

**Lettre professionnelle
Recherche et Référencement**

Numéro 161

Juillet-Août 2014



SARL RESEAU ABONDANCE - Olivier Andrieu - Consultant -

3 rue des Châteaux - 67140 Heiligenstein -

Tél : 03 88 08 83 26 - Fax : 03 88 08 83 27 -

E-mail : olivier@abondance.com - Web : www.abondance.com

Recherche et Référencement

Numéro 161 - Juillet-Août 2014

-- Lettre d'actualité mensuelle sur la recherche d'information et le référencement de sites Web --

Au sommaire ce mois-ci :

- > [Les personnes incontournables dans le "search" chez Google](#) (4 pages)
- > [Les services de Content Discovery \(3ème partie\)](#) (9 pages)
- > [Comment bien utiliser l'outil de planification des mots clés de Google ?](#) (8 pages)
- > [Comment utiliser intelligemment les Sitemaps XML ?](#) (11 pages)
- > [Google a-t-il le droit de "scrapen" nos contenus ?](#) (4 pages)

- > [Revue d'outils moteurs et SEO](#) (5 pages)
- > [Revue d'URL](#) (9 pages)

Lettre parue le mardi 15 juillet 2014 (prochaine parution : lundi 15 septembre 2014)

Le contenu de cette lettre est accessible sur la zone "Abonnés" du site Abondance, à l'adresse : <http://abonnes.abondance.com/>

La lettre "Recherche & Référencement" paraît aux alentours du 15 de chaque mois (un seul numéro pour les mois de juillet-août)
Pour tout renseignement : © Olivier Andrieu, olivier@abondance.com

IMPORTANT : Pour continuer de vous délivrer la meilleure information possible, la lettre "Recherche & Référencement" doit être utilisée au mieux par ses abonnés, qui doivent notamment se plier aux règles suivantes :

- * Il est **interdit d'envoyer, par mail** ou autres systèmes de transmission, un exemplaire de la lettre à quelque format que ce soit, à une personne qui ne serait pas abonnée.
- * Il est **interdit de diffuser à l'intérieur d'une entreprise, à plusieurs adresses e-mails ou sur un intranet**, le contenu de la lettre "Recherche & Référencement" si vous n'avez pas souscrit un **abonnement "Entreprise"** (<http://lettres.abondance.com/rechref.html>) qui est fait pour cela.
- * Il est **interdit de transmettre vos login / mot de passe** à des personnes qui ne seraient pas abonnées. Pour information, les adresses IP des machines se connectant sont relevées et tout abus pourrait être sanctionné si ce phénomène ne pouvait être justifié par vos services techniques.
- * Il est **interdit d'utiliser à des fins privées ou publiques tout contenu** de cette lettre sans notre autorisation expresse et écrite : envoi à des clients, des prospects, voire reprise intégrale du contenu en changeant le nom de la lettre et de la source (eh, oui, ça arrive...), etc.

De nombreux abus ont été constatés ces dernières semaines (merci à ceux qui nous les ont signalés) aussi, nous nous permettons de vous rappeler ces quelques règles de "savoir lire" :-)...

Merci pour votre compréhension et bonne lecture.

Toute l'information contenue de cette lettre est © Abondance / Olivier Andrieu. Toute reproduction ou distribution des informations, données et textes de cette lettre est interdite sans l'autorisation expresse du site Abondance. Toute reproduction ou représentation, intégrale ou partielle de ce contenu, faite sans le consentement de l'auteur, serait illicite. La loi n'autorise, que les copies ou reproductions réservées à l'usage privé du copiste et non destinées à l'utilisation collective, d'une part, et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration.

Sommaire du numéro 161 - Juillet-Août 2014

Les personnes incontournables dans le "search" chez Google	4
Les services de Content Discovery (3ème partie)	8
Comment bien utiliser l'outil de planification des mots clés de Google ?	17
Comment utiliser intelligemment les Sitemaps XML ?	31
Google a-t-il le droit de "scrapen" nos contenus ?	42
Revue d'Outils Moteurs et SEO	46
Revue d'URL	51

Les personnes incontournables dans le "search" chez Google

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Le moteur de recherche Google est conçu et maintenu chaque jour par des milliers d'ingénieurs. Mais, parmi ceux-ci, certains sont des rouages essentiels de la machine, les « têtes pensantes » qui ont conçu et développent encore les nouveautés qui apparaissent pour innover et améliorer l'outil. Voici la présentation de 5 d'entre eux (en dehors, bien sûr, des classiques Larry Page, Sergey Brin et Matt Cutts, trop « célèbres » pour être listés ici). Les connaissez-vous tous ? Pas si sûr. En fait, nous sommes certains que non...

Avec l'été et les congés souvent associés, on aime lire des choses plus légères et moins techniques. Nous avons décidé de ne pas déroger à cette tradition, en vous proposant quelques portraits des personnes qui sont importantes dans le « search » chez Google. Pour chacune de ces personnes, nous vous expliquerons ci-dessous en quelques phrases son rôle chez Google, et ses grands faits d'armes. Certains sont importants au sein du dispositif de la firme de Mountain View parce qu'ils ont fait de grandes contributions techniques, d'autres font simplement « tourner la boutique », mais tous sont essentiels dans les rouages de la machine.

Vous remarquerez qu'il manque quelques grands noms que vous connaissez tous. Point de Sergey Brin ou de Larry Page ici, encore moins de Matt Cutts, c'est totalement voulu, car vous savez déjà parfaitement qui sont ces gens et il n'est pas obligatoire de les présenter ici...

Amit Singhal

Le cheminement qui a amené Amit Singhal chez Google est assez amusant. Il a commencé ses études en Inde, et comme beaucoup, est parti aux États-Unis pour obtenir un master en sciences. C'est à l'Université du Minnesota qu'il va arriver, pour mieux rebondir puisque c'est là qu'un de ses professeurs l'aiguille vers la recherche et plus particulièrement vers Gerard Salton. C'est sous la direction de ce dernier qu'il va faire une thèse sur la recherche d'information, qu'il soutiendra en 1996.



Si vous lisez la lettre d'Abondance régulièrement, le nom de Salton (et son fameux cosinus : <http://recherche-referencement.abondance.com/2014/02/le-cosinus-de-salton-un-classique.html>) ne vous est pas inconnu. Et il est évident qu'une grande partie des choix algorithmiques fait par Google doivent beaucoup à la continuité de pensée entre Salton et Singhal, même si ce dernier a réussi à s'affranchir de son maître en poussant le knowledge graph à partir de 2010.

Sa carrière n'a cependant pas commencé chez Google, puisqu'avant il avait été chercheur aux AT&T Labs (ce que le grand public connaît plus sous le nom de Bell Labs). Il y travaillait déjà sur les mêmes problématiques : classement de documents, création automatique de résumé, mécanisme de réponse automatique, compréhension du langage parlé.

En 2000, il va être embauché par Google, et c'est là que sa carrière va décoller de manière très spectaculaire. Il devient rapidement (et reste toujours) le principal architecte de l'algorithmique du moteur. Il est actuellement vice président senior et Google Fellow, ce dernier titre étant réservé à l'élite des ingénieurs travaillant chez le géant du web.

Amit Singhal est raisonnablement connu du grand public, en particulier depuis qu'un article de 2007 du New York Times l'a présenté son un jour très flatteur.

Jeff Dean

La plupart des lecteurs de ces lignes n'auront sans doute jamais entendu parlé de lui, mais Jeff Dean est une légende dans le milieu de l'informatique. C'est à lui plus qu'à n'importe qui d'autre que Brin et Page doivent leurs milliards.



Vous trouvez que Google est efficace pour le crawl ? C'est lui qui a conçu et écrit le crawler. Vous trouvez que le système de fourniture des pubs est bien fait ? C'est lui. Vous trouvez que MapReduce est une bonne idée ? C'est lui qui en a fait la mise en place (l'idée était déjà connue). BigTable, Distbelief, AdSense for content, etc. Jeff Dean est partout dès qu'il y a une brique technique à mettre en place et à rendre rapide et efficace. Si vous avez cinq minutes pour chercher et voir ses réalisations, regardez donc sa page chez Google, c'est édifiant.

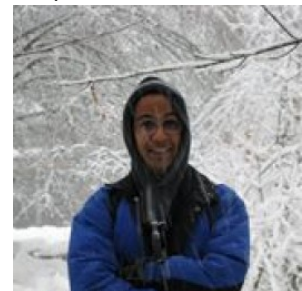
Comme beaucoup, il a une thèse en informatique, mais la sienne est dans un domaine très technique : l'optimisation du code pour les langages orientés objets. Très rapidement, il quitte le monde académique et se retrouve en 1999 chez Google, qui n'a alors qu'une vingtaine d'employés. C'est sous sa houlette que la plupart des gros développements cruciaux pour Google vont se faire, et son talent va permettre de faire la différence. Il va ainsi mettre en place une architecture logicielle ultra performante pour coller le plus possible à l'architecture algorithmique mise en place par les « theory guys », permettant des temps de réponse très courts sur une masse d'information impressionnante.

Dans le milieu de l'ingénierie informatique, on appelle les personnes comme Jeff Dean des « 10x ». Le terme a souvent été utilisé par Larry Page pour dire que chez Google il faut faire les choses 10 fois mieux qu'ailleurs. Concernant Jeff Dean, cela veut dire qu'il est capable de faire seul ce que feraient 10 ingénieurs travaillant en équipe. C'est sans doute une vision un peu trop romancée de sa personne, mais il est indéniable qu'il est l'un des meilleurs ingénieurs en informatique sur le marché. Certaines rumeurs annonçant d'ailleurs que Google le paierait 3 millions de dollars par an pour le garder. Pour clore sur Jeff Dean, nous vous laissons lire les Jeff Dean Facts (équivalent des Chuck Norris Facts) mais nous ne résistons pas à vous donner celui là : « Jeff Dean écrit ses programmes directement en binaire, puis il écrit le code source pour servir de documentation aux autres développeurs ». C'est tout dire...

Navneet Panda

Autant les référenceurs ne connaissent pas forcément Jeff Dean, autant le nom de Panda est plus que connu (malheureusement, diront certains). Il est cependant assez difficile de trouver des informations fiables sur cet ingénieur dont Amit Singhal a parlé une seule fois : « *Well, we named it internally after an engineer, and his name is **Panda**. So internally we called a big Panda. He was one of the key guys. He basically came up with the breakthrough a few months back that made it possible.* »

On sait de Navneet Panda qu'il a fait une thèse à l'Université de Californie à Santa Barbara, sous la direction de Edward Chang, dans le domaine de l'apprentissage et de la fouille de données (en anglais : *Machine Learning* et *Data Mining*). Ces travaux publiés dans les meilleures conférences internationales sont de très bonne qualité, et il est très clairement un spécialiste de l'accélération des méthodes d'apprentissage et classification. Tout cela est parfaitement cohérent avec ce qu'est Panda : un algorithme de classification ultra-rapide sur un index de taille très conséquente.



Depuis la sortie de Panda, Navneet Panda s'est fait très discret (est-ce vraiment lui sur la photo ci-dessus à droite, trouvée sur le Web ?). La plupart des informations le concernant ont disparu, y compris sa page personnelle qui était hébergée sur Googlepages. Il faut dire qu'il n'est pas très bien vu par tout le monde...

Chez Google, il y a un autre Panda, il s'agit de Biswanath Panda (ci-contre). Il travaille sur l'accélération à grande échelle de la fouille de données et ses algorithmes pourraient être utilisés pour faire de la détection de webspam à grande échelle. Mais les choix techniques semblent indiquer qu'à l'heure actuelle le travail de Navneet fait tourner le filtre Panda, tandis que celui de Biswanath serait plutôt utilisé pour la gestion des publicités.



Ben Gomes

Ben Gomes est lui aussi un Google Fellow. Comme bien des ingénieurs de la firme de Mountain View, il a obtenu une thèse en informatique, dans le domaine de l'intelligence artificielle (c'était en 1997). Le lecteur avisé remarquera aussi que sa thèse a été obtenue dans une université de la Californie (ici il s'agit de Berkeley), comme la plupart des *top googlers*.



Il est chez Google depuis les débuts, alors qu'il ne s'agissait encore que d'une petite startup, et il a contribué à un nombre incalculable d'améliorations du moteur, depuis le rajout de la correction orthographique jusque des changements algorithmiques plus conséquents. On lui doit par exemple la mise en place de la recherche instantanée.

Il est en charge de l'ingénierie globale du moteur, c'est-à-dire que c'est lui qui gère le « package » : interface, composants algorithmiques, surcouches diverses, etc. Son travail est donc un travail un peu ingrat : il doit s'assurer que le moteur fonctionne parfaitement quand des troupes d'ingénieurs surexcités collent toutes sortes de choses dessus, mais c'est un job totalement indispensable.

Et parce que c'est l'été et que l'été on fait dans le people, voici un petit fait amusant : un des amis d'enfance de Ben Gomes est Krishna Bharat, qui a développé Google News. Bon, c'est du people à la sauce Google...

Jon Wiley

Le dernier Googler que nous souhaitons vous présenter est Jon Wiley. C'est le monsieur UX de chez Google : il est en charge de l'expérience utilisateur. Il est chez Google depuis 2006, et travaille de manière transversale, sur tous les produits Google.



Au début de cet article, nous parlions des gens qui sont là pour faire tourner la boutique, même si ils ne sont pas des top ingénieurs. Jon Wiley est de ceux là. Avant de bosser pour Google, il travaillait chez Hoovers, qui est un data broker. Encore avant ? Il était administrateur d'un site web (celui du gouverneur du Texas), et précédemment, développeur web. Bref, il n'est pas un chercheur de renom, il n'est pas un grand technicien, mais il est l'homme qui fait que le moteur est simple à utiliser, il est aussi l'homme responsable des logos (doodles).

Jon Wiley est au final aussi important que les autres, car c'est grâce à lui que la première adhésion au moteur se réalise : l'outil est simple à utiliser, donc il est utilisé.

Conclusion

Pour cet article, pas vraiment de conclusion si ce n'est que nous vous souhaitons un bon été avec du repos pour les vacanciers, et des bonnes affaires pour les autres.

Sources

Page web d'Amit Singhal. <http://singhal.info/>

Modern Information Retrieval: A Brief Overview. Amit Singhal. IEEE Data Eng. Bull. 24(4):35-43 (2001). <http://act.buaa.edu.cn/hsun/IR2013/ref/mir.pdf>

Article du New York Times sur Amit Singhal :
<http://www.nytimes.com/2007/06/03/business/yourmoney/03google.html?pagewanted=1&r=0>

Jeff Dean chez Google : <http://research.google.com/people/jeff/>
<http://www.itworld.com/cloud-computing/399839/man-or-myth-3-million-google-engineer>

Jeff Dean Facts : <http://www.businessinsider.com/astounding-facts-about-googles-most-badass-engineer-jeff-dean-2012-1?op=1>

CV de Navneet Panda. <http://www.shmula.com/wp-content/uploads/2011/05/navneet-panda-google-pandaresume.pdf>

Page web de Navneet Panda : <http://alumni.cs.ucsb.edu/~panda/>

Biswanath Panda : <http://research.google.com/pubs/author38399.html>

Guillaume Peyronnet est gérant de **Nalrem Médias**. **Sylvain Peyronnet** est cofondateur et responsable des **ix-labs**, un laboratoire de recherche privé. Ensemble, ils font des formations, pour en savoir plus :
<http://www.peyronnet.eu/blog/>

Les services de Content Discovery (3ème partie)

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Nous sommes tous connectés à des réseaux sociaux comme Twitter ou Facebook, mais, parfois, leur utilisation à des fins de veille ou de community management est complexe, l'information étant souvent noyée dans la masse. C'est là qu'interviennent les outils de "content discovery" qui permettent, comme leur nom l'indique, de découvrir des contenus intéressants en traitant et filtrant les flux de données proposés par ces outils. Après l'exploration en profondeur de sept d'entre eux le mois dernier et le précédent, en voici deux nouveaux ce mois-ci, parmi les plus intéressants, pour clôturer cette série d'articles...

Depuis quelques années maintenant, l'animation de communautés sur le web ou « Community Management » (d'une page Facebook, d'un groupe LinkedIn, d'un compte Twitter) est devenue une profession à part entière. Si les interactions du community manager avec ses membres sous forme de réponses ou de discussions sont une des conditions de leur satisfaction, le fait d'apporter des contenus susceptibles de les intéresser sous forme d'articles, de vidéos ou de photos en est une autre. Encore faut-il identifier ce contenu. Les systèmes d'alertes par mots-clés de type Google Alertes jouent ce rôle depuis longtemps (pour d'autres solutions similaires voir ce tableau comparatif de 23 services : <http://socialcompare.com/fr/comparison/services-gratuits-d-alertes-par-mots-cles-1pp8mohd>) mais on voit arriver depuis quelques mois de nombreux autres services qui tentent d'« alimenter » le community manager (ou le veilleur bien sûr) en contenus pertinents en utilisant des techniques de repérage variées. Nous vous proposons d'en découvrir encore deux dans cette troisième partie de notre article sur les outils de "content discovery", après les sept déjà vus en mai et juin.

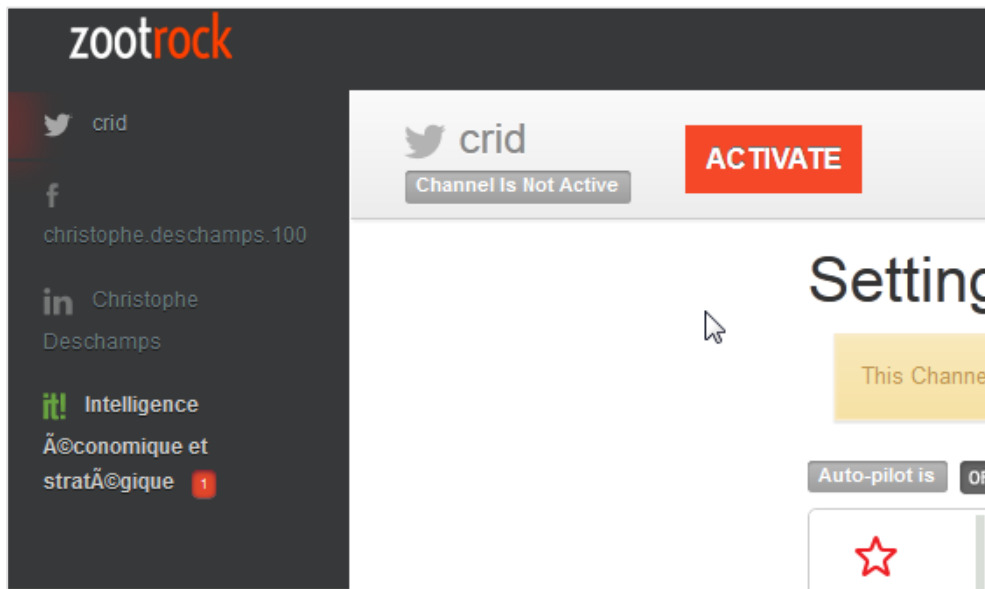
Après les services découverts durant ces deux derniers mois, il aurait pu sembler difficile d'en trouver d'autres pour un nouvel article sur le sujet. Ce n'est pourtant pas un problème tant il semble exister de l'effervescence autour de cette famille d'outils, le contenu étant plus que jamais le carburant du web social, et il n'est qu'à consulter la liste que met à jour Robin Good pour s'en convaincre (<http://content-discovery-tools.zeef.com/robin.good>).

