle. Pour voir si le site est bien crawlé par Google, rien de plus simple : analysez vos logs et repérez l'activité de Googlebot. Est-ce qu'il a bien visité toutes les pages qui nous intéressent ? À quelle fréquence visite-t-il les pages ? Ces informations permettront de savoir si un potentiel manque de visibilité est dû ou non à des difficultés liées au crawl du site.

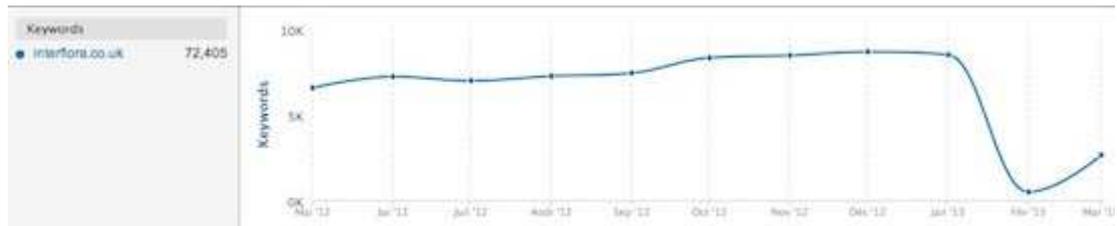
N'hésitez pas à créer des graphiques que vous diviserez selon les types de pages pour faciliter l'analyse du crawl.

### 2. L'évolution du positionnement :

Petit aparté : je fais de moins en moins attention au positionnement de mots clés particuliers (un comble pour un SEO n'est-ce pas ?) pour plusieurs raisons dont une particulièrement : entre la géolocalisation, l'historique de recherches, l'historique tout court, les préférences de recherches, le contenu de la boîte Gmail (parano, peut-être, mais pourquoi pas ?) le positionnement ne veut plus rien dire. Cela dit, je vérifie quand même toujours l'état de mes tops keywords 😊

Revenons-en à l'audit, malgré ce que je viens de dire, il faut tout de même reconnaître qu'une analyse du positionnement global s'avère nécessaire. J'entends par "positionnement global" l'ensemble des mots clés sur lesquels le site se positionne. [SEMrush](#), [SEMvisu](#) ou encore les [Webtools Yooda](#) sont des outils qui permettent de réaliser cette analyse, bien qu'ils soient tous les deux basés sur un échantillon de mots-clés et risquent de ne pas prendre en compte toutes les expressions de longue traîne.

L'idée ici est de voir l'évolution dans les SERP : y a-t-il de plus en plus de mots-clés qui se positionnent ? De moins en moins ? Y a-t-il eu une forte chute de visibilité à un moment donné ?



Bien qu'indiquant le nombre d'impressions et non pas le nombre de mots clés positionnés, la section Trafic ==> Requête de recherches de Google Webmaster Tools peut nous permettre d'effectuer une analyse approchante, sur trois mois maximum.

### 3. La répartition du positionnement :

Je vais ici prendre l'exemple de SEMrush que je connais bien. Cet outil nous liste les mots-clés pour lesquels le site se positionne sur les deux premières pages de résultats Google. Seulement voilà, entre une première position sur la première page et une dixième position sur la deuxième page, il y a tout un monde. L'idée est donc de savoir combien de mots-clés sont dans les top positions et combien de mots-clés sont un peu plus "reculés". Le graphique du paragraphe suivant illustre bien cela 😊 L'objectif ici est d'estimer l'évolution possible de la visibilité dans les SERP.

J'ai pris pour habitude de répartir le positionnement de cette façon : positions 1 à 3, positions 4 à 10 et positions 11 à 20 pour avoir un aperçu des possibilités d'évolution en fonction des positions "stratégiques".

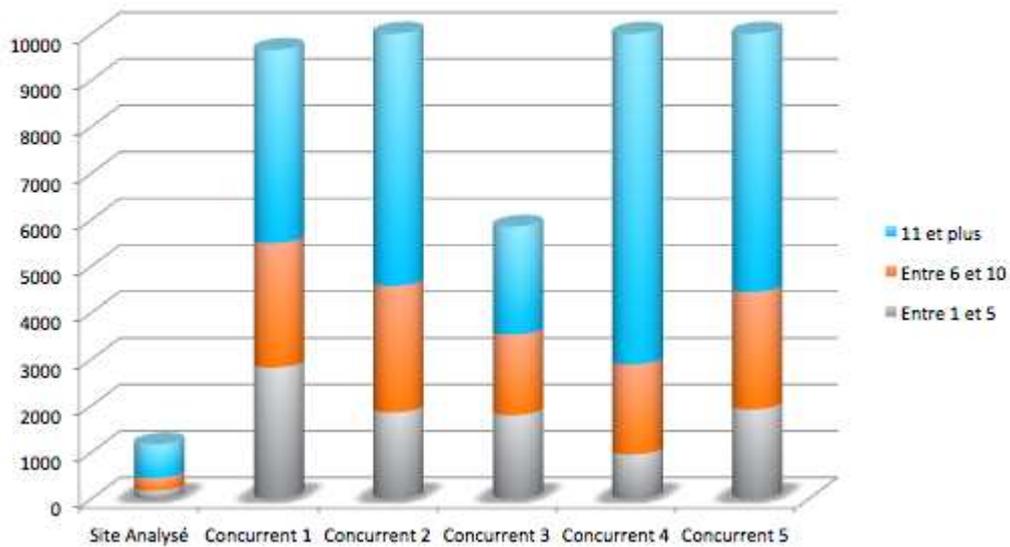
### 4. L'analyse de la concurrence :

On parle bien sûr ici de la concurrence dans les SERP qui peut être différente de la concurrence "de la vie réelle" (Wikipédia ou Comment Ça Marche ne sont pas des concurrents sur votre marché mais le sont dans les SERP).

Et bien ici nous allons pouvoir faire la même analyse que précédemment concernant la répartition du positionnement pour chacun des concurrents des SERP. Enfin, presque... Cas pratique : vous proposez des vacances en Allemagne, si vous cherchez "vacances Allemagne" sur Google, le premier résultat est dates-vacances-scolaires.com. Ce n'est pas un concurrent de la vie réelle mais c'est un concurrent dans les SERP... Mais pour cette expression seulement ! Si vous analysez ses mots-clés, vous allez tomber sur des expressions du type "vacances scolaires" qui ne vous concernent pas particulièrement. Il faudra donc faire preuve d'un peu de discernement et ne pas foncer tête baissée 😊

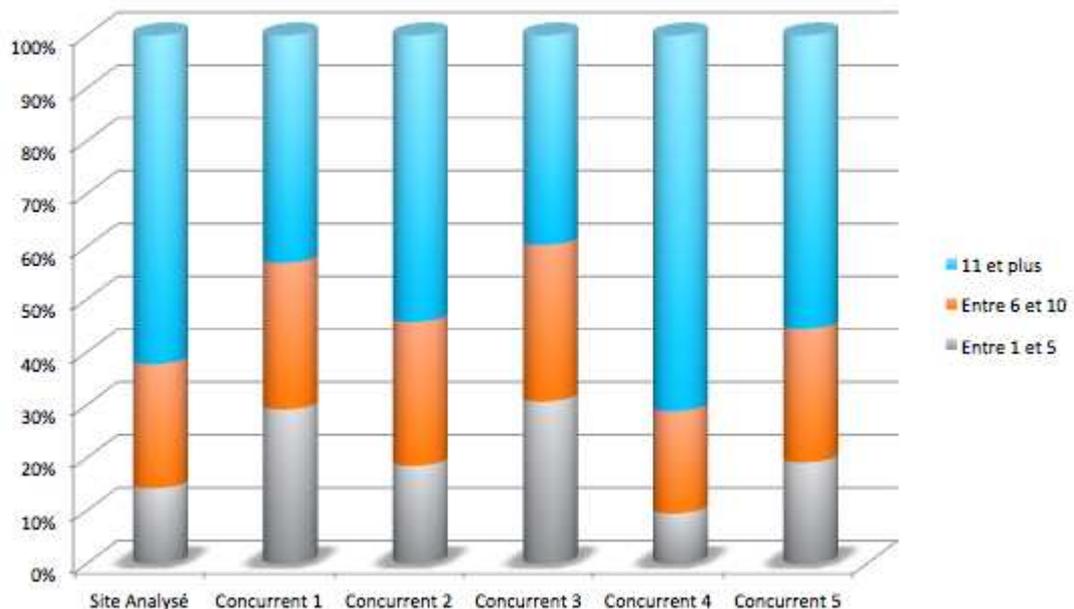
Donc une fois que vous avez repéré et analysé les concurrents, il va falloir mettre en place un système qui permettra de visualiser votre visibilité par rapport à la leur. Encore une fois, un graphique est toujours bien parlant :