ZootRock

Ce service (<http://www.zootrock.com/>) a été lancé en novembre 2013 avec pour objectif de permettre à chacun de récupérer des contenus provenant des réseaux sociaux auxquels il est connecté et de les rediffuser vers des pages de curation.

La première étape consiste à ouvrir ses profils sociaux à Zootrock afin de lui permettre d'y diffuser des contenus. Il est possible d'ajouter :

- Une page Facebook ;
- Un profil Facebook ;
- Un compte Twitter ;
- Un profil LinkedIn ;
- Une page LinkedIn ;
- Des pages Scoop.it ;
- Des flux RSS.



Une fois cela fait, le service vous propose d'affiner le profil des personnes qui vous suivent sur ces différents réseaux.

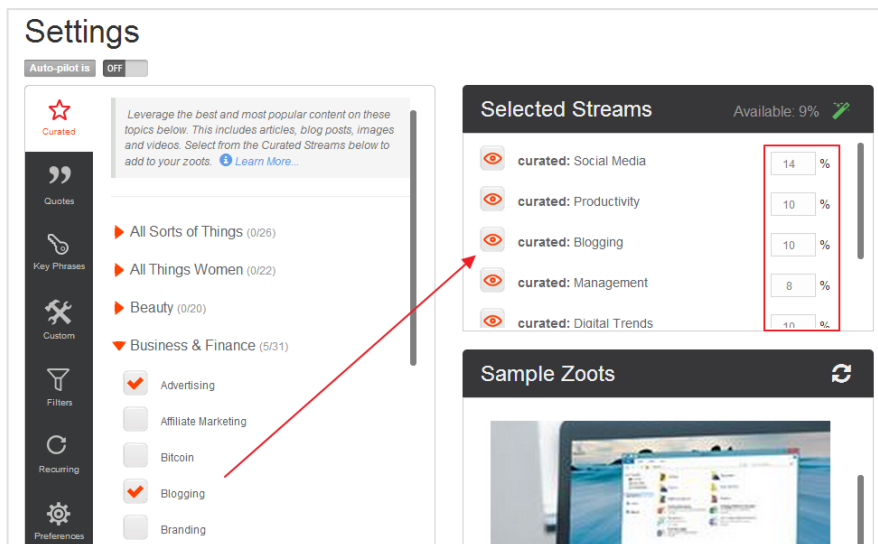
The image shows a form titled 'Who's Following You?' with a red person icon carrying a briefcase above the title. Below the title, there is a message: 'To ensure we find you the best possible content, please tell us a little more about you audience. (Just pick one for now, you can always add more variety later)'. The form contains two columns of checkboxes with labels: 'Artsy', 'Beauty Buff', 'Business' (checked), 'Entrepreneur', 'Explorer', 'Foodie', 'Funny', 'Health Conscious', 'Local Expert', 'Miscellaneous', 'Pop Culture Junkie', 'Sporty', 'Techie', 'The Intellectual', and 'Women'.

A noter que la version gratuite ne permet de diffuser que vers un seul des médias ci-dessus. Une fois cela fait, Zootrock vous propose des contenus issus de la presse anglo-saxonne ou de blogs et vous permet de les diffuser aisément vers les chaînes activées. Par exemple sur une page Scoop.it :

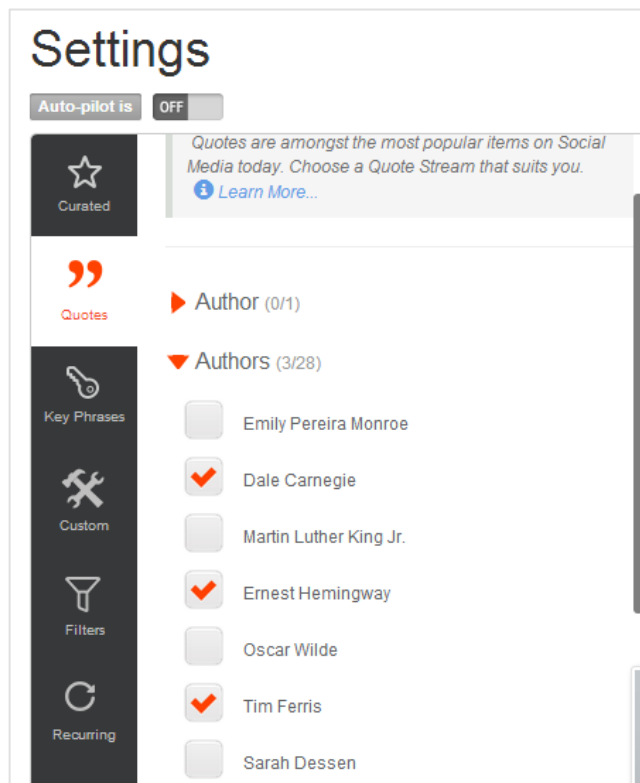


Il est possible d'améliorer les suggestions automatiques faites par l'outil chaîne par chaîne. Il faut pour cela en choisir une puis cliquer sur « Channel settings » en haut à droite de la page. On peut alors choisir d'être « alimenté » en :

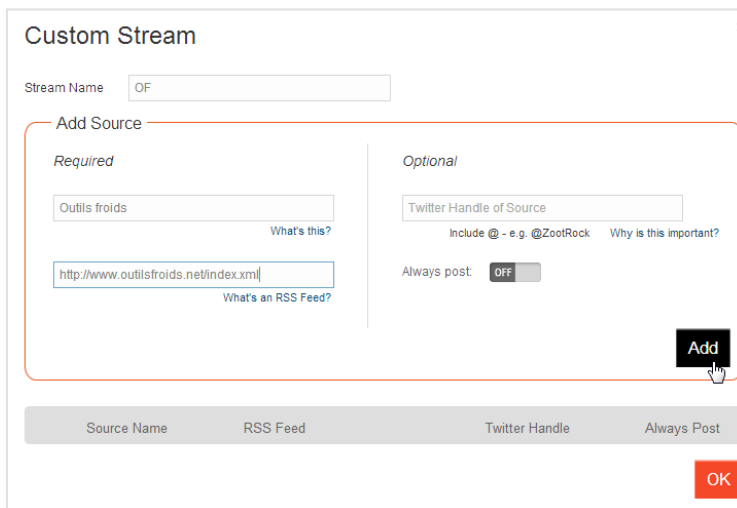
- actualités thématiques et attribuer un pourcentage de visibilité à chacune :



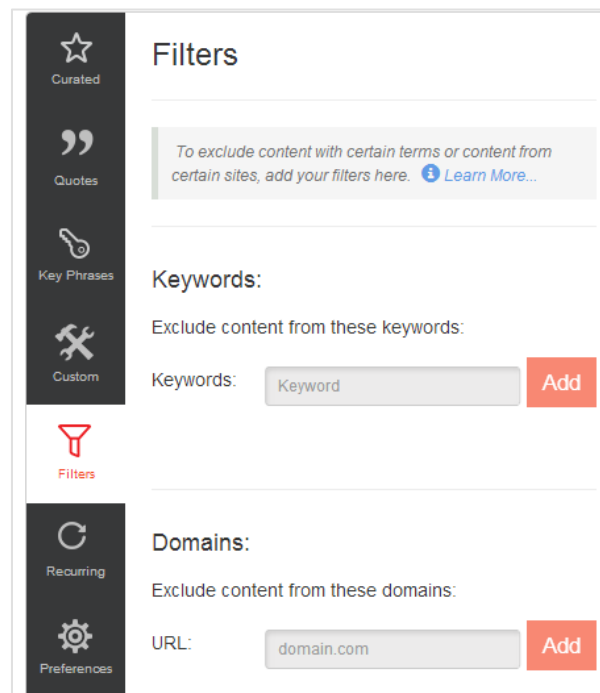
- en citations :



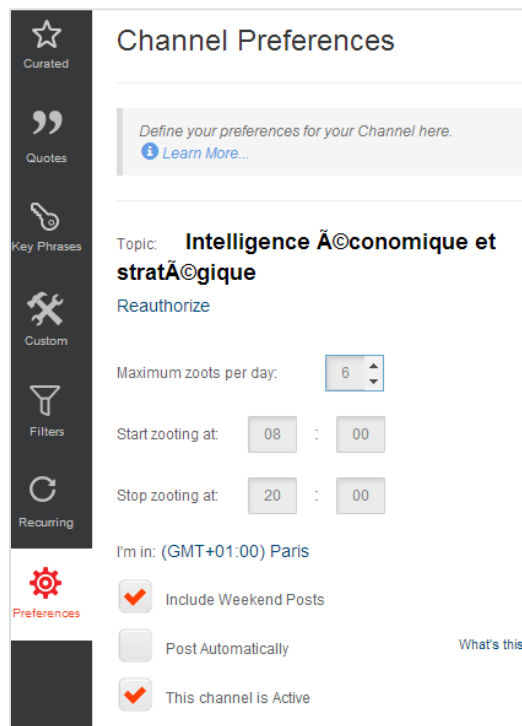
- en alertes par mots-clés dont les contenus seront apportés par Twitter ;
- en flux RSS que l'on configurera dans l'onglet « Custom ». Notez qu'il est possible de regrouper des flux par bouquets :



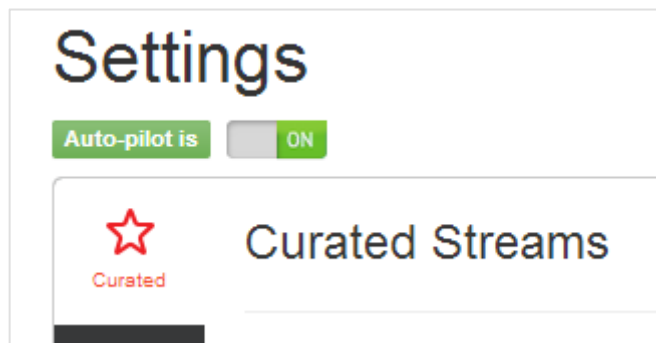
L'onglet « Filters » permet de supprimer les contenus contenant certains mots-clés ou issus de certains noms de domaine.



L'onglet « Recurring » permet d'ajouter des messages sur sa page profil et de les afficher de manière régulière. Enfin l'onglet « Preferences » permet de modifier plusieurs réglages quant au nombre de contenus publiés par jour sur une chaîne, les plages horaires ou encore le fuseau horaire.



A partir de là, Zootrock va vous proposer des contenus à partager. Si l'on fait suffisamment confiance aux réglages effectués et à ces propositions, il est même possible d'en automatiser complètement la diffusion.

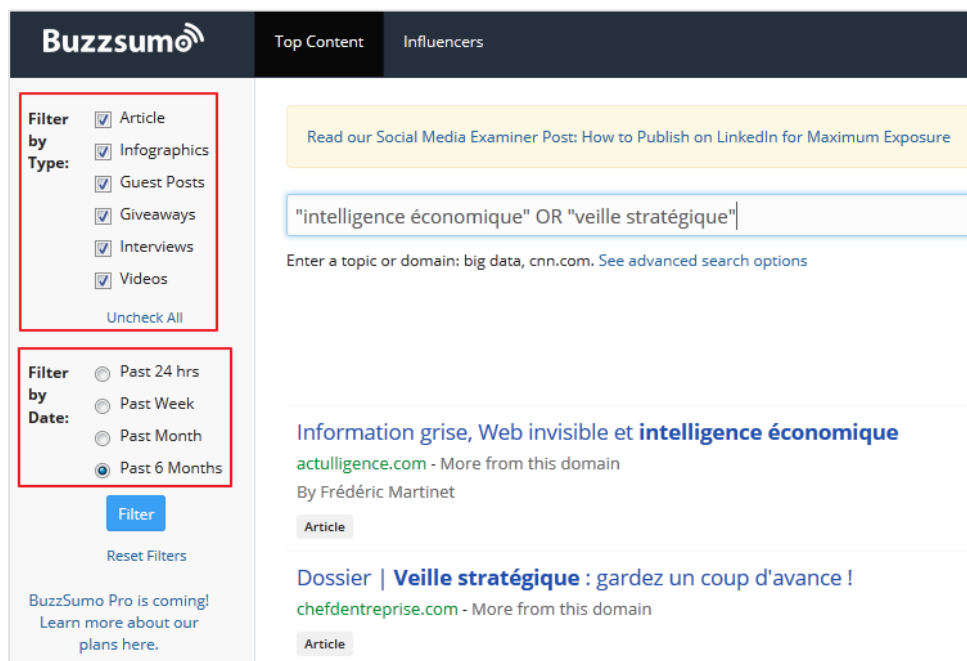


Au final, Zootrock est un service qui dispose donc de nombreuses fonctionnalités et qui se positionne comme un centre de réception/diffusion vers ses profils et pages sur les médias sociaux. On pourra toutefois regretter deux choses : d'une part un manque d'ergonomie qui entraîne pas mal de tâtonnements, d'autre part un système de pourcentage qui limite finalement la quantité de contenus à diffuser par thème, même s'il assure ainsi de limiter le temps passé dans le système. En gros, ne comptez pas sur Zootrock pour effectuer une veille exhaustive de vos thématiques.

Buzzsumo

Buzzsumo (<http://buzzsumo.com/>) est un service encore plus récent, puisqu'il a été lancé au début de l'année 2014. Il s'agit d'un autre service qui tire parti des réseaux sociaux pour faire remonter les contenus qui s'y partagent le plus. Étonnamment, on s'y connecte *via* un identifiant traditionnel et l'on peut ensuite ajouter un compte Twitter. Le service fonctionne ensuite comme un moteur de recherche que l'on interrogera à la façon d'un Google.

Première constatation, la langue française ne lui pose pas de problème :



La requête s'effectue sur plusieurs types de contenus et de périodes qu'il est possible de filtrer comme on le voit ci-dessus.

"intelligence économique" OR "veille stratégique"

Enter a topic or domain: big data, cnn.com. See advanced search options

Sort by: Total Shares Page 1 of 1

FACEBOOK SHARES LINKEDIN SHARES TWITTER SHARES PINTEREST SHARES GOOGLE+ SHARES TOTAL SHARES

Information grise, Web invisible et intelligence économique
actulligence.com - More from this domain
By Frédéric Martinet

View Sharers 42 44 127 0 15 228

Share

Article

Pour chaque contenu remonté, il est possible de voir le nombre de partages sur les principaux réseaux sociaux ainsi qu'un total de ces partages. C'est bien sûr à partir de ce total qu'est calculé le classement.

En cliquant sur le bouton « View sharers » on fait apparaître la liste de ceux qui, sur les réseaux sociaux, ont partagé le lien vers le contenu. Cette liste peut être utile pour découvrir des personnes influentes ou *a minima* intéressées par vos thématiques.

Filter by Type: Bloggers Influencers Companies Journalists Regular People

Ignore Broadcasters

Filter Reset Filters

Users that shared: "Information grise, Web invisible et intelligence économique" actulligence.com

Back

Sort by: Average Retweets Page 1 of 2

	PAGE AUTHORITY	DOMAIN AUTHORITY	FOLLOWERS	RETWEET RATIO	REPLY RATIO	AVERAGE RETWEETS
<p>Michèle Battisti @mbattisti64 bit.ly/6F6dbH Veille sur le droit de l'information. Blog Paralipomènes</p>	19	63	2,772	81%	2%	1.5
<p>Aref JDEY @demaïnlaveille demaïnlaveille.com Digital Trends & Transformation, Information Management, #SocialTech, Usability http://t.co/NQJPu6HEaG</p>	28	14	9,222	29%	8%	1.4
<p>David FERRON @innouvelle about.me/dferron #Innovation en Nord-Pas de Calais... et bien au-delà ! #compétitivité #design #3pod #madeinfrance #startup #prospective #E #CCI #Lille #npdc</p>	47	91	2,637	8%	2%	1.3

On constate par ailleurs qu'il est possible, en haut à gauche, de n'afficher que certains groupes d'acteurs : bloggeurs, journalistes, influenceurs, entreprises...

C'est une force de ce service que de « travailler » autant sur les contenus que sur ceux qui les partagent et on peut d'ailleurs démarrer une requête directement dans le volet « influenceurs ».

Buzzsumo Top Content **Influencers** API Help Email Us Blog Account

Read our Social Media Examiner Post: How to Publish on LinkedIn for Maximum Exposure

"intelligence économique"

Enter a topic or username: @buzzsumo, big data OR marketing. See advanced search options

Sort by: Relevancy Page

	PAGE AUTHORITY	DOMAIN AUTHORITY	FOLLOWERS	RETWEET RATIO	REPLY RATIO	AVERAGE RETWEETS
<p>Mugabowindekw Robert @mugrobert jambaasbl.com Jambo - Intelligence économique</p>	31	21	81	0%	0%	0.0
<p>Mohamed TOUMI @moh_toumi Consultant en intelligence économique</p>	0	0	2	0%	0%	0.0
<p>Farouk @Lfk vizify.com/farouk-1/twitter-video</p>	0	0	82	15%	0%	0.0

Nous pouvons ici découvrir des personnes influentes sur n'importe quelle thématique ajoutée au moteur et les suivre sur Twitter si l'on a connecté son compte à Buzzsumo, ou encore l'ajouter à une liste twitter.

Profile	Followers	Following	Tweets	Retweets
Alain KHEMILI @alain_khemili Responsable Département Industrie, Innovation et Intelligence économique, CCI France	62	73	247	23%
Zamogo Tuo @zamogot Intelligence économique - Veille - Webmarketing - Journalisme web - Marketing & Événementiel Sportifs	0	0	20	0%
F. JEANNE-BEYLOT @troover Gérant-fondateur Troover - #Veille & #Intelligence économique sur Internet Directeur associé @InMediatic - notoriété réputation #influence en ligne	17	753	30%	
Farid Tamsamani @faridtemsamani Consultant en intelligence économique spécialisé sur l'Afrique subsaharienne et stratégie politique. #Sevran #AgirPourSevran #UDI	0	0	168	0%

Qu'il s'agisse de personnes ou de contenus il est possible d'enregistrer une requête que l'on pourra relancer au coup par coup.

Filter Article
by Infographics
Type: Guest Posts
 Giveaways
 Interviews
 Videos
Uncheck All

Filter by Date: Past 24 hrs
 Past Week
 Past Month
 Past 6 Months
Filter
Reset Filters

Read our Social Media Examiner Post: How to Publish

"intelligence économique" OR "veille stratég

Enter a topic or domain: big data, cnn.com. See advance

Information grise, Web invisible et intel
actulligence.com - More from this domain
By Frédéric Martinet
Article

Dossier | Veille stratégique : gardez u
chefdentreprise.com - More from this domain
Article

Intelligence économique: un cyberpré
challenges.fr - More from this domain
By Par Challenges.fr Voir tous ses articles

Saved Searches - Edit
Run a saved search...
Run a saved search...
IE - Veille strat
Learn more about our plans here.