## Positionnement des mots clés



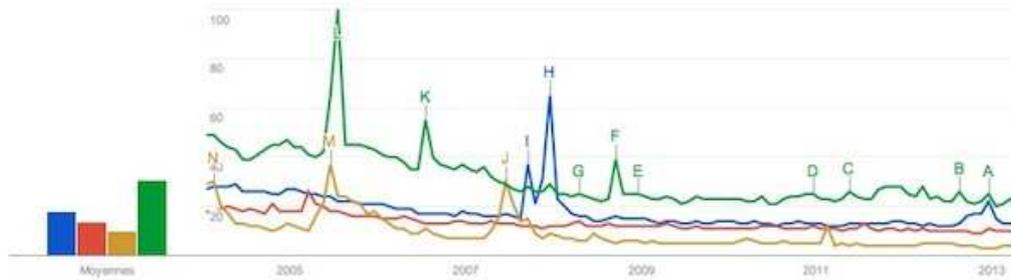
En plus de ce graphique, vous pouvez bien évidemment comparer la répartition du positionnement :

## Répartition du positionnement de mots clés



### 5. Les tendances de recherches :

On entre déjà dans les recommandations post-audit, je m'étais pourtant promis de ne pas en parler dans ce dossier, tant pis. Un petit passage par [Google Trends](#) vous permettra de vérifier l'évolution des recherches sur les expressions clés principales que vous aurez choisies. Ceci n'est pas réellement (ou pas uniquement) lié à la visibilité. En effet, ça a plutôt un effet sur le contenu. Et sur les statistiques. Et un peu sur la visibilité quand même.



En plus de ces tendances, Google Trends donne lui aussi une liste de recherches associées, pensez-y pour vos suggestions de mots clés !

Mieux encore, Google Trends nous permet d'avoir un aperçu du futur. Lorsque vous êtes sur le graphique de tendance des recherches, jetez un oeil en haut à droite et cochez la case "Prévision" pour voir comment la requête va évoluer dans le futur. Si ça ne donne pas vraiment d'indications sur la visibilité future du site, ça en donne sur la concurrence qui risque se mettre en place et sur vos efforts à fournir.

Ne prenez pas ces prévisions pour argent comptant, veillez aussi à l'actualité : par exemple je ne sais pas à quoi ressemble la courbe pour "hôtel Rio de Janeiro" ni ses prévisions, mais j'ai dans l'idée qu'avec les J.O. et la coupe du monde de foot, cette requête sera amenée à exploser 😊

On s'approche de la fin de cet audit, c'est pas trop tôt, ça fait déjà 5 jours qu'on est dessus ! Mais il faut avouer que ça vaut le coup, non ? Allez, on se retrouve demain pour l'audit de liens.

## 6ème partie : Audit SEO – Liens

17 avril 2013

Ça y est, on se rapproche encore et toujours de la fin de l'audit. Nous avons effectué l'audit technique, l'audit de contenu, l'analyse statistique et un audit de visibilité ; c'est le moment de s'attaquer aux liens du site.



Contrairement à ce qu'on peut lire parfois (de moins en moins, je vous l'accorde), il ne suffit pas de compter le nombre de liens entrants pour analyser la qualité du linking d'un site. Il y a bien sûr plein d'autres choses à prendre en compte, et c'est ce qu'on va voir ici.

### 1. Lister les liens entrants :

Je viens de dire que ça ne suffisait pas, mais il faut tout de même le faire ! Voici quelques unes des infos à récupérer :

- total des liens entrants
- total des domaines référents
- les pages du site qui récupèrent les liens
- les ancres de liens

Vous pouvez récupérer ces informations (et d'autres bien utiles que nous verrons un peu plus loin dans l'article) avec des outils comme Ahrefs, Open Site Explorer, Majestic SEO, Link Research Tool, etc.

Après quelques manipulations, Google Webmaster Tools vous permet également de récupérer ces informations gratuitement. Rendez-vous dans Trafic ==> Liens vers votre site.

### 2. Le type de liens :

Nous allons différencier les liens selon leurs types :

- ratio de liens follow/nofollow
- ratio de liens sitewide (i.e. sur toutes les pages du site)/non sitewide
- source du lien (texte, image, etc.)

On ne va pas s'embêter, Ahrefs fait lui-même ce travail



Ici on a déjà un bon aperçu du profil des liens : naturels, pas naturels ? Beaucoup trop de nofollow, pas assez ? Que du sitewide ?

### 3. Les TLD :

On va passer rapidement sur les TLD (Top Level Domain : .fr, .com, .net, .de, etc.), l'objectif ici est de repérer qu'on n'a pas trop de liens depuis des domaines

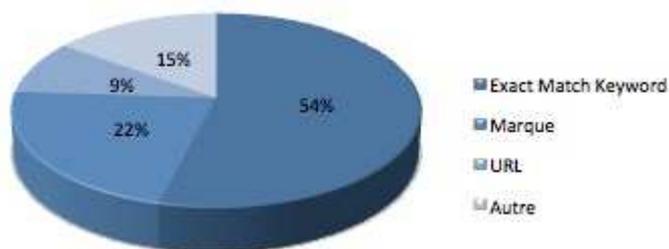
exotiques, ce qui pourrait signifier que du Black Hat a été effectué sur ce site, voire du Negative SEO (ce qui n'arrive que très rarement 😊).

#### 4. Les ancres des liens :

Ce que nous allons faire ici, c'est simplement différencier les types d'ancres pour déterminer si le profil est assez naturel et, le cas échéant, préparer un plan d'action pour la suite. Pour l'exemple, voici comment j'ai réparti les types d'ancres :

- Exact Match Anchor (l'ancre qui correspond exactement au mot-clé visé)
- Marque
- URL
- Autre (ancres génériques, liens sur des morceaux de phrases, etc.)

Et voilà le résultat :



#### 5. Les pages recevant les liens :

Plusieurs questions sont à se poser ici également :

- est-ce que la home reçoit la quasi totalité des liens ?
- est-ce qu'au contraire, elle ne reçoit que très peu de liens ?
- quel type de page reçoit beaucoup de liens (contenu éditorial, produit, catégorie, etc.) ?
- les pages qui reçoivent les liens répondent-elles correctement (pas de 404) ?
- sont-elles toujours d'actualité, à jour ?
- est-ce que ce sont bien des pages que l'on cherche à positionner ?

Maintenant que nous avons répondu à toutes ces questions, nous avons des informations sur :

- le profil du référenceur qui est passé avant nous
- le type de contenu qui est apprécié, et qui est donc partagé

Le plan d'action pour le futur se précise donc petit à petit.