Le vrai "plus" serait bien sûr de pouvoir transformer cette requête en une alerte par email. Des fonctionnalités *premium* sont annoncées pour bientôt et il y a fort à parier que cette possibilité en fasse partie. C'est vraiment le seul manque d'un service simple et très efficace selon nous.

Conclusion

Ce tour d'horizon de ces services depuis trois numéros nous a amené à découvrir des services qui, s'ils ont tous pour point commun la découverte de contenus, fonctionnent souvent selon des modalités très différentes. Là où certains utilisent les médias et blogs connus comme sources, d'autres préfèrent s'intéresser à ce qui se partage sur les réseaux sociaux, d'autres encore se focalisent sur ce que partagent vos propres amis et contacts sur ces mêmes réseaux, au risque de limiter vos possibilités de découvertes originales mais en vous assurant ainsi des contenus qui rentreront dans le « scope » de vos intérêts. D'autres enfin sont un joyeux mix de tout cela.

Les nombreuses passerelles avec les services de partage et de curation assurent selon nous à ces services une durabilité évidente. Nous ne parlons pas forcément des 9 que nous avons testé (même si certains sont très bons) mais plutôt de cette famille d'outils qui permet en quelque sorte de boucler la boucle de production/diffusion/consommation d'informations par tous, pour tous. Une promesse tenue du web 2.0.

Christophe Deschamps

Consultant et formateur en gestion de l'information.

Responsable du blog Outils Froids (<http://www.outilsfroids.net/>)

Comment bien utiliser l'outil de planification des mots clés de Google ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Depuis le 6 Septembre 2014, Google a définitivement remplacé son « outil de génération de mots clés » par l'« outil de planification des mots clés » (ou « Google Keyword Planner »). Toujours utile pour le référencement naturel, l'outil a cependant fortement évolué. Dans cet article, nous allons aborder toutes les techniques pour réellement tirer parti de cet outil en SEO, notamment la fusion des listes de mots clés, le ciblage géographique ou encore le traitement et l'analyse des données fournies par Google....

L'outil de génération de mots clés de Google a toujours été utile pour savoir combien de fois un mot clé était recherché sur Google. Premier problème du nouvel outil (par rapport à l'ancien), disponible à l'adresse <https://adwords.google.fr/KeywordPlanner> : vous aurez besoin obligatoirement d'un compte Adwords (même inactif) pour y accéder. De même, l'interface a changé et il faudra vous y adapter.

Voyons maintenant ce qu'il est possible de faire avec cet outil dans le cadre du référencement naturel de votre site. Vous trouverez en effet 4 fonctionnalités différentes :

- la recherche de mots clés (ce qui correspond à l'ancien générateur de mots clés) ;
- l'obtention de volumes de recherche pour une liste de mots clés ;
- la prévision de trafic pour une liste de mots clés ;
- un outil pour fusionner des listes de mots clés.

Outil de planification des mots clés

Planifier votre prochaine campagne sur le Réseau de Recherche

Que souhaitez-vous faire ?

- ▶ Rechercher de nouvelles idées de mots clés et de groupes d'annonces
- ▶ Obtenir les statistiques relatives au volume de recherche d'une liste de mots clés ou les répartir par groupes d'annonces
- ▶ Obtenir des prévisions de trafic pour une liste de mots clés
- ▶ Multiplier les listes de mots clés pour obtenir de nouvelles idées de mots clés

Pour être utile en référencement ou dans un audit, il faut en effet respecter quelques règles de base que nous allons détailler tout au long de cet article :

- avoir une liste prédéfinie de mots clés (fournie par le client ou une analyse marketing) ;
- savoir correctement cibler vos recherches ;
- bien utiliser et savoir analyser les données affichées lors la recherche de mots clés ;

- comprendre l'intérêt de l'outil de fusion des mots clés, surtout pour les variantes d'expressions ou pour le référencement local.

C'est par ce dernier élément que nous commencerons. Mais juste avant, il est impératif de comprendre le système de ciblage et de filtres de l'ensemble des 4 outils.

Le ciblage des données

Contrairement à la précédente version de l'outil, l'outil de planification des mots clés va toujours vous montrer les résultats pour la zone géographique que vous ciblez, et non plus les résultats globaux.

La clé pour votre référencement est donc de correctement cibler les données qui seront affichées. Chacun des quatre outils vous donne en effet accès à un nombre plus ou moins restreint de filtres. En tout, vous pouvez filtrer :

- par zone géographique ;
- par langue ;
- par source (soit le moteur de recherche, soit le moteur de recherche et son réseau display Adsense) ;
- par volume de recherche ;
- par CPC, concurrence et/ou taux d'impression dans Adwords ;
- par mot clé à exclusion.

Pensez donc toujours à ne cibler que les zones qui vous intéressent, par exemple en français si votre site est en français ou encore « Belgique » si vous ne ciblez que ce pays-là.

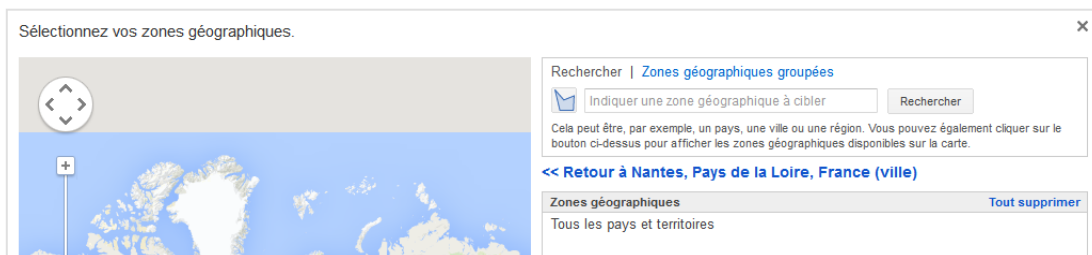
Le ciblage géographique

Dans cet outil, l'outil de ciblage géographique a été amélioré. Vous pouvez ainsi cibler par pays, par région, par département et par ville, ou encore un mix des 4 (par exemple cibler la France, Londres et le sud de l'Espagne en même temps). Vous allez aussi pouvoir cibler les données dans un rayon autour d'un lieu précis.

Pour cela, cliquez sur le ciblage géographique, puis cliquez sur « Recherche Avancée » :

The screenshot shows the 'Ciblage' (Targeting) section of the Google AdWords interface. At the top, there are two tabs: 'Ciblage' and 'Personnaliser votre recherche'. The 'Ciblage' tab is active. Below it, there is a section for 'Zones géographiques' (Geographic zones) with a dropdown menu currently set to 'Tous les pays et territoires' (All countries and territories). To the right of this dropdown is a 'Tout supprimer' (Remove all) button. Below the dropdown is a text input field labeled 'Indiquer une zone géographique à cibler' (Specify a geographic zone to target) and a blue link for 'Recherche avancée' (Advanced search). Below this input field is a small instruction: 'Par exemple, un pays, une ville ou une région.' (For example, a country, a city or a region). Below the 'Zones géographiques' section is the 'Période' (Period) section, which includes a button to 'Afficher le nombre moy. de recherches mensuelles pour : Les 12 derniers mois' (Show the average number of monthly searches for: Last 12 months). To the right of this is the 'mon compte' (My account) section, which includes a button to 'Masquer les mots clés dans mon plan' (Hide keywords in my plan) and a button for 'Inclure/Exclure' (Include/Exclude).

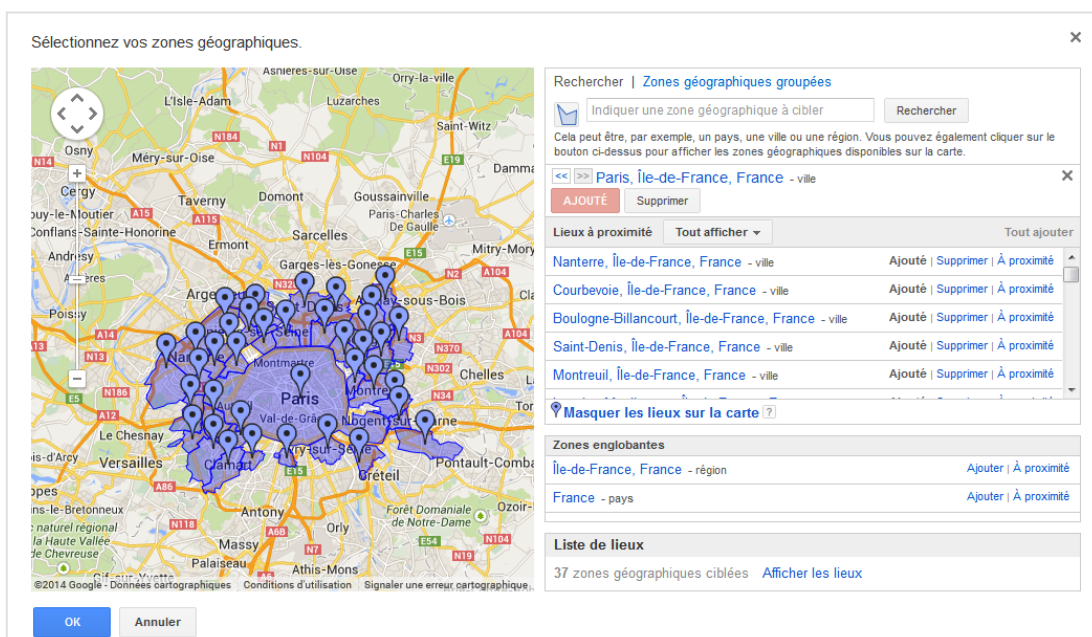
Par défaut, on voit que « Tous les pays et territoires » sont sélectionnés. Cliquez donc avant tout sur « Tout supprimer » pour démarrer votre ciblage, sinon cela ne servira à rien.



Utilisez alors le barre de recherche pour cibler vos résultats. Lors de chaque recherche, Google vous proposera alors les résultats directs et les zones géographiques qui contiennent votre recherche :



Méfiez-vous cependant des liens « A proximité » qui n'affichent parfois pas toute la totalité des zones autour de votre cible. Par exemple, si l'on prend « A proximité de Paris » et que l'on ajoute tous les résultats, nous aurons quelques blancs :



Pensez donc toujours à bien vérifier que votre zone ciblée est complète.

Dans l'onglet « Zones géographiques », sachez que l'on peut ajouter de grandes zones directement. Vous devrez d'abord entrer l'identifiant du code pays (FR pour le France par

exemple) avant de pouvoir chercher ces zones. Vous en trouverez la liste par pays et par type, ici : <https://developers.google.com/adwords/api/docs/appendix/geotargeting>.

Attention à ne pas trop cibler non plus, car une partie des outils qui l'on va détailler ensuite est limitée et ne vous affichera que les statistiques pour les 10 premières zones géographiques sélectionnées.

Une fois que vous aurez ajouté toutes les villes et zones désirées, cliquez simplement sur OK.

L'outil de multiplication de listes de mots

Voyons maintenant en détail le premier outil proposé.

Préalable : Avant d'expliquer en détail de quoi il retourne, nous partons du principe que vous avez déjà une liste de mots clés fournie par une analyse des concurrents de votre secteur d'activité et par votre client (que ce soit le directeur, des vendeurs, des techniciens, des personnes du service après-vente...). Plus votre liste sera exhaustive, plus les données que vous allez pouvoir traiter seront utiles et pertinentes par la suite.

Fonctionnement :

Lorsque vous commencerez à utiliser cet outil, démarrez toujours par l'outil de « multiplication des mots clés » car elle va permettre d'élargir votre base initiale d'expressions, en évitant que l'outil de recherche n'en oublie certains. L'outil de multiplication permet en effet de mélanger deux groupes de mots clés distincts.

Imaginons que vous vouliez coupler des types de produits à des caractéristiques, sans tout taper à la main : c'est exactement ce que cet outil va permettre de faire. Par exemple, vous vendez des camions, des voitures et des vélos. Chacun de ces 3 produits est vendu sur Nantes, Paris et Strasbourg. En entrant ces 2 groupes de mots clés dans l'outil, on pourra immédiatement fusionner les deux listes et obtenir les données qui y sont associées :

▼ Multiplier les listes de mots clés pour obtenir de nouvelles idées de mots clés

Liste 1

camion
voiture
vélo

Liste 2

nantes
paris
strasbourg

Ciblage ?

Tous les pays et territoires

Google

Mots clés à exclure

Obtenir des estimations

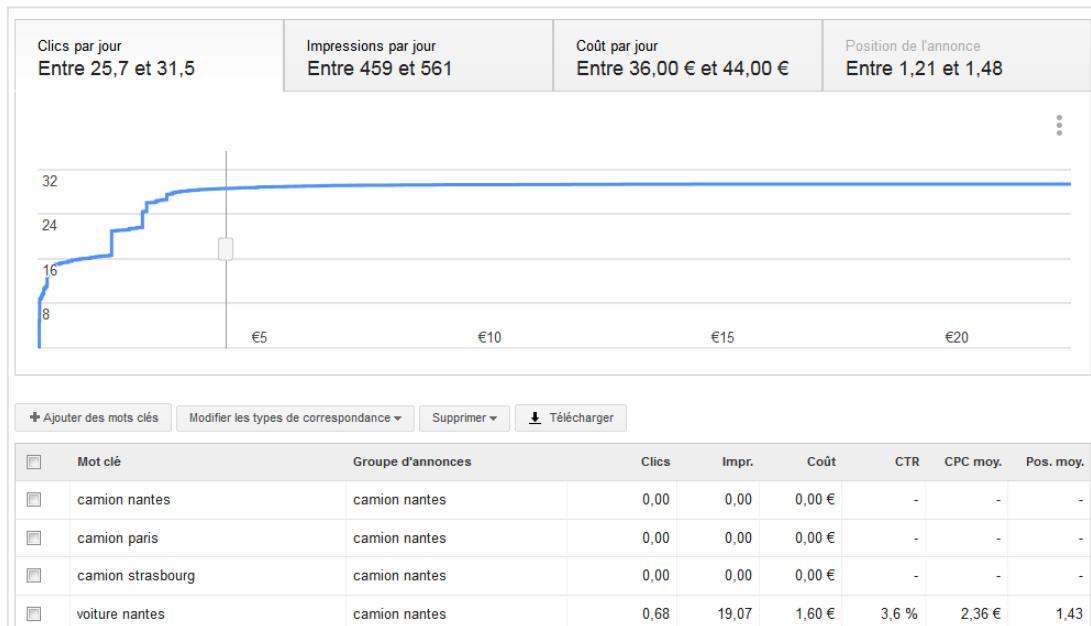
Obtenir le volume de recherche

Créez des combinaisons de mots clés et obtenez des statistiques relatives aux volumes de recherche ou des prévisions de trafic.

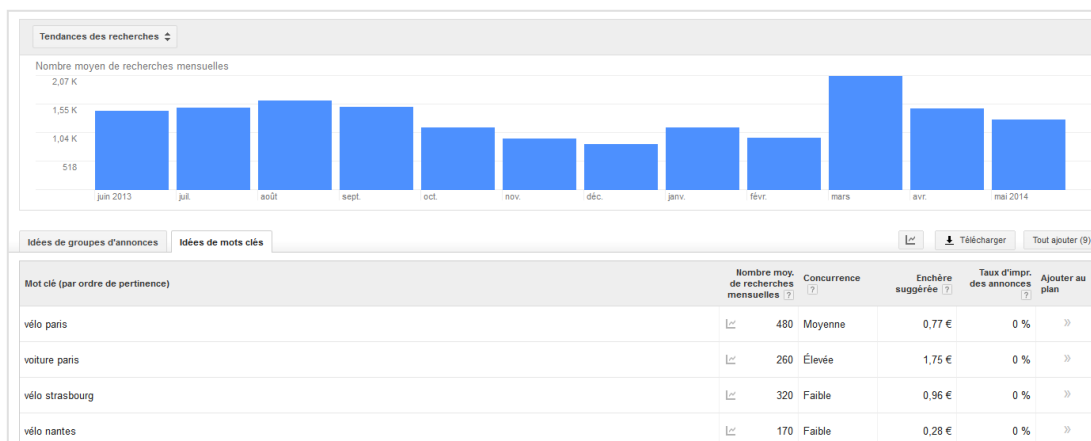
Une fois le ciblage déterminé en bas à gauche, vous pourrez alors cliquer sur deux options différentes. Sachez juste que, par défaut, les deux options vous montreront des résultats globaux, c'est à dire des volumes ou des estimations de trafic pour l'ensemble

des mots clés générés. N'oubliez donc pas de cliquer sur les onglets « mots clés » pour avoir le détail de chaque mot clé ou expressions. Voici les deux options concernées :

- Obtenir des estimations de clics pour Adwords (pas utile en référencement naturel) :



- Obtenir le volume de recherche du mots clés, c'est à dire connaître le trafic actuel moyen de chacun des mots clés :



Cette deuxième option vous affichera en haut de page la saisonnalité de l'ensemble des mots clés, puis pour chaque ligne suivante le volume moyen de recherches par mois sur la zone ciblée. L'interface étant identique à celle de l'outil « rechercher des mots clés », nous en reparlerons plus loin.

En tant que référencier, cet outil va avoir comme intérêt majeur de pouvoir étoffer immédiatement votre liste de mots clés. Imaginons que votre client ou votre site vende des milliers de références produits dans des boutiques partout en France : vous pourriez très bien fusionner ces types de produits avec des centaines de localisations géographiques. Vous pourriez aussi fusionner des formats ou attributs à certains produits (« rouge », « pas cher », « xxl »...).

Mieux encore, si vous référenciez le site d'une entreprise implantée à un seul endroit, vous allez pouvoir ainsi générer facilement la liste de l'ensemble des produits et prestations de

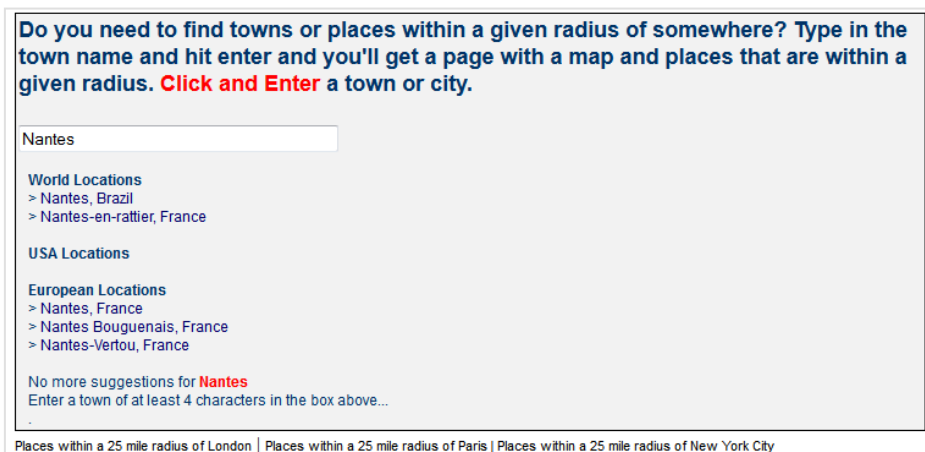
cette société avec l'ensemble des villes situées dans un rayon de X kilomètres (par exemple « voiture paris », « voiture versailles », « voiture nanterre »...).

Trouver une liste de villes à proximité d'un lieu

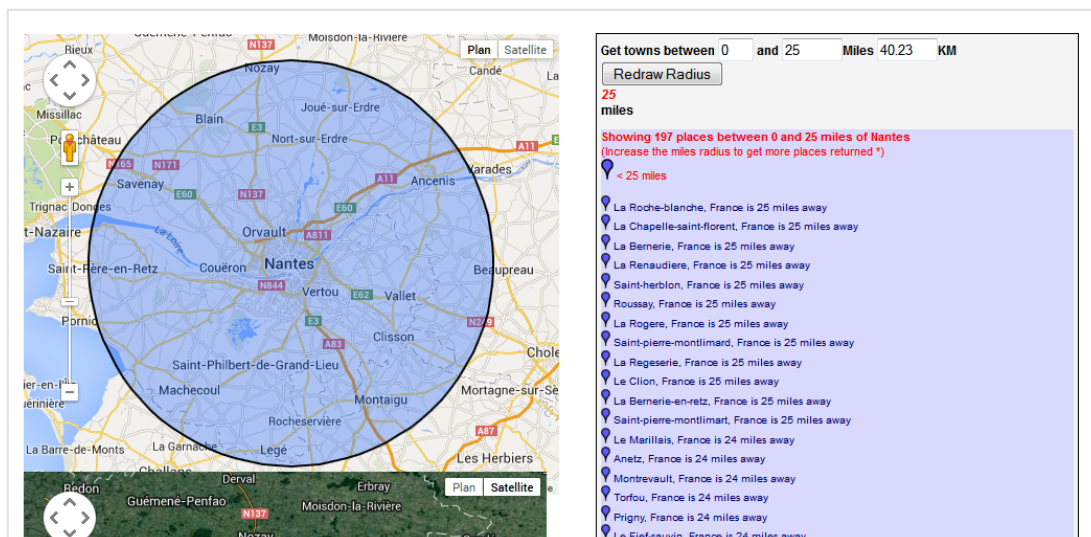
Pour trouver la liste de villes situées dans un rayon prédéfini autour d'une autre ville pour l'utiliser dans l'outil de multiplication des mots clés, rendez-vous sur :

<http://www.distance-calculator.co.uk/towns-within-a-radius-of.php>

Cliquez sur « Click and Enter », entrez le nom de la ville souhaitée et tapez « Entrée ».

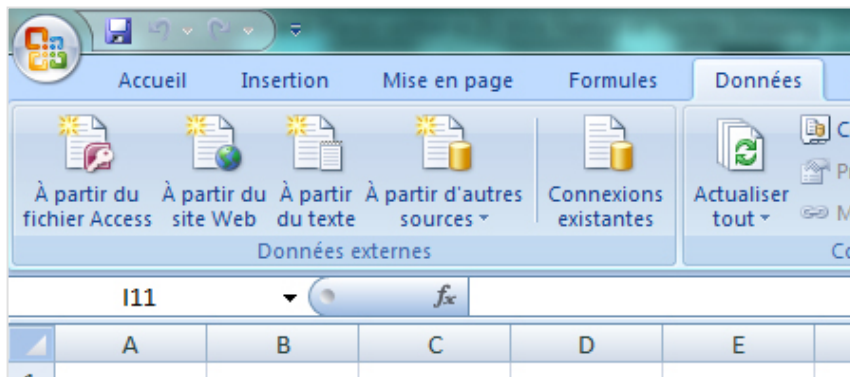


Ensuite, en haut à droite de la liste des villes vous pourrez choisir le rayon désiré.

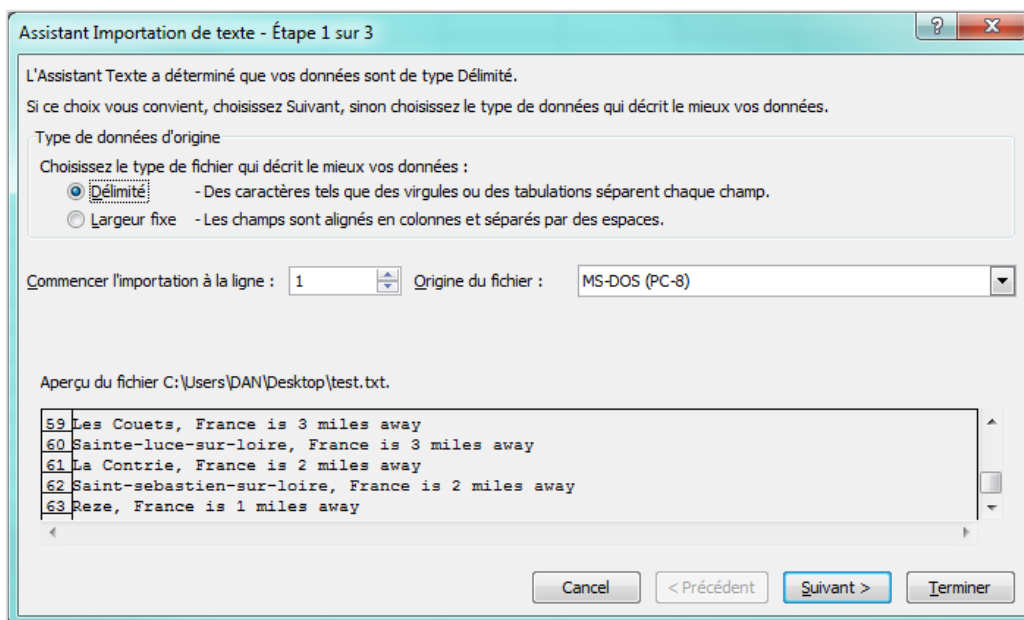


Maintenant, nous allons devoir traiter les données :

- copiez-collez la liste des villes dans un fichier txt ;
- ouvrez ensuite Microsoft Excel ;
- rendez-vous dans « Données => A partir du texte » ;



- dans le menu suivant, choisissez « Délimité » ;



- dans le menu suivant, choisissez « virgule » puis cliquez en bas à droite sur « Terminer ».

Vous devriez avoir alors la liste de toutes les villes ciblées dans la première colonne.

	A	
1	Casson	Fr
2	Le Cellier	Fr
3	Le Landreau	Fr
4	Le Pallet	Fr
5	Montbert	Fr
6	Saint-etienne-de-montluc	Fr
7	Curete	Fr
8	Grandchamp-des-fontaines	Fr
9	La Chapelle-basse-mer	Fr
10	La Paquelais	Fr
11	Maisdon	Fr
12	Maisdon-sur-sevre	Fr
13	Monnieres	Fr
14	Port-saint-pere	Fr
15	Saint-mars-de-coutais	Fr
16	Vigneux-de-bretagne	Fr
17	La Chapelle-heulin	Fr
18	La Pepiere	Fr
19	Le Loroux-bottereau	Fr
20	Saint-leger-les-vignes	Fr
21	Bouche	Fr
22	Brains	Fr
23	Chateauthebaud	Fr
24	Coueron	Fr
25	La Chevroliere	Fr
26	La Poivetiere	Fr
27	La Sanglaire	Fr

Libre à vous ensuite de l'utiliser pour générer de nouvelles idées de mots clés, surtout pour votre référencement local.

Remarque : attention, la liste des villes n'est pas parfaite. L'outil ajoute en effet des tirets qui ne sont parfois pas nécessaire. Par exemple, la ville « Saint Jean de Boiseau » est indiquée « Saint-jean-de-boiseau ». Pensez donc à remplacer les tirets par des espaces. N'oubliez pas non plus de réécrire les variantes de villes, notamment « Saint » que les gens peuvent avoir l'habitude d'écrire « St ».

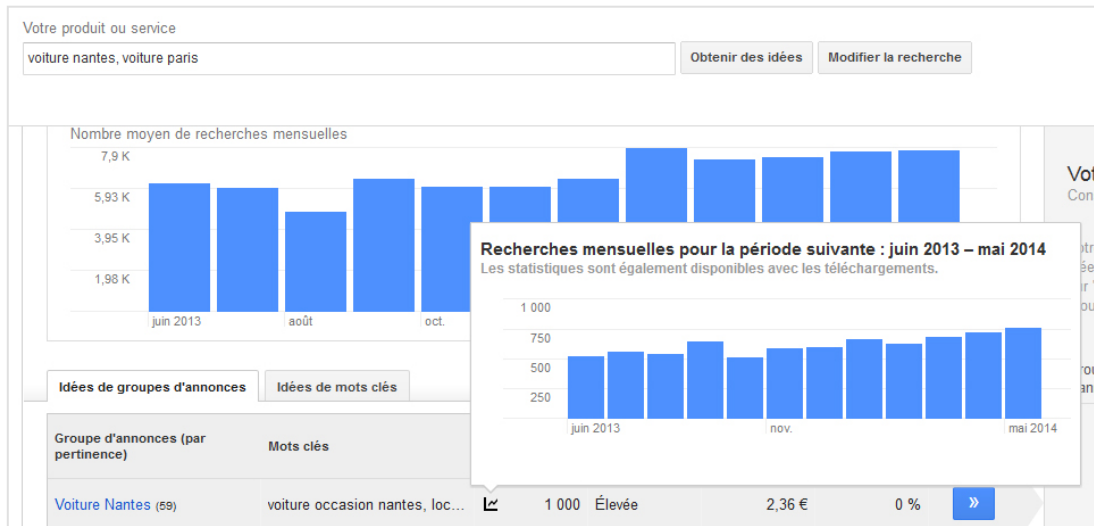
L'outil de Recherche des mots clés

Comme expliqué précédemment, l'interface du bouton « Volume de recherche » est identique ou presque à l'outil « Rechercher des mots clés ».

Voici comment s'en servir à bon escient dans une stratégie ou un audit SEO : cet outil permet de trouver de nouvelles idées de mots clés à partir d'une liste déjà existante, d'une page web ou d'une catégorie de « produit » (par thématique, en somme). Et pour l'ensemble de vos mots clés et des mots clés suggérés, l'outil vous donnera un volume de recherche mensuel moyen.

Le fonctionnement est simple : renseignez soit un ou plusieurs mots clés, soit une adresse web. Ensuite, paramétrez impérativement les différents filtres en dessous pour affiner vos résultats en fonction de votre zone de chalandise ou de vos objectifs. N'oubliez donc pas de correctement définir la langue et la zone géographique souhaitée.

Il faut savoir impérativement plusieurs choses : la première est que le nombre moyen de recherches mensuelles a été lissé sur les 12 derniers mois. Vous ne voyons donc pas ici la saisonnalité du mot clé ou du groupe d'annonce, et il faudra survoler le petit icône en forme de graphique de chaque ligne pour la connaître :



Sachez également qu'il y aura toujours un décalage entre les données affichées et les données actuelles. Par exemple, les captures d'écran réalisées pour cet article ne nous montrent les statistiques que jusqu'à mai 2014, alors que nous sommes déjà en Juillet. Si vous souhaitez avoir une analyse plus récente, utilisez plutôt l'outil Google Trends : <http://www.google.fr/trends/>

Ensuite, il faut bien comprendre les données du tableau :

- le taux d'impression correspond au taux d'impression de vos annonces pour ce mot clé dans Adwords. Sans campagne, il sera donc toujours à 0% ;
- l'enchère suggérée est souvent fautive voire aberrante. Elle affiche parfois des mots clés de niche à plus de 100€ du clic. Il vaut donc mieux ne pas utiliser cette donnée dans vos analyses ;
- il n'est plus possible de différencier les résultats entre les ordinateurs de bureau, les tablettes et les téléphones ;
- dans le tableau, le volume de recherche est affiché avec un ciblage « Exact ».

Rappels sur le ciblage

Petit rappel obligatoire sur le ciblage des mots clés, car cela peut fausser réellement toute votre analyse, sachant que Google a modifié sa façon de considérer le ciblage le plus précis. Comme expliqué, les volumes de recherche affichés dans le tableau sont faits en mode « Exact ». Il existe en réalité deux autres modes : le ciblage par « expression » et le ciblage « large ».

- **Exact** : les volumes de recherche correspondent à l'expression tapée exactement de cette façon, ainsi que ses variantes proches (sigles, pluriels, fautes d'orthographe...) ;
- **Par expression** : les volumes correspondent à toutes les recherches contenant l'expression ciblée dans le même ordre. Les mots clés apparaîtront dans l'interface Adwords sous la forme "mot clé" ;
- **Exact** : les volumes correspondent à toutes les recherches contenant l'expression ciblée, dans n'importe quel ordre et avec les synonymes. Les mots clés apparaîtront sous la forme [mot clé].

Prenons un exemple issu du centre d'aide Adwords.

- Ciblage exact « chapeaux femmes » : les données vont inclure « chapeaux femmes » et « chapeau femme ».
- Ciblage par expression « chapeaux femmes » : les données vont inclure « chapeaux femmes », « chapeau femme », « achat chapeau femme »...

- Ciblage large « chapeaux femmes » : les données vont inclure « chapeaux femmes », « chapeau femme », « achat chapeau femme », « femme achat chapeau », « achat casquette femme »,...

Les données qui s'affichent au centre dans le tableau sont au format « Exact », donc le plus proche du mot clé. Attention : cela veut donc dire que sont éliminés les résultats qu'une recherche par expression nous aurait donnés, et qui auraient pu être intéressants. Il faut donc bien avoir conscience que la liste de mots clés de base que l'on teste dans l'outil pour trouver de nouveaux mots clés doit être la plus exhaustive possible.

Dans l'onglet « Idées de mots clés », vous verrez deux tableaux l'un en dessous de l'autre : le premier ne contiendra que les mots clés que vous avez cherchés, tandis que le second est celui qui contient toutes les expressions suggérées. D'ailleurs, n'hésitez pas à augmenter le nombre de mots clés par page avec la petite flèche déroulante pour ne pas avoir à naviguer entre des dizaines de pages différentes.



Exporter et traiter les données

Pour utiliser ensuite ces mots clés dans Adwords ou pour les exporter pour un autre outil, vous pouvez utiliser les boutons disponibles au bout de chaque ligne « Ajout au plan » ou le bouton « Tout ajouter ». Vous verrez alors que vos données seront ajoutées dans la colonne de droite.

Nous conseillons en général de commencer à ajouter manuellement tous les mots clés que vous avez tapés, c'est-à-dire ceux du premier tableau. Vous verrez ainsi à droite un groupe d'annonce se créer « Mes idées de mots clés ». Cela vous permettra de distinguer vos mots clés de ceux suggérés par l'outil.

Faites très attention : dès que vous ajouterez un mot clé ou un groupe de mots clés ici, le comportement par défaut de Google va changer. En effet, tout ce que vous retrouverez dans la colonne de droite contiendra des données « larges » et non plus au format « exact ». Pensez donc à bien cliquer sur le bouton en forme de crayon « définir le type de correspondance » pour choisir « exact ».

Votre plan
Conservé jusqu'au 14 juil. 2014

Plage d'enchères [?](#)
Entre 0,00 € et 22,08 €

Estimations quotidiennes
Entre 0 et 60,2 Clics
Entre 0,00 € et 377 € Coût

Groupes d'annonces : 54

- Voiture (66)**
- De Voiture (6)
- Nantes (7)
- Voiture Paris (25)
- Location Voiture (141)
- Sans Permis (5)
- Louer (6)
- Louer Utilitaire (5)
- Auto Occasion (8)

[Consulter les estimations](#)

Vous pourrez alors cliquer sur « Consulter les estimations » pour télécharger le tableau complet de données.

Téléchargez votre plan
Groupes d'annonces : 54, mots clés : 802

Statistiques historiques Inclure le nombre moyen de recherches mensuelles, le niveau de concurrence et d'autres statistiques
 Segmenter par mois

Estimations du trafic Inclure le nombre de clics, le nombre d'impressions et d'autres estimations
Vous ne pouvez pas télécharger d'estimations avant d'avoir défini une enchère.

Format de fichier CSV (AdWords Editor)
Vous pourrez importer ce format dans AdWords sous forme de feuille de calcul. [En savoir plus](#)
 CSV (Excel)

[Télécharger](#) [Annuler](#)

Remarques

Pour chaque recherche, l'outil ne vous affichera que 800 résultats au total, ce qui peut parfois limiter le nombre de mots clés suggérés.

Sachez aussi que pour Google, un mot clé avec accent est considéré comme différent du même mot clé sans accent (ce qui est d'ailleurs logique car les résultats dans le moteur de recherche changent aussi). Pour tous vos mots clés avec accent, vous devrez tester les

variantes sans. Si par exemple vous vendez des vélos, il faudra tester le mot clé « velo » et le mot clé « vélo ».

Il faut également savoir que l'exportation de vos données est limitée à 3 000 mots clés par fichier CSV. Cette limite peut être rapidement atteinte. Pensez donc parfois à exporter une première liste de mots clés, puis en refaire une seconde. Il sera alors toujours temps de supprimer les doublons avec Excel.

L'outil de volumes de recherche

Cet outil ne vous proposera pas de suggestions de nouveaux mots clés, mais uniquement les volumes de recherche pour une liste donnée. On pourrait se poser la question de l'intérêt de cet outil par rapport au précédent. En fait, il y en a principalement un : vous n'aurez plus la limite de 800 mots clés, et vous pourrez donc récupérer le volume de trafic de plusieurs milliers de mots clés d'un seul coup.

Si vous désirez des prévisions de trafic, sachez que vous pouvez entrer directement vos mots clés avec le ciblage désiré. Par exemple :

- "voiture nantes" : ciblage par expression ;
- [voiture nantes] : ciblage exact ;
- voiture nantes : ciblage large.

L'outil de prévision de trafic

Nous ne nous attarderons pas sur cet outil car il ne servira qu'à gérer vos prévisions de trafic sur Adwords, et ne sera donc d'aucune utilité pour votre référencement naturel.

Que faire avec ces données ?

Une fois que vous avez utilisé tous ces outils, vous devriez avoir une longue liste de mots clés pertinents liés à votre activité avec le volume de recherche associé. Ces mots clés - vous pouvez théoriquement en avoir plusieurs centaines voire plusieurs milliers - sont normalement disponibles dans un ou plusieurs fichiers csv.

Etape 1 : importer toutes vos données

Commencez par ouvrir Excel, puis importez tous vos fichiers avec « Données > A partir du texte » dans le même onglet. Si vous avez plusieurs fichiers, pensez à supprimer la ou les premières lignes de légendes à chaque nouvel import.