En couplant ces informations avec celles récupérées lors de l'analyse statistique, particulièrement les sites référents, nous avons pas mal de billes supplémentaires pour le plan d'action de linking 😊

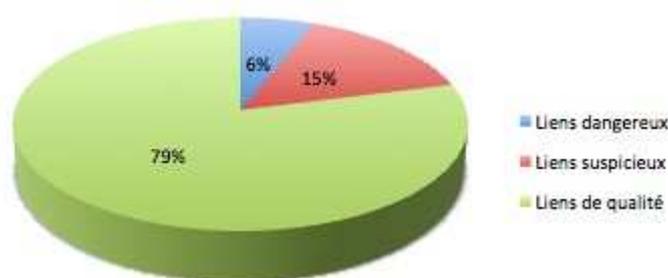
#### 6. Déterminer la qualité des liens :

C'est ici que ça va commencer à se corser : comment déterminer la qualité des liens ? Si nous ne cherchons que des infos "officielles" nous allons être vite limités : ce sera le PageRank de la page liante et de la home qui nous donnera des indications sur la qualité du lien. Attention tout de même à ne pas crier au loup dès qu'un lien

provient d'un petit PR : il s'agit peut-être d'un très beau lien sur un nouveau site. Ce qui nous emmène à la règle d'or : **toujours vérifier les liens manuellement**, au moins quand on a un doute 😊

Cette vérification permettra également d'appliquer une sorte de coefficient de pondération : un lien au sein d'un texte, dont la page et le site complet parlent de la même thématique que le site audité vaut mieux qu'un lien en commentaire sur un site hors sujet.

Pour nous faciliter le travail, certains outils ont déjà établi une sorte de classement. C'est le cas de Ahrefs qui a mis en place un Ahrefs Rank pour les pages et un Ahrefs Domain Rank pour les domaines. Le problème de ces systèmes est qu'on ne sait pas réellement sur quoi ils se basent, mais on peut prendre le pari que s'ils ont décidé d'implémenter ces classements dans leur outil, c'est qu'ils estiment qu'ils sont sérieux.



Voici le résultat auquel vous pouvez arriver une fois ce travail effectué.

Tout comme Ahrefs, [l'API Mozscape](#) a mis en place une notation des domaines. Sauf qu'elle, il est possible de l'utiliser gratuitement même sur de gros volumes après quelques paramètres 😊

## 7. Le nombre de liens par rapport à l'âge du site :

Ici encore, je vais surtout parler de l'après-audit, ce que je n'étais pas censé faire. Effectuer un ratio nombre de liens/âge du site permet d'affiner un peu plus le plan d'action. Admettons que le site audité existe depuis un an. À ce jour, il a 120 liens entrants, ce qui signifie en moyenne 10 liens par mois. Pour rester naturel, ne vous éloignez pas trop de ce nombre de 10 liens par mois, tout en l'augmentant tout de même. Vous pourrez donc proposer 13 ou 14 liens par mois pour les mois suivants, tout en augmentant petit à petit ce nombre.

Vous pouvez même être plus précis en vous basant sur la qualité des liens vues dans le paragraphe précédent. Et oui, le linking naturel, c'est beaucoup de réflexion !

Allez, plus que deux petits articles et nous arrivons à la fin de ce dossier, dans deux jours vous serez prêt à effectuer un bel audit, un vrai, un qui rendra vos cheveux lisses et soyeux et qui fera de vous une rockstar. À demain pour l'activité sociale !

## 7ème partie : Audit SEO – Social

18 avril 2013

La partie sociale constitue le dernier checkpoint avant la grande ligne droite pour terminer cet audit de référencement qui nous aura tout de même occupé pendant une semaine. Avant toute chose, je tiens à préciser que ce n'est pas un audit de la stratégie sociale que l'on va faire ici mais plutôt un audit de la relation existante entre le site et les médias sociaux. Ça, ça en jette.



Une seconde précision : très souvent, la thématique du site fait que cette partie de l'audit ne s'avèrera pas nécessaire. D'ailleurs, je n'ai jamais eu à la réaliser. Mais ça ne m'empêche pas d'y avoir réfléchi 😊 Et voilà comment j'organiserais ça, **tout en restant dans le cadre de l'audit SEO** sans déborder sur un "audit SMO".