Vous devriez avoir ce type de données :

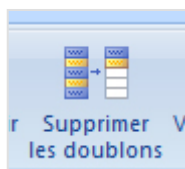
Ad Group	Keyword	Keyword Type	Currency	Segmentation	Avg. Monthly Searches (exact match only)	Competition	Suggested bid	Impr. Share
			EUR	Total - Tablet Only	1370			
			EUR	Total - Desktop Only	8028			
			EUR	Total - Mobile Only	1071			
			EUR	Total - Treillieres				
			EUR	Total - La Chapelle-sur-Erdre	553			
			EUR	Total - Saint-Sebastien-sur-Loire	832			
			EUR	Total - Bouguenais	462			
			EUR	Total - Reze	2719			
			EUR	Total - Nantes				
			EUR	Total - Basse-Goulaine	295			
			EUR	Total - Vertou	727			
			EUR	Total - Saint-Herblain	1906			
			EUR	Total - Carquefou	1663			
			EUR	Total - Sainte-Luce-sur-Loire	356			
			EUR	Total - Orvault	951			
17	Mes Idées De Mots Clés camion nantes	Exact	EUR		0	0	0	0
18	Mes Idées De Mots Clés camion strasbourg	Exact	EUR		0	0	0	0
19	Mes Idées De Mots Clés camion paris	Exact	EUR		10	0	0	0
20	Mes Idées De Mots Clés voiture strasbourg	Exact	EUR		0	0	0	0
21	Mes Idées De Mots Clés voiture nantes	Exact	EUR		10	0,61	0	0
22	Mes Idées De Mots Clés velo strasbourg	Exact	EUR		10	0	0	0
23	Mes Idées De Mots Clés vélo nantes	Exact	EUR		10	0,17	0,31	0
24	Mes Idées De Mots Clés voiture paris	Exact	EUR		10	0,82	0	0
25	Mes Idées De Mots Clés velo nantes	Exact	EUR		30	0,19	0	0
26	Mes Idées De Mots Clés vélo paris	Exact	EUR		10	0	0	0
27	Mes Idées De Mots Clés velo paris	Exact	EUR		10	1	0	0
28	Vélo Paris Itineraire velo paris	Exact	EUR		0	0	0	0
29	Vélo Paris paris rando velo	Exact	EUR		10	0	0	0
30	Vélo Paris paris a velo	Exact	EUR		0	0	0	0
31	Vélo Paris velo a paris	Exact	EUR		10	0	0	0
32	Vélo Paris trajet velo paris	Exact	EUR		0	0	0	0
33	Vélo Paris velo occasion paris	Exact	EUR		0	0	0	0
34	Vélo Paris paris vélo	Exact	EUR		0	0	0	0
35	Vélo Paris velo hollandais paris	Exact	EUR		10	0	0	0
36	Vélo Paris location velo paris	Exact	EUR		0	0	0	0
37	Vélo Paris velo electrique paris	Exact	EUR		10	1	1	0
38	Vélo Paris boutique velo paris	Exact	EUR		10	0	0	0
39	Vélo Paris balade velo paris	Exact	EUR		0	0	0	0
40	Vélo Paris coursier velo paris	Exact	EUR		0	0	0	0
41	Vélo Paris velo d occasion paris	Exact	EUR		0	0	0	0

Remarque : en cas de problème de caractères spéciaux lors de l'import, recommencez l'opération et changez l'origine du fichier du format ANSI au format UTF-8.

Etape 2 : doublons et données inutiles

Supprimez tout d'abord les lignes de Total situées au début, ainsi que toutes les colonnes sauf « Ad Group », « KeyWord » et « Avg. Monthly Researches... ».

Sélectionnez ensuite tout le tableau, et utilisez la fonction de suppression des doublons d'Excel (attention, faites le tri uniquement sur la colonne « Keyword »).



L'étape suivante consiste à réduire le « bruit », c'est-à-dire tous les mots clés peu importants car très peu tapés (surtout si vous avez des milliers de mots clés lors de l'importation). En règle générale, nous conseillons de supprimer purement et simplement tous les mots clés tapés en moyenne moins de 90 fois par mois. Cela représente en effet à peine 4 recherches par jour, sachant qu'une position en top 1 ne garantit même pas d'obtenir l'ensemble de ce trafic.

Sur des audits à plusieurs milliers de mots clés, on peut même masquer tous les mots clés tapés moins de 100, 200 voire 300 fois par mois pour ne garder que les plus pertinents et rémunérateurs.

Etape 3 : segmentation

Pour être utile, il faut ensuite segmenter votre liste par thématiques. Google vous mêche en partie le travail avec la colonne « Ad Group » censée classer vos mots clés. Vous vous en rendez-compte assez vite mais ce tri est parfois douteux.

Vous allez donc devoir passer en revue manuellement chaque mot clé pour savoir s'il est réellement pertinent (les suggestions de l'outil de planification peuvent parfois être hors-sujet par rapport à la thématique de votre site), et vous devrez ensuite les classer par thématique.

Si vous êtes référenceur en agence, faites participer vos clients à ce tri, car ils seront les plus à même de trouver des thématiques communes parmi toutes vos expressions.

Etape 4 : utilisation de ces données

C'est ce classement thématique qui va vous permettre de réellement utiliser ces données. Une fois terminé, vous pourrez en effet lister tous les mots clés pertinents pour chaque thématique ou gamme de produits. Dans chaque thème, vous connaîtrez vos mots clés phares et ceux de longue traîne. Libre à vous ensuite de les utiliser pour :

- remanier la structure du site ;
- créer ou optimiser les contenus ;
- savoir vers quelles pages et avec quelles ancrs orienter votre travail de netlinking ;
- mettre en place un suivi de positionnement ;
- mettre en place une veille concurrentielle et technologique sur votre secteur d'activité ;
- créer un site spécifique pour certaines thématiques très importantes en termes de trafic et/ou de rentabilité ;
- ...

Conclusion

L'outil de planification des mots clés d'Adwords est réellement un très bon outil quand on travaille le référencement d'un site Internet.

Il permet de comprendre rapidement les expressions d'un secteur d'activité ou d'un site en particulier ainsi que le comportement de l'internaute. Cela permet ensuite de pouvoir cibler et améliorer son travail de référencement sur un ou plusieurs groupes précis d'expressions.

La seule règle à suivre est de savoir correctement filtrer puis lire les données que l'outil nous donne, sous peine de mal les comprendre et de définir une mauvaise stratégie pour son site (que ce soit pour l'optimisation ou la rédaction de contenus, la structure ou le netlinking).

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

Comment utiliser intelligemment les Sitemaps XML ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Les fichiers Sitemaps XML fêteront l'année prochaine leurs 10 ans d'existence. Parfois décriés par certains au niveau de leur utilité pour le référencement, ils s'avèrent pourtant très importants dans de nombreux cas. Encore faut-il bien les mettre en place, dans les règles de l'art, et connaître leurs nombreuses possibilités. Voici un article détaillé sur les bonnes pratiques en la matière qui devrait aider votre site à être mieux crawlé et indexé par Google, Bing et leur confrères...

L'utilité des sitemaps XML n'a jamais totalement fait l'unanimité au sein de la communauté des experts SEO. Leur impact sur le référencement naturel des sites a également souvent été remis en question. Pourtant, si on comprend leur objectif réel et si on sait comment les implémenter, un référenceur peut tirer un réel bénéfice de la soumission de ces fichiers d'URL au format XML.

L'objectif de cet article est donc de faire le point sur les fonctions assurées par les sitemaps, sur les gains réels qu'ils apportent, et sur les précautions à prendre lorsqu'on les construit. Au passage, nous rappellerons quelques astuces avancées, parfois méconnues par les webmasters.

Pourquoi le protocole Sitemap a-t-il été inventé ?

Les moteurs de recherche comme Google ont abandonné très tôt le principe de la soumission d'URL aux moteurs, pour leur préférer une exploration automatique des pages du web en suivant les liens découverts dans les pages crawlées.

Mais lorsque l'on utilise ce mode d'exploration, deux défis apparaissent : assurer une "couverture" aussi parfaite que possible du web, et garantir la "fraîcheur" des données explorées.

Le problème de la couverture

Une partie des pages se situe dans une zone du web que l'on appelle souvent le "web invisible". Il s'agit de pages que les moteurs de recherche ne peuvent pas découvrir en suivant des liens placés sur des pages explorables. La structure des sites web et de nombreux problèmes techniques peuvent créer des situations (volontaires ou involontaires) où une page n'est liée à aucune page accessible aux moteurs. Le problème s'aggrave si on considère que les moteurs décident aussi d'"oublier" parfois volontairement d'explorer des pages qu'ils considèrent comme inutiles (doublons, pages extrêmement profondes etc.).

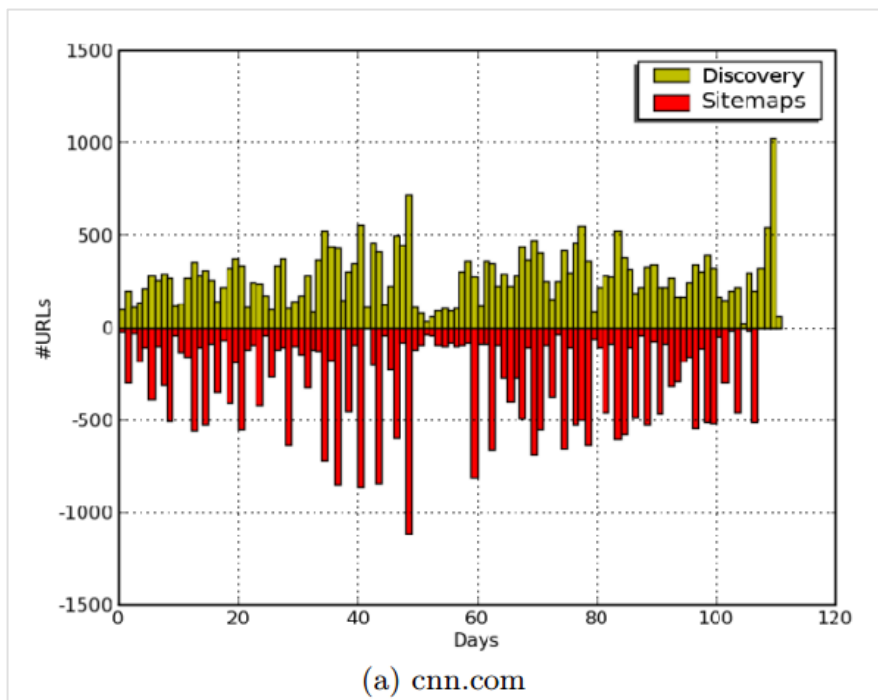
La "couverture" d'un robot d'exploration d'un moteur de recherche est caractérisée par le ratio entre les pages web explorées et les pages web intéressantes.

Améliorer la couverture est apparu comme une priorité absolue aux ingénieurs de Google, dès le début des années 2000, tant le volume des pages "ignorées" semblait grand.

La "fraîcheur" des données

L'autre défi posé par l'exploration automatique des pages web à l'aide d'un spider, est de garantir la fraîcheur des données. En effet, ce processus prend du temps, et dans la

pratique, il peut s'écouler des semaines entre deux explorations d'une même page web. Si un lien vers une nouvelle page apparaît sur un site, il peut donc s'écouler beaucoup trop de temps avant que cette page ne soit détectée et explorée. Sur certains types de sites, des liens éphémères peuvent même ne jamais être détectés.



Une illustration tirée de l'article « sitemaps : beyond and above the crawl of duty ». L'analyse du nombre d'URL réellement explorées par Google sur le site CNN (en vert), comparées à celle mentionnées dans le sitemap (en rouge) montre que l'exploration passe à côté d'un grand nombre de pages. Pour un site dont le contenu évolue rapidement, les sitemaps sont indispensables pour assurer à la fois une bonne couverture et garantir la fraîcheur des résultats.

Jun 2005 : Google introduit le protocole sitemap XML

Les ingénieurs de Google se sont rendu compte que le seul moyen d'améliorer la couverture et la fraîcheur des URL explorées étaient de mettre les webmasters des sites à contribution, en leur suggérant de soumettre la liste des URL de leur site au moteur.

En Juin 2005, Google a donc annoncé (<http://actu.abondance.com/2005-23/google-sitemaps.php>) le support d'un nouveau protocole baptisé "Sitemaps" (plans de site) conçu à cette fin, sur le modèle du protocole robots.txt. Un site officiel permet de s'informer sur le format à respecter pour construire un fichier sitemap : <http://www.sitemaps.org/fr/>

Qui supporte le format Sitemaps ?

Bing supporte le format depuis 2007. Baïdu et Yandex gèrent aujourd'hui aussi parfaitement le protocole sitemaps. Orange et Exalead également.

Le format sitemap XML

Décrire dans le détail le format des sitemaps XML dépasse la portée de cet article : on se reportera donc utilement à la documentation fournie sur le site du protocole sitemaps mentionné ci-dessus, et à l'aide de Google sur les sitemaps :

<https://support.google.com/webmasters/answer/156184?hl=fr>

Rappelons simplement que les sitemaps XML doivent être encodés en UTF-8, et respecter le format habituel des fichiers XML, notamment l'obligation d'échapper les caractères réservés. Voici un exemple basique de fichier sitemap XML

```
<?xml version="1.0" encoding="utf-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd">
  <url>
    <loc>http://example.com/</loc>
    <lastmod>2006-11-18</lastmod>
    <changefreq>daily</changefreq>
    <priority>0.8</priority>
  </url>
</urlset>
```

Les deux autres formats autorisés pour le sitemaps XML

Le format XML reste le format recommandé et c'est également le plus utilisé par les webmasters (à 90% d'après un chiffre communiqué par Google en 2009). Mais le protocole Sitemaps prévoit le support de deux autres formats de fichier :

- Le format texte

Il est possible de stocker les URL à soumettre dans un fichier texte brut, à raison d'une ligne par URL. Le fichier doit être encodé en UTF-8. Le format des URL doit respecter les directives défini par le protocole sitemaps, et les fichiers sont soumis aux mêmes limitations de taille que les sitemaps XML.

- Le format "flux de syndication" (flux RSS)

On peut également se servir des formats définis pour les flux de syndication, à savoir : Atom 0.3 ou 1.0, et RSS 2.0. Cette possibilité s'avère très pratique lorsque l'on utilise des CMS qui savent créer facilement des flux RSS pour le contenu d'un site mais ne disposent pas d'un outil générant des sitemaps xml automatiquement.

Le problème de la taille des sitemaps

Les fichiers sitemaps sont limités en taille :

- ils sont limités à 50 000 URL par fichier.
- et ne peuvent pas dépasser un poids de plus de 10 Mo.

Notons qu'il est possible, et même hautement conseillé, de compresser les fichiers sitemaps au format gzip. Donc dans la pratique, la limite d'URL sera atteinte avant la limite de poids du fichier.

La solution pour les "gros" sites : le fichier "sitemap index"

Pour un site de taille respectable, la limite des 50 000 URL est vite atteinte. La solution consiste donc à éclater les sitemaps en plusieurs fichier et à utiliser un fichier "sitemap index" qui indique la liste des différents fichiers xml sitemaps créés. Le format du fichier sitemap index est légèrement différent :

```
<?xml version="1.0" encoding="UTF-8"?>
<sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <sitemap>
    <loc>http://www.example.com/sitemap1.xml.gz</loc>
    <lastmod>2004-10-01T18:23:17+00:00</lastmod>
```

```
</sitemap>
<sitemap>
  <loc>http://www.example.com/sitemap2.xml.gz</loc>
  <lastmod>2005-01-01</lastmod>
</sitemap>
</sitemapindex>
```

Les formats de sitemaps spécifiques

Au fil du temps, le protocole sitemap a été enrichi pour permettre l'exploration de données spécifiques. Ces formats particuliers concernent :

- les vidéos ;
- les images ;
- les actualités ;
- les URL mobile ;
- les pages multilingues.

Le format sitemap video :

Ce format spécifique est détaillé ici :

<https://support.google.com/webmasters/answer/80472?hl=fr>

Google a créé ce format car l'exploration automatique des vidéos se heurte à de nombreux problèmes techniques. Beaucoup d'implémentations de lecteurs vidéo empêchent une exploration normale des flux, il est donc fortement recommandé de construire un sitemap vidéo.

On peut également utiliser des fichiers au format mRSS.

L'ajout d'information sur les images :

Il est possible d'utiliser des balises complémentaires pour soumettre des données sur les images que l'on souhaite voir indexer pour ses pages. Ces balises sont décrites ici :

<https://support.google.com/webmasters/answer/178636?hl=fr>

Voici un exemple de sitemap enrichi par les balises pour les images :

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:image="http://www.google.com/schemas/sitemap-image/1.1">
<url>
  <loc>http://example.com/sample.html</loc>
  <image:image>
    <image:loc>http://example.com/image.jpg</image:loc>
  </image:image>
  <image:image>
    <image:loc>http://example.com/photo.jpg</image:loc>
  </image:image>
</url>
</urlset>
```

On peut déclarer au maximum 1000 images pour une page.

Les sitemaps "news" :

Un format spécifique a été développé par Google pour les pages d'actualités. Ce sitemap est conçu pour faciliter l'exploration et l'indexation des pages dans le moteur vertical Google News. Compte tenu des spécificités du crawler de Google News (et de ses limites techniques ainsi que de son caractère "capricieux"), construire un sitemap news quand on est référencé dans un site d'actualité est fortement recommandé.

Les spécificités de ce format sont détaillées ici :

<https://support.google.com/news/publisher/answer/74288?hl=fr>

Voici un exemple de fichier exploitant ce format :

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:news="http://www.google.com/schemas/sitemap-news/0.9">
  <url>
    <loc>http://www.example.org/business/article55.html</loc>
    <news:news>
      <news:publication>
        <news:name>Journal L'Exemple</news:name>
        <news:language>fr</news:language>
      </news:publication>
      <news:access>Subscription</news:access>
      <news:genres>PressRelease, Blog</news:genres>
      <news:publication_date>2008-12-23</news:publication_date>
      <news:title>Les entreprises A et B envisagent une fusion</news:title>
      <news:keywords>économie, fusion, acquisition, A, B</news:keywords>
      <news:stock_tickers>NASDAQ:A, NASDAQ:B</news:stock_tickers>
    </news:news>
  </url>
</urlset>
```

Indiquer les URL pour mobile dans un sitemap :

Lorsqu'un site dispose d'URL destinées aux appareils mobiles, il est possible d'indiquer cette spécificité à l'aide de balises supplémentaires.

Ce format est détaillé ici :

<https://support.google.com/webmasters/answer/34648?hl=fr>

```
<?xml version="1.0" encoding="UTF-8" ?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:mobile="http://www.google.com/schemas/sitemap-mobile/1.0">

  <url>
    <loc>http://mobile.example.com/article100.html</loc>
    <mobile:mobile/>
  </url>
</urlset>
```

Indiquer les pages multilingues dans un sitemap :

Enfin, il est possible d'ajouter des informations de type "hreflang" dans les sitemaps, pour indiquer qu'une page existe en plusieurs versions linguistiques.