### 1. État des lieux du partage existant :

Celui-ci, vous vous en doutiez. Réaliser un état des lieux de partage sur chaque page est le point de départ quasi-obligatoire. Vous pouvez utiliser un outil générique (type Social Crawlytics qui décompte les partages sur Facebook, Twitter, Pinterest, Google+, LinkedIn, Delicious e StumbleUpon) pour effectuer cet état des lieux ou développer vous-même votre propre outil : très souvent, les réseaux sociaux partagent une API permettant aux développeurs de faciliter leur travail d'analyse.



Trois objectifs se cachent derrière cette action : connaître l'engagement des lecteurs envers le contenu du site, déterminer quel type de contenu est le plus apprécié et évaluer l'impact social sur le SEO. Si le premier reste relativement simple à comprendre, nous allons détailler les deux suivants.

Si vous auditez un site sous WordPress, vous pouvez installer l'extension Social Metrics qui existe en version gratuite (mais limitée) ou en version payante, elle vous fera le décompte des partages sur les principales plateformes.

### 2. Quel type de contenu est partagé :

Une fois l'état des lieux des partages effectué, nous allons classer les pages des plus partagées aux moins partagées. Si tout va bien, nous aurons à l'arrivée une typologie de ce qui plaît aux lecteurs, c'est donc une bille supplémentaire pour la mise en place d'une stratégie future.

Si une stratégie SMO existe (ou existera), pensez à déterminer quel média amène le plus de partages : ce sera le média à privilégier dans cette stratégie.

### 3. L'impact social sur le SEO :

Grand débat que voilà, les anti-social (facile, très très facile...) ne me suivront pas là-dedans, pourtant les réseaux sociaux peuvent avoir un impact sur le SEO, plus ou moins fort. Pour évaluer cet impact, il faudra diviser l'analyse en fonction de chaque réseau :

- les liens partagés sur le réseau sont-ils visibles ? Il faudra faire une estimation du pourcentage de profils privés sur le réseau pour avoir une estimation du nombre de partages visibles (allez convaincre un anti-social -oui elle me plait cette vanne- avec ça...),
- les liens sur le réseau sont-ils crawlables ? À ma connaissance il n'y en a pas, mais peut-être que certains réseaux partagent leurs liens en javascript par exemple,
- les liens sur le réseau sont-ils en nofollow ? Très souvent oui, mais pas toujours 😊
- les liens sur le réseau sont-ils cachés derrière des redirections ?
- le réseau propose-t-il un flux RSS pour chaque compte ? Auquel cas, il est probable que le flux reprenne les liens partagés,
- les liens sur le réseau sont-ils permanents ou temporaires ?

Nous avons ici pas mal de points à vérifier pour déterminer si oui ou non, les partages ont un impact sur le SEO du site cible.

### 4. Les profils sociaux :

Je l'annonce dès maintenant, ou plutôt je le répète (parce que je vous vois venir avec vos gros sabots), on regarde ici **uniquement** l'impact SEO et non l'impact social, exit donc le nombre de fans/followers/encercleurs/reliés et autres, exit la qualité de la communauté (fake ou pas), exit le taux d'engagement, etc

On va donc se poser plusieurs questions :

- sur combien de réseaux le site a-t-il un profil ?
- ces profils sont-ils visibles ?
- y a-t-il possibilité de mettre en place des liens sur ces profils ?
- est-ce que ça a été fait ?
- des pages du site ont-elles également été partagées sur ces profils ?
- etc.

En résumé, est-ce que toutes les opportunités SEO ont été saisies ?

On dérape vers le SMO, mais est-ce que le site est présent sur les réseaux de sa thématique (Soundcloud ou Youtube pour une boîte de prod, LinkedIn pour un bureau de recrutement, etc.) ? N'oubliez pas qu'en SEO, le contexte du lien importe également 😊

Je rappelle une fois de plus que cette partie ne tient compte que des bénéfices SEO des médias sociaux, le travail sur la stratégie en place, le taux d'engagement, les résultats obtenus, etc. concernent le SMO qui n'est pas le sujet ici 😊 Je vous retrouve demain pour le dernier article de la série ?