Voici un exemple de ce format :

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:xhtml="http://www.w3.org/1999/xhtml">
  <url>
    <loc>http://www.mon-domaine.com/english/</loc>
    <xhtml:link rel="alternate" hreflang="de" href="http://www.mon-
domaine.com/deutsch/" />
    <xhtml:link rel="alternate" hreflang="de-ch href="http://www.mon-
domaine.com/schweiz-deutsch/" />
    <xhtml:link rel="alternate" hreflang="en" href="http://www.mon-
domaine.com/english/" />
```

```
</url>
<url>
  <loc>http://www.mon-domaine.com/deutsch/</loc>
  <xhtml:link rel="alternate" hreflang="en" href="http://www.mon-
domaine.com/english/" />
  <xhtml:link rel="alternate" hreflang="de-ch" href="http://www.mon-
domaine.com/schweiz-deutsch/" />
  <xhtml:link rel="alternate" hreflang="de" href="http://www.mon-
domaine.com/deutsch/" />
</url>
<url>
  <loc>http://www.mon-domaine.com/schweiz-deutsch/</loc>
  <xhtml:link rel="alternate" hreflang="de" href="http://www.mon-
domaine.com/deutsch/" />
  <xhtml:link rel="alternate" hreflang="en" href="http://www.mon-
domaine.com/english/" />
  <xhtml:link rel="alternate" hreflang="de-ch" href="http://www.mon-
domaine.com/schweiz-deutsch/" />
</url>
</urlset>
```

Ce format est décrit ici :

<https://support.google.com/webmasters/answer/2620865>

Remarque : la plupart de ces formats sont spécifiques à Google. Mais ils respectent tous la norme du protocole sitemap, donc ils sont tous partiellement supportés par les autres moteurs (au pire, les informations complémentaires sont ignorées).

Où placer les fichiers sitemaps ?

Le protocole sitemaps introduit une contrainte importante : seules les URL situées dans le même répertoire que le sitemap lui-même seront prises en compte. Par exemple, si le fichier sitemap est placé dans ce répertoire :

www.domaine.com/rep1/

Alors ce fichier peut contenir des URL de ce type

www.domaine.com/rep1/page.html ou www.domaine.com/rep1/ssrep/page.html

mais pas

www.domaine.com/repdifferent/page.html

Le fichier sitemap index peut lui être placé n'importe où sur le site. Par contre, les fichiers sitemap individuels qu'il mentionne doivent respecter la règle ci-dessus.

Cette contrainte peut s'avérer problématique s'il est impossible de placer le fichier sitemap à la racine du site.

Google propose heureusement une solution : si le fichier est soumis via Google Webmaster Tools, les URL appartenant à un host partageant le même propriétaire validé sont autorisés. Dans ce cas, un sitemap peut contenir des URL appartenant à différents sous domaines et même à différents domaines !

<http://hote1.example.com>

<http://hote2.example.com>

<http://hote3.example.com>

<http://hote1.example1.com>

<http://hote1.example.ch>

Utilisez "lastmod" à bon escient

Eclater les fichiers sitemaps en plusieurs fichiers présente un avantage supplémentaire. Il peut être intéressant de placer dans des fichiers séparés :

- les URL statiques (zones du site qui ne changent pas au fil du temps) ;

- les URL récentes (pages nouvellement ajoutées) ;
- les URL éphémères (zones du site où des pages apparaissent et disparaissent fréquemment).

Cette séparation permet de jouer sur l'effet de la balise xml "lastmod", qui indique la date de dernière modification du fichier. La mise à jour correcte de cette balise permet d'assurer la prise en compte rapide des URL récentes et des URL éphémères.

Attention à la génération automatique des sitemaps

Il existe beaucoup de solutions permettant de générer automatiquement les sitemaps pour un site internet.

Ces solutions sont de deux types :

- **les plugins ou fonctionnalités internes** : beaucoup de CMS disposent de solutions permettant de générer un sitemap automatiquement à partir de leur base de données de contenu. Le problème est que ces solutions ignorent toutes les pages du site qui n'auraient pas été créées à l'aide du CMS : le Sitemap peut donc s'avérer incomplet dans certains cas.
- **les crawlers générateurs de sitemap** : dans ce cas, le sitemap est généré à partir d'un crawl du site. Le sitemap contient donc toutes les URL... explorables. Cette solution présente donc l'inconvénient d'être incapable d'inclure des URL situées dans des zones invisibles du site (cachées derrière un formulaire, ou derrière un script ajax ou actionscript). Et il faudra générer un nouveau sitemap à chaque changement sur le site (création, modification, suppression de page), pour avoir un fichier à jour, ce qui peut rapidement s'avérer fastidieux.

Il est important de ne pas accorder trop de confiance à ces générateurs automatiques de sitemaps. Un examen attentif des résultats montre :

- qu'ils sont parfois bogués, et génèrent des syntaxes d'URL erronées ;
- qu'ils sont souvent limités, et ne savent pas générer des formats spécifiques ;
- qu'ils ne tiennent pas compte des pages bloquées par le robots.txt ;
- qu'ils ne savent pas proposer une version canonique.

Dans la pratique, il faudra donc :

- soit corriger les fichiers produits par ces générateurs ;
- soit modifier les scripts pour obtenir un résultat conforme.

Eviter l'indexation des sitemaps

Les sitemaps sont des fichiers web qui peuvent aussi être explorés et indexés par des moteurs de recherche. Une requête sur "sitemap.xml" permet de vérifier que cette situation n'est pas exceptionnelle. Lorsque cela arrive, les webmasters veulent en général faire disparaître ce résultat, pour des raisons de sécurité ou pour éviter de faciliter le vol de contenu.

La solution consiste à utiliser l'attribut noindex dans une directive x-robots-tag placée dans l'en-tête http: du fichier sitemap.

Le format de cette directive est décrit ici :

https://developers.google.com/webmasters/control-crawl-index/docs/robots_meta_tag?hl=fr

Soumettre les sitemaps

Les fichiers sitemap / sitemap index ne peuvent pas être découverts automatiquement par les moteurs, car contrairement au robots.txt, le nom de ce fichier et son emplacement ne sont pas fixes.

On peut "soumettre" le fichier de deux façons différentes :
- soit en le déclarant dans le fichier robots.txt ;
- soit en le soumettant via un compte webmaster tools.

La soumission via le fichier robots.txt

La syntaxe est très simple :

Sitemap: <http://www.mon-domaine.fr/sitemap1.xml>
Sitemap: <http://www.mon-domaine.fr/sitemap2.xml.gz>

Pensez à séparer par un caractère retour chariot créant une ligne blanche les directives concernant les sitemaps. On peut déclarer autant de sitemaps qu'on le souhaite dans un fichier robots.txt

Cette solution fonctionne avec tous les moteurs de recherche supportant le protocole sitemaps.

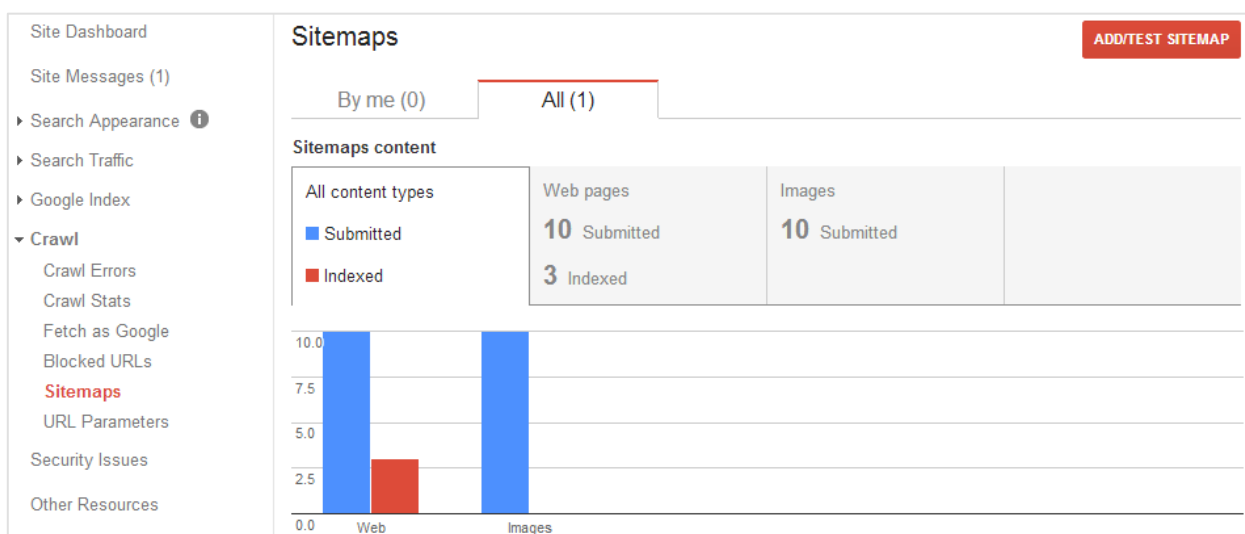
La soumission via les webmaster tools

Certains moteurs (par exemple Google, mais également Bing) permettent la soumission des sitemaps via le compte Google Webmaster Tools. Avoir recours à cette solution n'empêche pas d'utiliser aussi la soumission via robots.txt.

La soumission via Webmaster Tools permet d'accéder à un grand nombre d'informations utiles comme :

- la confirmation de la prise en compte du fichier par le moteur ;
- la liste des erreurs rencontrée par le moteur ;
- le nombre d'URL indexées.

Il est donc possible de savoir facilement quel est le taux d'indexation d'un groupe d'URL défini par un fichier sitemap, ce qui peut fournir des renseignements utiles pour le référencement.



Vue de la rubrique sitemaps de Google Webmaster Tools : ici le sitemap soumis contient 10 URL, mais seulement 3 sont indexées.

Le Ping de sitemaps pour accélérer leur prise en compte

Il peut s'écouler un certain temps (plusieurs heures, voire plusieurs jours selon les sites) entre la soumission d'un fichier sitemap et sa prise en compte par les moteurs. C'est surtout vrai si le fichier est soumis via le fichier robots.txt uniquement.

Une façon très efficace de signaler à un moteur la présence d'un nouveau fichier sitemap est d'utiliser les URL de ping des moteurs.

Par exemple pour Google la syntaxe est :

www.google.com/webmasters/tools/ping?sitemap=http%3A%2F%2Fwww.yoursite.com%2Fsitemap.gz

Via le compte webmaster tools de Google, il suffit de cliquer sur le bouton "soumettre à nouveau" et le moteur sera averti que vous avez rafraîchi le contenu du sitemap.

Les sitemaps sont-ils réellement utiles ?

L'utilité des sitemaps pour le référencement a fait l'objet dans le passé de débats riches et passionnés. Deux camps se sont longtemps opposés. Certains experts, dont Rand Fishkin de Moz, ont longtemps défendu l'idée que les sitemaps étaient potentiellement plus nuisibles qu'autre chose. A l'inverse, d'autres ont défendu l'idée que les sitemaps xml étaient indispensables pour le référencement.

Avec le temps et l'expérience, un consensus a commencé à se dégager en faveur de l'utilité des sitemaps.

L'un des arguments des "anti" sitemaps était l'existence d'un phénomène d'"aplatissement" de la structure du site. En effet, la notion d'arborescence disparaît totalement dans les sitemaps, qui ne contiennent aucune information sur la manière dont les pages sont reliées entre elles. Alors que cette information est évidemment collectée lors de l'exploration automatique traditionnelle.

Mais en réalité, l'information sur les liens entre les pages, ou l'arborescence n'est pas "écrasée" en cas de recours aux sitemaps. Les moteurs ajoutent simplement à la liste des URL stockées dans l'ordonnanceur (le composant logiciel qui décide des URL à crawler) les URL manquantes découvertes dans les sitemaps.

Une autre critique adressée aux sitemaps était leur inefficacité en termes d'amélioration du référencement et du positionnement. Force est de constater que les sitemaps ont pour mission unique de résoudre des problèmes de couverture et de fraîcheur lors des crawls, et que leur impact sur le référencement ne peut être qu'indirect, en favorisant la découverte et l'indexation de pages qui auraient été ignorées sans leur présence.

Avec le temps, le camp des "anti" s'est dégarni (Rand Fishkin s'est rangé dans le camp des pro-sitemaps après deux ans d'expérience), et une majorité des experts SEO, à l'instar des moteurs eux-mêmes, recommandent de construire et de soumettre des fichiers sitemaps.

La balise "priority" sert-elle à quelque chose ?

Si on ne peut pas indiquer dans un sitemap xml la position d'une url dans l'arborescence, le protocole prévoit, dans son format xml, de pouvoir indiquer via le tag "priority" une valeur entre 0,0 et 1,0 (la valeur de priorité par défaut est de 0,5 pour une url) indiquant le niveau d'importance de la page selon les webmasters. Il s'agit d'une indication fournie aux moteurs que ces derniers exploitent... ou non. Une analyse des logs montre en effet souvent que le comportement de crawl de Google n'est pas impacté par des changements de valeur de priorité. En réalité, cette balise comme la balise <changefreq> qui déclare la fréquence de changement des pages, sont ignorées si le crawler estime les données issues de l'exploration normale plus fiable. Ce comportement frustrant ne doit pas vous empêcher d'indiquer des valeurs cohérentes pour ces balises.

Des sitemaps indispensables dans les phases de lancement de sites / de création de pages

Certaines étapes de la vie d'un site peuvent être facilitées grâce à la soumission de sitemaps adaptés : il s'agit de toutes les étapes de lancement des sites ou d'ajout massif de nouvelles URL. Dans ces contextes, un fichier sitemap peut aider le moteur à découvrir et indexer rapidement les URL. C'est d'autant plus indispensable si le moteur crawle peu fréquemment votre site.

Indiquez l'url canonique dans vos sitemaps

Il faut signaler un comportement particulier de Google dans la prise en compte des URL incluses dans un sitemap : le moteur considère que la syntaxe incluse dans un sitemap est celle que le webmaster déclare préférer. Cela signifie qu'en cas de doublon d'URL, la simple mention d'une des deux syntaxes dans un sitemap pourra conduire Google à "canonicaliser" les syntaxes dupliquées, pour retenir la syntaxe du sitemap (par exemple, si le sitemap contient des URL en https, et si le site contient aussi des versions http des mêmes pages, la version indexée sera la version https).

Compte tenu de ce comportement, il est conseillé de rester cohérent, et d'adopter des conventions similaires dans les link rel canonical. Par exemple :

Si l'on retient l'URL <http://www.domaine.com/page.html?param=1> comme URL préférentielle et qu'il existe une URL avec un contenu identique dont l'URL serait <http://www.domaine.com/page.html?param=2> alors la première url doit être placée dans le sitemap (pas la seconde), et la deuxième URL peut contenir un *link rel=canonical* pointant vers la première URL (et pas l'inverse).

Faut-il inclure des pages en noindex ou en nofollow dans le sitemap ?

Les sitemaps servent à indiquer la liste des URL que les moteurs doivent explorer. Cela n'a rien à voir avec l'indexation. Par conséquent :

- des pages contenant une meta robots avec comme valeur d'attributs noindex, follow ont parfaitement leur place dans un fichier sitemap ;
- par contre, les pages en nofollow, logiquement, doivent être exclues ;
- de même les URL dont la syntaxe est bloquée par une directive disallow dans le robots.txt n'ont rien à faire dans le sitemaps ;

Attention à ne pas mettre en ligne des sitemaps bourrés d'erreur

Si vous voulez que les sitemaps soient correctement pris en compte par les moteurs, il est important de faire la chasse aux erreurs qu'ils peuvent contenir. En particulier

- les URL dans les sitemaps ne doivent pas renvoyer des codes 404, 500 ou tout code autre que 200 ;
- les URL dans les sitemaps ne doivent pas être redirigées ;
- les URL dans les sitemaps doivent correspondre à des pages uniques, autant que possible ;
- les URL ne doivent pas être en nofollow ou bloquées par un robots.txt.

Si le sitemap contient trop d'erreurs, le contenu du fichier ne sera pas pris en compte par le moteur de recherche. Ce taux est connu pour Bing : 10%.

Informations détaillées sur les sitemaps								← Tester le sitemap	
Sitemap : /fr/sitemap.xml									
Ce sitemap a été envoyé le 23 août 2011 et traité le 19 avr. 2012.									
Informations sur l'erreur : 2 erreurs, 0 avertissements.									
Afficher :		<input type="button" value="Tous"/> <input type="button" value="Erreurs"/> <input type="button" value="Avertissements"/>		Afficher		25 lignes ▼		1 à 2 sur 2 < >	
#	Type	Problème	Description	Nombre de problèmes	Exemple	Ligne	Détectée		
1	Erreurs	Erreur	Une erreur s'est produite lors de la tentative d'accès à votre sitemap. Veuillez vous assurer qu'il respecte nos consignes et qu'il est accessible à l'emplacement que vous avez indiqué, puis envoyez-le à nouveau.	1	Erreur HTTP générique: 404 introuvable Erreur HTTP : 404	-	4 juil. 2014		
2	Erreurs	Le sitemap est un fichier HTML.	Votre sitemap semble correspondre à une page HTML. Veuillez utiliser un format de sitemap pris en charge.	1	Tag : html	2	19 avr. 2012		

Dans la rubrique sitemaps, Google Webmaster Tools retourne les problèmes d'exploitation des sitemaps en les classant en deux types : les « erreurs » qui empêchent la prise en compte du fichier entier, et les avertissements qui concernent des URL individuelles.

Conclusion

Soumettre un Sitemap XML pour les moteurs de recherche se révèle être toujours une excellente idée. Les sitemaps news, ou video, sont même indispensables pour un bon référencement dans les moteurs verticaux de Google. Mais il ne faut pas s'attendre à un impact direct sur le positionnement des pages web de votre site : leur objectif est simplement d'aider les moteurs de recherche à découvrir l'intégralité du contenu de votre site.