## 8ème partie : Audit SEO – Et après ?

19 avril 2013

Enfin ! Dernier article de dossier sur l'audit de référencement ! Vous avez aimé ? Vous en voulez encore ? Nous allons commencer ici par faire un récapitulatif de tout ce qui s'est dit jusqu'à présent, nous allons ensuite évoquer les outils utilisés, puis nous allons parler de la façon de présenter les résultats pour enfin terminer avec "l'après audit".



Avant toute chose, je rappelle qu'un audit doit être réalisé **en fonction des objectifs définis avec le client** : sans objectif, aucun audit digne de ce nom ne peut être réalisé. Deuxième précision : vous avez dans ce dossier la liste quasi-exhaustive des points à prendre en compte, mais derrière tout ça il va falloir montrer les défauts, pointer du doigt les éléments bloquants, proposer des solutions, et tout ceci demande une réelle connaissance du SEO, si j'osais je parlerais même d'expertise, et ça vous ne pourrez pas l'acquérir en lisant des blogs, des livres, des forums, etc.

### L'audit SEO de A à Z

Comme annoncé en introduction, nous allons faire un petit récapitulatif des différentes étapes de cet audit :

- [Les pré-requis](#) : tout ce qu'il faut avoir en sa possession pour réaliser l'audit le plus complet possible, des différents accès aux outils nécessaires
- [l'audit technique](#) : les points essentiels à analyser d'un point de vue purement technique
- [l'audit de contenu](#) : comment effectuer un état des lieux du contenu du site, s'il est adapté aux objectifs, etc.
- [l'analyse statistique](#) : quelles sont les métriques à prendre en compte lors d'un audit SEO ? Quelles sont celles qui importent et qui nous apporteront des informations pertinentes ?
- [l'audit de visibilité](#) : comment se situe le site par rapport à la concurrence ? Comment peut-il évoluer ?
- [l'audit de netlinking](#) : quelle est la popularité du site sur le web ? Comment la définir ?
- [l'audit social](#) : s'il est concerné, comment évaluer la portée sociale du site ?

## Les outils de l'audit

Ce paragraphe n'était pas prévu à l'origine, les outils existants étant déjà listés un peu partout sur le web. Mais comme il semblerait que vous en soyez friands, voici une liste non-exhaustive d'outils existants sur le marché et qui vous seront utiles pour réaliser un audit. Petite précision : l'habit ne fait pas le moine, avoir les bons outils ne suffit pas à faire un bon audit !

### 1. Outils de crawl

- Screaming Frog SEO Spider
- Xenu's Link Sleuth
- Inegrity

### 2. Outils d'analyses de liens

- Ahrefs
- Majestic SEO
- Open Site Explorer
- Link Research Tool

### 3. Outils de visibilité

- SEMrush
- SEMvisu

### 4. Outils de positionnement

- Ranks
- MyPoseo

### 5. Outils de performances

- Yslow
- Page Speed Insights
- Pingdom Tools
- GT Metrix

### 6. Outils d'analyse sociale

- Social Crawlytics

### 7. Les outils multi-usage

- Yooda SeeURank
- Advanced Web Ranking

Ajoutez à cela des outils de statistiques qui normalement sont déjà intégrés sur le site, et la suite Office (ou assimilé) pour analyser et présenter le tout. D'ailleurs, en parlant de ça...

## Comment présenter l'audit ?

Plusieurs paramètres entrent en jeu : allez-vous présenter votre audit face au client ? Ce serait bien sûr l'idéal, mais ce n'est pas toujours possible. Donc, présentation PowerPoint ou dossier Word ? Et bien il n'y a pas de règle, vous choisissez. Ou plutôt, si, il y a une règle : **la présentation doit être claire et facilement compréhensible**. Le client ne