Parfois décriés, les sitemaps ne présentent pas d'inconvénients majeurs pour le référencement, et ils ne présentent que des avantages si les sitemaps ne contiennent pas d'erreurs. Le soit-disant impact négatif des sitemaps pour le SEO n'est qu'une légende urbaine. En réalité, la tendance lourde est même aujourd'hui d'enrichir le contenu possible des sitemaps pour élargir leurs fonctionnalités...

Après bientôt dix ans d'existence, l'utilité des sitemaps ne cesse de progresser et ils vont probablement continuer à jouer un rôle important à l'avenir pour le référencement.

Bibliographie

Sites utiles :

Site officiel du protocole sitemaps
<http://www.sitemaps.org>

Page de support de Google sur les sitemaps
<https://support.google.com/webmasters/answer/156184?hl=fr>

Article scientifique :

Sitemaps: Above and Beyond the Crawl of Duty
<http://www2009.eprints.org/100/1/p991.pdf>
Uri Schonfeld, UCLA Computer Science Department. Narayanan Shivakumar, Google Inc.

Philippe YONNET, Directeur Général de l'agence Search-Foresight
(<http://www.search-foresight.com>).

Google a-t-il le droit de "scaper" nos contenus ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Google propose de plus en plus dans ses résultats des extraits de pages web, comme il l'a montré récemment en reprenant purement et simplement des tutoriels directement dans ses SERP sans demander l'avis des éditeurs les ayant créés sur leurs sites respectifs. A-t-il juridiquement le droit de mettre en place de telles pratiques ? Peut-il être attaqué pour "vol de contenu", ce à quoi peut être assimilé le "scraping" ? Réponses dans cet article...

Le post de Matt Cutts, en décembre 2013, sur le scraping et la mise à disposition d'un formulaire de signalement de ce type de cas (<http://www.abondance.com/actualites/20131205-13450-matt-cutts-et-le-scraping-de-textes.html>) a rouvert le débat autour de cette pratique. Et ce, d'autant plus étrangement que, concomitamment et de plus en plus, Google recopie quelques lignes de contenu de certains sites (comme Wikipédia) dans le cadre de la fourniture de résultats naturels (<http://www.abondance.com/actualites/20140626-14037-knowledge-graph-commence-afficher-tutoriels-scrape-web-vergogne.html>). La nature et les risques juridiques de cette pratique sont pourtant précis et ne souffrent pas de discussion. Les exemples de procès perdus par Google Actualités dans le cadre de la reprise de filets d'actualités, sont là pour en témoigner.

Le concept de « scraping »

Le terme « scraping » vient de l'anglais et peut signifier « racler » ou « prélever ». Ce terme est notamment utilisé dans un sens médical.

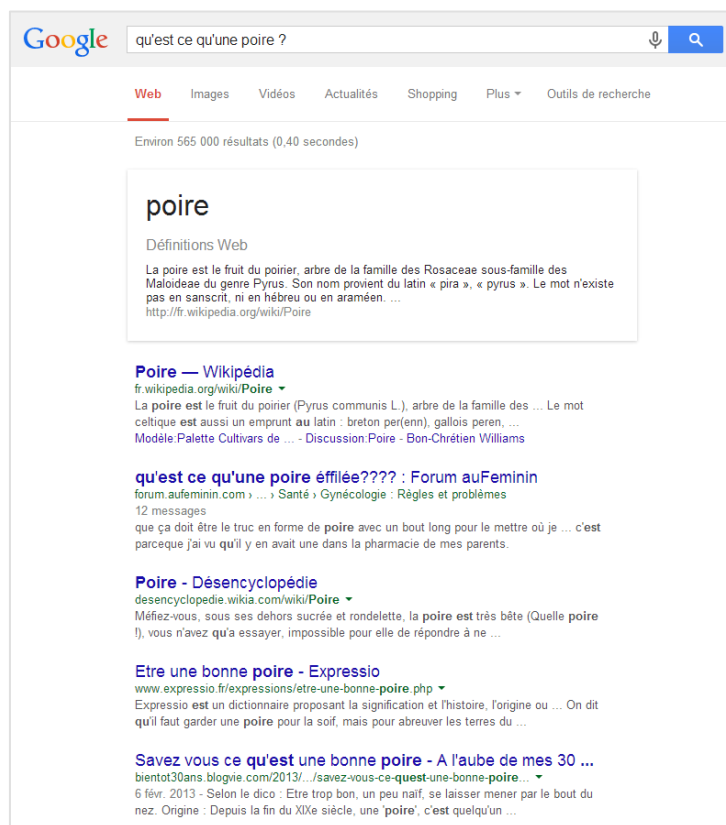
Il semble exister plusieurs cas de scraping :

- Selon Wikipedia (http://fr.wikipedia.org/wiki/Web_scraping), le *Web scraping* (parfois appelé *Harvesting*) est une technique d'extraction du contenu de sites Web, via un script ou un programme, dans le but de le transformer pour permettre son utilisation dans un autre contexte, par exemple le référencement ;
- Le scraping peut concerner uniquement les images ou vidéos, voire des écrans d'impression de sites ;
- Le scraping peut viser une technique d'aspiration de données structurées, dans une base de données, qui peut être réalisée entre deux applications (par exemple, en cas de clusterisation de données).

La notion induite est donc la copie de données.

Concrètement, le scraping peut viser tout ou partie d'un site.

- Il semble exister des pratiques consistant à recopier intégralement un site ;
- De nombreux sites ou moteurs de recherche reprennent des extraits de site (comme dans l'exemple ci-dessous avec Google reprenant les premières lignes du contenu de la page Wikipédia).



Google peut-il « scraper » le contenu de tiers ?

En faisant apparaître un filet de quelques lignes reprenant le contenu d'un site tiers, Google pratique donc clairement le scraping.

D'un point de vue juridique, cette reprise pourrait apparaître comme de la violation de droit d'auteur (ou copyright dans les pays anglo-saxons).

En pratique, le régime du droit d'auteur exige que l'œuvre copiée soit originale et une création de forme perceptible par les sens (la vue pour ce qui est des écrits). La condition d'originalité requiert que l'œuvre porte l'empreinte de la « personnalité » de son auteur. L'originalité est un concept distinct de celui de « nouveauté ». Ainsi, le contenu d'un site, un texte, une photo, etc... n'est protégé par le droit d'auteur que s'il est original dans le sens qu'il porte l'empreinte de la « personnalité » de son auteur. Les tribunaux font attention à ce principe directeur lorsqu'ils déterminent si oui ou non une œuvre est protégée. Ainsi, contrairement à ce que beaucoup de personnes pensent, un contenu banal qui n'a rien d'original n'est pas protégé par le droit d'auteur et l'auteur n'a pas le droit d'interdire à un tiers de le copier.

Si un contenu est protégé au titre du droit d'auteur, il n'en reste pas moins qu'un tiers a toutefois le droit d'utiliser une courte partie de ce contenu dans certaines conditions : l'article L122-5 du Code de la Propriété Intellectuelle dispose précisément que « *lorsque l'œuvre a été divulguée, l'auteur ne peut interdire (...), sous réserve que soient indiqués clairement le nom de l'auteur et la source les analyses et courtes citations justifiées par le caractère critique, polémique, pédagogique, scientifique ou d'information de l'œuvre à laquelle elles sont incorporées (...)* », étant précisé que cette exception ne peut « *porter atteinte à l'exploitation normale de l'œuvre ni causer un préjudice injustifié aux intérêts légitimes de l'auteur* ».

Il faut alors :

- Que le nom de l'auteur et la source soient mentionnés clairement ;
- Que l'extrait de l'œuvre originelle soit « incorporé » dans un ensemble plus large. Cette incorporation suppose que l'ensemble plus large ait un caractère « critique, polémique, pédagogique, scientifique ou d'information » qui suppose un travail d'analyse plus général où les extraits ont vocation à exposer le propos. ;
- Que la citation réponde à l'exigence de brièveté, n'être qu'un accessoire de l'œuvre dans laquelle elles s'insèrent : pour en juger, le critère traditionnellement reçu consiste à apprécier si l'œuvre dite seconde, celle qui recourt aux citations, peut survivre sans dénaturation à la suppression desdites citations. Cependant la loi ne prévoit pas de "taille mesurable" pour cette notion de "citation", qu'il est du coup complexe de quantifier en termes de mots ou de caractères mais simplement en "exigence de brièveté"... Ainsi, dans le passé, certains tribunaux ont accepté la reprise de 50% d'un passage, et d'autres courts ont condamné pour 10% repris...

Si la première condition est remplie correctement par Google (que ce soit dans l'exemple du résultat naturel ou dans le cadre de Google Actualités), la troisième condition peut parfois prêter à discussion, notamment quand le contenu visé est si court que la reprise consiste, en réalité, en une copie totale du contenu. Enfin, et surtout, la deuxième condition nous semble extrêmement discutable dans de nombreux exemples de services de Google. En effet, l'exemple du résultat naturel ne s'inscrit pas dans un cadre plus général et structuré. Au demeurant, Google Actualités a connu, sur le terrain de la courte citation, plusieurs batailles souvent perdues ou transigées. En effet, Google reprend quelques lignes de chaque article et agrège le tout. L'AFP avait ainsi ouvert le bal des procès pour la reprise des titres, quelques lignes et photos, la procédure se terminant par une transaction évidemment secrète. Toutefois, dans d'autres pays de droits d'auteur, comme la Belgique ou le Brésil, Google a connu des déconvenues et a été contraint de désindexer les articles de nombreux journaux nationaux. L'argument de la courte citation n'a manifestement pas été reçu par les tribunaux.

En conséquence, Google ne peut pas juridiquement reprendre des contenus de tiers de manière systématique et automatisée sans l'accord des auteurs sur le terrain de la courte citation.

Si le droit d'auteur est très vaste et permet à un juge de dire parfois oui et parfois non à la protection d'une œuvre, il existe toutefois une autre voie : celle du parasitisme. Trois conditions sont nécessaires pour se voir reconnaître responsable d'un fait :

- Faire une faute ;
- Qu'il y ait un préjudice pour quelqu'un ;
- Qu'il y ait un lien de causalité entre la faute et le préjudice.

Ainsi, si la reproduction d'une œuvre n'est pas une contrefaçon, elle peut constituer parfois une faute. Certains se souviendront que les Conditions Générales de Vente du site VentePrivee.com avaient été copiées et utilisées par un concurrent. Ce concurrent a été condamné à 30.000 € pour parasitisme : la Cour d'appel de Paris a estimé que les Conditions Générales de Vente ne sont pas protégées par le droit d'auteur, mais que le fait de les copier est un acte de parasitisme condamnable.

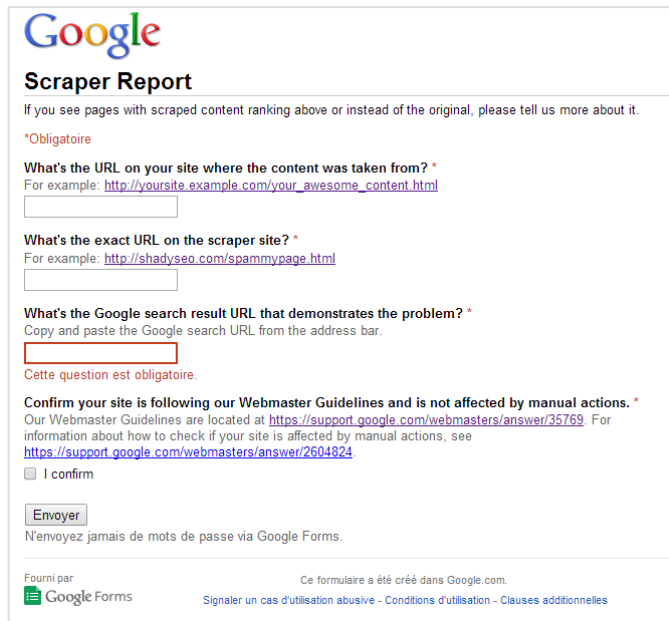
Des internautes peuvent-il « scraper » le contenu de tiers ?

Par ailleurs, tout un chacun peut parfaitement être concerné par la question : soit votre site a été « scrapé », soit vous « scrapez » le contenu d'un tiers.

Les règles juridiques sont exactement les mêmes que pour Google : le droit d'auteur protège un contenu original et le copier est pénalement et/ou civilement répréhensible. L'exception de courte citation peut permettre de copier de courts extraits, mais dans le cadre exposé ci-avant.

Les internautes doivent, eux, faire attention à un point supplémentaire : les règlements de Google. En effet, Matt Cutts a annoncé, en février 2014, la mise en place d'un formulaire « Scraper Report »

(https://docs.google.com/a/google.com/forms/d/1Pw1KVOVRyr4a7ezj_6SHghnX1Y6bp1SOVmy60QjkF0Y/viewform) permettant de signaler à Google des sites qui ont été copiés.



The image shows a Google Forms interface titled "Scrapper Report". At the top is the Google logo. Below it, the title "Scrapper Report" is followed by a brief instruction: "If you see pages with scraped content ranking above or instead of the original, please tell us more about it." A red asterisk indicates a required field. The first question is "What's the URL on your site where the content was taken from?*" with an example URL and a text input field. The second question is "What's the exact URL on the scraper site?*" with an example URL and a text input field. The third question is "What's the Google search result URL that demonstrates the problem?*" with a text input field and a note "Cette question est obligatoire." The fourth question is "Confirm your site is following our Webmaster Guidelines and is not affected by manual actions.*" with a checkbox labeled "I confirm" and a link to Google's Webmaster Guidelines. At the bottom, there is an "Envoyer" button, a note "N'envoyez jamais de mots de passe via Google Forms.", and footer information including "Fourni par Google Forms" and "Ce formulaire a été créé dans Google.com." with links for reporting abuse, terms of use, and additional clauses.

Rien n'interdit de plus, de signaler la copie du site par le biais du formulaire DMCA de Google (<https://support.google.com/legal/troubleshooter/1114905?rd=2-ts=1115655,1282900,1115846,1117010,1697925>) dans la mesure où le scraping porte souvent sur du contenu protégé.

Enfin, et de manière plus large, reste la question de savoir si le scraping est un acte de « negative SEO ». En effet, la copie servile de site tiers peut avoir pour fonction de « pousser » certaines pages en référencement. Nous estimons que la réponse ne peut être que positive car, tant dans l'esprit que dans les manipulations techniques, le scraping peut viser à optimiser de manière illicite le référencement. Au demeurant, la rubrique « Contenu de Qualité » des « Conseils pour Webmasters » est claire à ce titre : le scraping n'apporte rien, peut porter atteinte à des droits de tiers et induit que le contenu du site servile ne sera pas de qualité (et donc répréhensible dans le langage de Google) (<https://support.google.com/webmasters/answer/2721312?hl=en>).

En conséquence, en plus de risquer d'éventuels problèmes juridiques, les internautes scrapant un site pour le copier dans un autre, risquent une sanction de la part de Google en termes de référencement.

Alexandre Diehl

Avocat à la Cour, cabinet Lawint (<http://www.lawint.com/>)

Revue d'Outils Moteurs et SEO

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Cette rubrique signale chaque mois 5 outils (site web, logiciel, plug-ins Firefox...) qui nous ont parus intéressants (ou parfois étonnants...) pour le référencement naturel ou la recherche d'information. Des outils pas obligatoirement très récents, mais parfois méconnus ou qui gagnent à être utilisés selon nous. Chaque outil est décrit en quelques phrases, sous la forme d'une fiche signalétique vous permettant de le tester au plus vite...

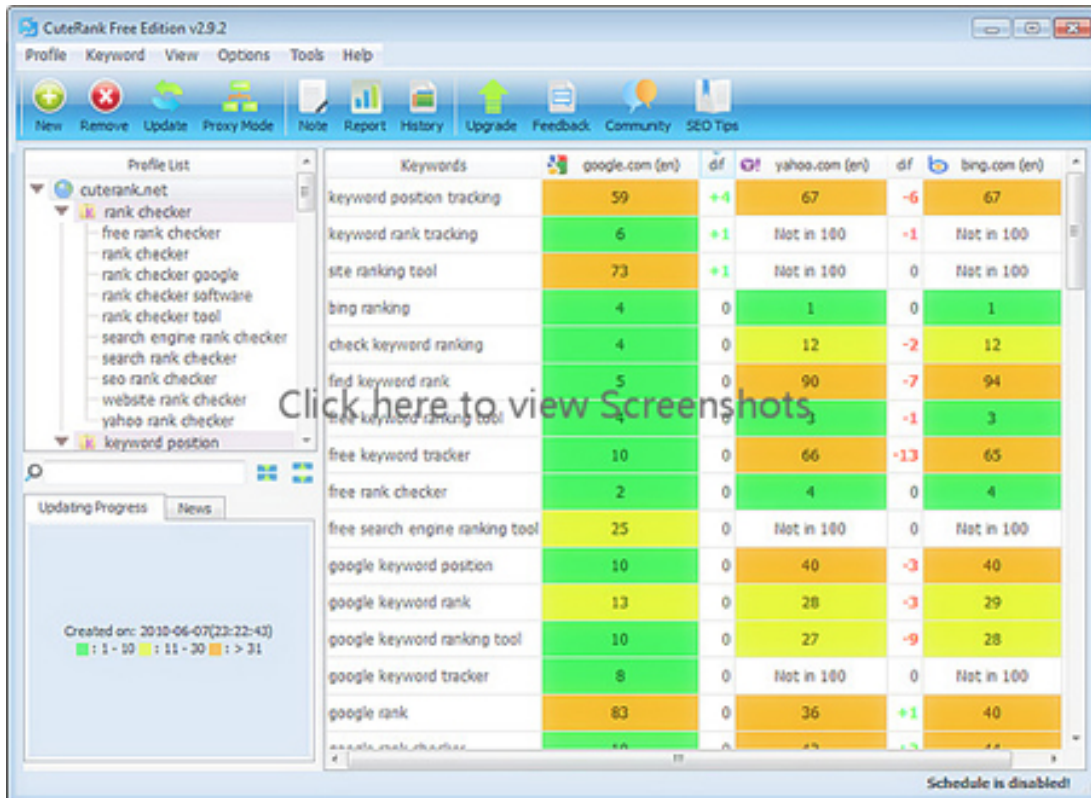
Nom : Cute Rank

URL : <http://cuterank.net/>

Domaine : SEO

Type : Logiciel Windows/Mac

Descriptif : Comme chaque mois, ou presque, un outil de vérification de vos positionnements sur Google. Ici un logiciel qui a le mérite de fonctionner à la fois sur PC et Mac...



Nom : Keyword Strategy Studio

URL : <http://www.goodkeywords.com/internet-strategy-studio/#kwsspro-more>

Domaine : SEO

Type : Logiciel + site web

Descriptif : Un outil permettant de rechercher des mots clés pour votre SEO... A voir aussi les autres outils sur la même page et notamment le "*Keyword and Product Idea Generator*" (<http://www.goodkeywords.com/internet-strategy-studio/#kwideas-more>).

Keyword Ideas	#	Keyword Ideas
From Recent News	1	Domain Name Manager
From Products That Sell	2	Name Manager Solution
From Product Search	3	Name Manager Offer
From Buy & Sell Marketplace	4	Domain Name Manager Solution
From Recent Social Messages	5	Free Domain Name
From Top Ranked Page Titles	6	Domain Name Manager Offer
From Semantic Words	7	Domain Name Registration
	8	Cost Domain Name
	9	Charge Domain Name
	10	Go Daddy Help
	11	Free Domain Name Manager
	12	Cost Domain Name Manager
	13	Advanced Domain Manager
	14	Dns Records Add
	15	Records Add Geopl
	16	Free-of-charge Domain Name

Nom : Microdata Reveal

URL :
<https://chrome.google.com/webstore/detail/microdatareveal/olapakiakkbldaaajcifglndni kpdh?hl=en>

Domaine : SEO

Type : Extensions Chrome

Descriptif : Plusieurs extensions permettent de visualiser d'éventuelles balises sémantiques dans le code HTML d'un page web. *Microdata Reveal* est l'un d'eux, mais on peut également citer *Microdata/JSON-LD sniffer*

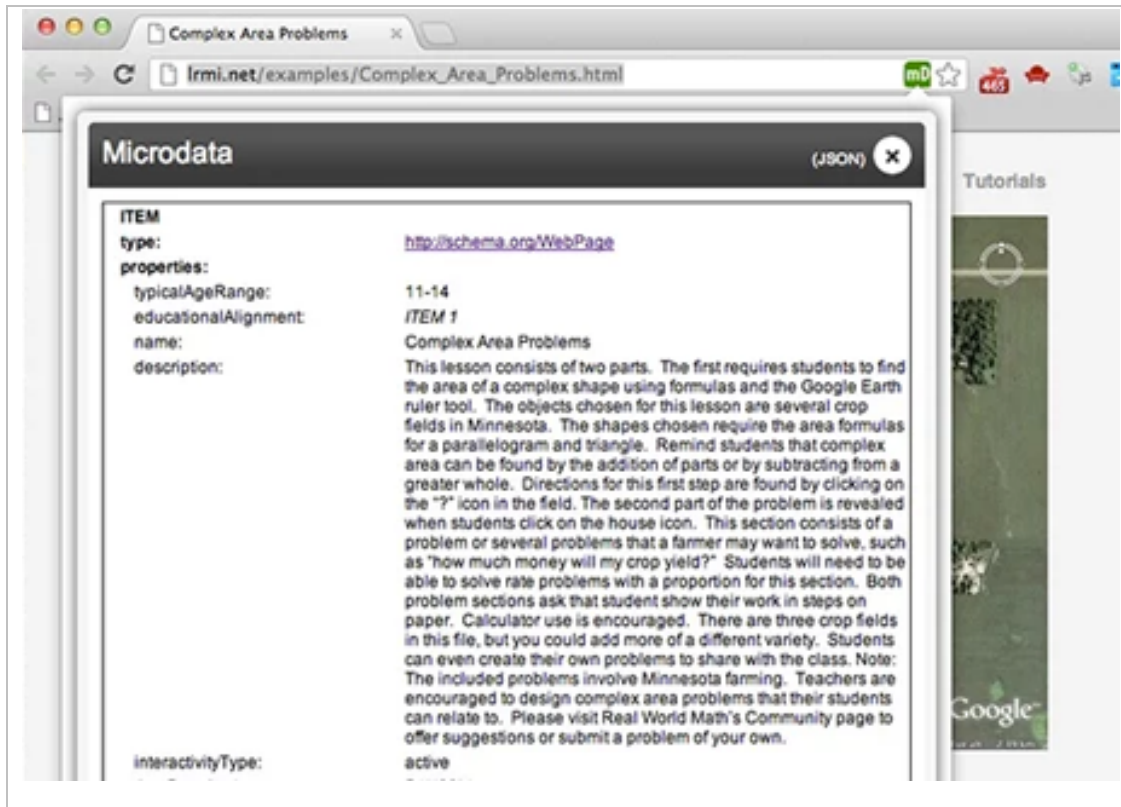
(<https://chrome.google.com/webstore/detail/microdatajson-ld-sniffer/djgonlcfpeededjhpqijdckhgddcfgf?hl=en>), *Semantic inspector*

(<https://chrome.google.com/webstore/detail/semantic-inspector/jobakbebljifplmcapcooffdbdmfdbjh?hl=en>), *META SEO inspector*

(<https://chrome.google.com/webstore/detail/meta-seo-inspector/ibkclpciafdglkjcibmohobjkcfkaef?hl=en>) ou *Green Turtle RDFa*

(<https://chrome.google.com/webstore/detail/green-turtle-rdfa/loggcajcfkpdeoaehclldihfefijjam?hl=en>).

Tous ces outils sont présentés ici : <http://searchengineland.com/see-entities-web-page-tools-help-194710>



Nom : Mnogosearch

URL : <http://www.mnogosearch.org/>

Domaine : SEO et recherche

Type : Logiciel

Descriptif : Code open source pour créer un moteur de recherche web. Site (non officiel et plus mis à jour) en français ici : <http://mnogosearch.free.fr/>. Page Wikipedia : <http://fr.wikipedia.org/wiki/MnoGoSearch>



Nom : SEO Stats

URL : <https://github.com/eyecatchup/SEOstats>

Domaine : SEO

Type : Librairie PHP


Descriptif : Présentation de cet outil par Jean-Benoit Moingt (Pages Jaunes) sur le JDN (<http://www.journaldunet.com/solutions/seo-referencement/outils-seo-gratuits-open-source/seostats.shtml>) : "SEOstats est une librairie PHP incontournable pour ceux qui développent leurs propres outils SEO. SEOstats propose en effet de nombreuses fonctions qui permettent de gagner un temps précieux : récupération du PageRank d'une page, du nombre de pages indexées, résultats de SERPs... SEOstats permet également de récupérer des données d'Alexa, de SEMrush, de Mozscape... Un must !"








SEOstats is a powerful open source PHP library to request a bunch of SEO relevant metrics.

116 commits 4 branches 2 releases 6 contributors

branch: master **SEOstats** / +

Hello 2014

 **eyecatchup** authored on 19 Jan latest commit [d6f8dc189f](#)

 SEOstats	Fix #62: Fix Google+ +1 share count	6 months ago
 example	Update autoload instructions in example files.	7 months ago
 .gitignore	Added .travis.yml	7 months ago
 .travis.yml	Added .travis.yml	7 months ago
 LICENSE	Initial (dev branch) commit of 2.5.2	11 months ago
 README.md	Hello 2014	6 months ago
 composer.json	Add required PHP extensions to composer.json.	7 months ago

Rubrique proposée par Olivier Andrieu, éditeur du site Abondance.com.

Revue d'URL

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé



Le dossier du mois : Authorship et photo d'auteur

Le dossier du mois concerne la disparition de la photo des auteurs dans l'authorship de Google

(<http://www.abondance.com/actualites/20140626-14039-authorship-google-supprime-photo-lauteur-ses-serp.html>). Parmi les dizaines d'articles qui sont sortis à cette occasion, en voici quelques-uns qui ont retenu notre attention :

-> Comment je filoutais sur Google avec l'authorship ?

La fête est finie, les photos d'auteurs ne s'affichent plus sur Google. Ceci ne veut pas dire que l'authorship ne fonctionne plus, mais juste que la photo n'est plus visible...

<http://blog.axe-net.fr/comment-je-filoutais-sur-google-avec-lauthorship/>

-> Pourquoi Google arrête-t-il d'afficher les photos Authorship ?

Comme nous le savons tous maintenant, Google a fait le choix de retirer la photo des auteurs présente dans les résultats de recherche...

<http://www.maxime-denizon.fr/pourquoi-google-arrete-photo-autorship/>

-> Hey Google, t'as un problème ? Tu veux ma photo ?

Si vous faites une veille digne de ce nom, vous n'avez pas pu passer à côté de l'information du jour (non ce n'est pas la qualification de l'EDF suite à un « splendide » match face à l'équateur)...

<http://www.visibilite-referencement.fr/blog/suppression-authorship>

-> Google Removes Author Photos From Search: Why And What Does It Mean?

The search world was rocked by yesterday's announcement by Google's Webmaster trends analyst John Mueller that Google will be removing all author photos and Google+ view counts from search results on both desktop and mobile. This change will roll out to all users over the next few days...

<http://searchengineland.com/google-removes-author-photos-search-mean-195236>

-> Google Announces the End of Author Photos in Search: What You Should Know

Even so, it came as a surprise when John Mueller announced Google is dropping authorship photos from most search results...

<http://moz.com/blog/bye-bye-author-pics>

Revue d'URL en français... ■ ■

-> Link building : 3 mauvaises habitudes à perdre

Tous les gens qui sont spécialisés dans le référencement depuis des années savent que le temps où l'obtention de centaines de liens en quelques heures pouvait améliorer le référencement d'un site est révolu (en tout cas, si on veut être référencé sur le long terme)...

<http://www.referenceur.be/link-building-3-mauvaises-habitudes-a-perdre-13134.html>

-> Les avantages de Google Webmaster Tools

L'utilisation de Google Analytics est devenue assez commune, celle de Google Webmaster Tools beaucoup moins. Pourtant les avantages de cet outil sont multiples bien que méconnus. Voici 6 avantages liés à l'utilisation de Google Webmaster Tools...

<http://www.carrementplus.net/2014/07/07/les-avantages-dutiliser-google-webmaster-tools/>

-> Quelle est la Formule pour Créer le Titre Parfait d'un Article ?
Savez-vous que 8 personnes sur 10 lisent un titre mais que seulement 2 sur 10 lisent le reste ? Si vous êtes un lecteur régulier du blog, nous avons abordé plusieurs fois ce sujet: le titre est l'élément le plus important d'un billet car il détermine le taux de clic et influence le comportement de vos lecteurs quant à la poursuite éventuelle de votre article...

<http://www.emarketinglicious.fr/blogging/formule-titre-parfait-article>

-> Choisir ses mots-clés pour le référencement de son site
Il faut bien avouer que la recherche de mots-clés n'est pas une tâche des plus amusantes dans le référencement d'un site ou d'une page et c'est pourtant une étape décisive...

<http://www.pole-position-seo.com/2014/07/choisir-ses-mots-cles-pour-le-referencement-de-son-site.html>

-> L'audit SEO ou la psychanalyse digitale (vu au SMX Paris 2014)
C'est un sujet extrêmement vaste, complexe et non linéaire. L'idée n'est pas de vous apprendre comment réaliser un audit par vous-même. 7500 mots ou une heure de présentation ne suffisent pas. C'est un vrai métier, je le répète inlassablement...

<http://www.htitipi.com/blog/audit-seo-smx.html>

-> Moteur de recherche : 11 alternatives à Google
Google est le moteur de recherche absolu en France et dans beaucoup de pays dans le monde. Et soyons honnêtes, il fonctionne quand même très bien...

<http://www.webmarketing-com.com/2014/06/24/28742-moteur-recherche-11-alternatives-google>

-> Les 5 erreurs à éviter après une formation en référencement
Erreur 1 : J'ai pris des notes, je vais les ranger dans l'étagère 1, dossier 4, intercalaire 8... En quelques heures, vous avez ingurgité une grande quantité d'informations...

<http://optimiz.me/les-5-erreurs-eviter-apres-formation-en-referencement/>

-> Guest blogging et SEO : les limites à ne plus franchir
Il y a quelque mois, Google fustigeait bruyamment les excès du guest blogging. Des pénalités ont suivi. Faut-il pour autant oublier totalement ce moyen d'obtenir des liens ?...

<http://www.journaldunet.com/solutions/seo-referencement/guest-blogging-et-seo-les-penalites-de-google.shtml>

-> Google Pingouin : comment le surmonter ?
Mes deux derniers articles relataient sur des outils permettant de voir si un site a été touché par une quelconque mise à jour de Google et la façon de récupérer son trafic dû à une pénalité de Panda. Dans cette logique, nous allons voir cette fois-ci comment se remettre d'une pénalité de Google Pingouin...

<http://ludismedia.com/google-pingouin-comment-le-surmonter/>

-> Méthodologie : analyse de la structure d'un site
J'embarque assez souvent dans mes prestations une analyse de la structure interne des sites clients. On va parler aujourd'hui de ma méthodologie et je vais vous présenter quelques-uns des outils (développés maison) que j'utilise...

<http://www.didcode.com/tuto/structure-interne.html>

-> Comment poster un message efficace sur Google+
Pour réussir sur Google+, vous devez attirer l'attention des internautes sur vos posts. Il y a des techniques à exploiter et des erreurs à ne pas commettre. Voici des conseils pour optimiser vos posts Google+...

<http://www.webrankinfo.com/dossiers/google-plus/post-parfait>

-> Comment Bing choisit-il les balises Title de vos pages web ?
Si Bing n'a pas la même influence que Google sur les internautes européens, ce moteur de recherche est reconnu comme le deuxième le plus utilisé en Europe et dans le monde d'après les statistiques de Statcounter. Il peut donc tout à fait apporter une visibilité

supplémentaire à votre entreprise, même si c'est dans une moindre mesure par rapport à Google...

<http://www.referenceur.be/comment-bing-choisit-balises-pages-web-13085.html>

-> Le standard du fichier robots.txt fête ses 20 ans, l'occasion de lister des conseils
Le fichier robots.txt, qui permet de bloquer le crawl des moteurs de recherche pour certaines URL ou parties d'un site, fête ses 20 ans : il fut créé le 30 juin 1994 par Martijn Koster. Voici plein de conseils que vous ne connaissiez peut-être pas...

<http://www.webrankinfo.com/dossiers/indexation/20-ans-robots-txt>

-> 13 conseils pour rédiger du contenu efficace en référencement

La rédaction est un des éléments les plus importants du référencement naturel. Pour se positionner au mieux dans les résultats des moteurs de recherche, il est effectivement nécessaire d'optimiser son contenu pour les moteurs de recherche tout en l'écrivant pour les lecteurs...

<http://www.referenceur.be/13-conseils-rediger-contenu-efficace-referencement-13060.html>

-> Comment sortir d'une sanction manuelle Google

Vous devez être certainement déjà au courant que Supernova a été sanctionné manuellement par Google début mai. Ce sont en fait mes 3 annuaires qui sont tombés simultanément, Supernova, Private et Luxury ! Une sanction en demi-mesure puisque le PR de Supernova est passé de 5 à 2. Côté positions par contre, ça a été la chute libre avec une disparition des mots clés principaux sur les 10 premières pages minimum. Inutile de préciser que les Adsense se sont effondrées elles aussi...

<http://www.hdv-referencement.fr/penalite-google-pour-liens-sortants-factices/>

-> Découvrez notre livre blanc « Guide du débutant en SEO »

Si vous êtes un fidèle lecteur de Carrément Plus Net, vous avez, sans doute, dû percevoir un certain changement dans notre ligne éditoriale...

<http://www.carrementplus.net/2014/06/25/decouvrez-notre-livre-blanc-guide-du-debutant-en-seo/>

-> Ces services de Google méconnus du grand public

Gmail, Maps, Google Watches... Le géant de l'Internet est devenu bien plus qu'un simple moteur de recherches. Mais tous ses services ne sont pas forcément connus du grand public. Développer votre culture musicale, aider les ONG, services d'alerte, visite de Mars...

<http://www.journaldunet.com/ebusiness/le-net/google-service-inconnus/>

-> 10 pièges à éviter en référencement naturel

Comme nous le savons tous, le référencement naturel n'est pas une science exacte. Bien qu'il s'agit souvent d'une affaire de bon sens et d'une connaissance approfondie des outils que l'on utilise. Il n'est pas rare d'observer des erreurs grossières par manque d'informations ou de logique. Je ne blâme personne, car moi le premier, je me suis déjà fait abuser par des pratiques ou des idées reprises par la majeure partie de la communauté SEO. En bref, n'hésitez pas à réfléchir par vous-même et à approfondir chaque idée avant de l'appliquer...

<http://www.maxime-denizon.fr/10-pieges-eviter-en-referencement-naturel/>

-> 8 moteurs de recherche alternatifs, insolites et utiles

Le premier réflexe aujourd'hui, lorsque l'on cherche une information, c'est de la chercher sur internet, et tout particulièrement sur Google. Cependant, il existe des alternatives au moteur de Mountain View qui sont souvent plus performantes selon le type d'information que l'on recherche. Mais il existe de nombreux moteurs de recherche. C'est pour cela qu'on a décidé de vous faire une sélection de quelques uns de ces moteurs dits « alternatifs », parfois étonnants et souvent très utiles...

<http://www.pullseo.com/8-moteurs-de-recherche-alternatifs-insolites-et-utiles/>

-> Yandex ne fournit plus les mots-clés recherchés (not provided)

Après Google et Yahoo, Yandex bloque lui aussi des données sur les referrers, si bien que les webmasters ne peuvent plus connaître les mots-clés générateurs de trafic (SEO).

<http://www.webrankinfo.com/dossiers/yandex/cryptage-requetes>

-> 5 conseils pour du référencement local à succès

A l'heure actuelle, de plus en plus d'entreprises cherchent à gagner une certaine visibilité locale pour leur site internet. Cela signifie que celui-ci devra se positionner de manière idéale pour des recherches incluant des régions ou des villes spécifiques...

<http://www.referenceur.be/5-conseils-referencement-local-succes-13014.html>

-> Top 10 des erreurs SEO

Voici la liste des 10 erreurs à ne pas commettre en référencement naturel...

<http://www.alexco.fr/conseils-referencement-naturel,top-10-erreurs-seo.html>

-> Web responsive : Google publie un kit de démarrage

Pour encourager davantage les développeurs Web à adopter le design responsive, Google a publié un ensemble de modèles disponibles au téléchargement...

<http://pro.clubic.com/creation-de-site-web/langage-programmation/html-5/actualite-710977-web-responsive-google-kit-demarrage.html>

-> Comment utiliser l'algorithme Colibri de Google dans son référencement ?

En septembre 2013, Google a annoncé l'existence de son nouvel algorithme Colibri (ou Hummingbird) destiné à améliorer l'expérience utilisateur. Comment ? En fournissant des réponses plus pertinentes dans les résultats de recherche à partir de l'analyse de requêtes conversationnelles, c'est-à-dire de phrases interrogatives précises. Comment fonctionne cet algorithme et quelles actions faut-il mettre en place ?...

<http://www.webzeen.fr/internet/referencement/comment-utiliser-lalgorithme-colibri-de-google-dans-son-referencement/13328>

-> Les Graphes du webmarketing

Les graphes du web, qu'est-ce c'est ? A quoi servent-ils ? Quels sont les outils (NodeXL, gephi..) qui exploitent la data ? Dans un monde numérisé, ou tout se compte, décompte, se regroupe, s'agrège, il est parfois difficile d'avoir une représentation globale de tous ces bruits qui environnent notre présence sur internet. Représenter des acteurs dans un but analytique, des objets, des contenus et leur interrelation pourrait être un des objectifs de ces projections graphiques, mais pas seulement, le but peut être aussi esthétique, avoir des bulles dessinées au lieu de barres ou autres camemberts cela peut être plus harmonieux...

<http://www.mauricelarger.com/matrice-a-bulles-pour-comprendre-les-donnees/>

-> Le Pagerank thématique : qu'est-ce que c'est ?

On aborde de plus en plus souvent le thème du Pagerank sémantique (ou Pagerank thématique) lorsque l'on parle de référencement. Ce Pagerank est impacté par l'optimisation sémantique Onpage (sur le site), et influencera le netlinking (Offpage)...

<http://blog.axe-net.fr/pagerank-thematique/>

-> Créer son simulateur de robot d'indexation en PHP

Je suis actuellement en pleine écriture d'un prochain livre sur le référencement technique et le suivi-audit SEO avec Alexandra Martin, notre Miss SEO Girl nationale. Pour cet ouvrage, je code beaucoup en PHP (mais aussi en Javascript, jQuery voire en VBScript pour ASP parfois... :D) afin de vous fournir un maximum d'informations techniques qui nous manquent souvent lorsque nous travaillons...

<http://blog.internet-formation.fr/2014/06/creer-son-simulateur-de-robot-dindexation-en-php/>

-> 15+ sites d'images libres de droits pour votre blog

Lorsque l'on rédige du contenu régulier sur son blog, il n'y a que deux façons de trouver des images pour illustrer son site...

<http://www.pullseo.com/15-sources-images-libres-de-droits/>

-> SEO en agence vs SEO chez l'annonceur

Faisant suite à la table ronde à laquelle j'ai eu le plaisir de participer lors du SMX Paris avec Béatrice Payonne et Mael Montarou, et ayant œuvré dans les deux mondes, je vous livre aujourd'hui ma réflexion personnelle sur les différences entre le SEO coté agence et le SEO coté annonceur, du point de vue du référenceur essentiellement. Mais les entreprises sauront également en tirer des enseignements ;)...

<http://www.sebastien-billard.fr/seo/post/SEO-en-agence-vs-SEO-chez-l-annonceur-18-06-2014>

-> SMX Paris : comment sortir d'une pénalité manuelle ou l'éviter
Que faire pour sortir d'une pénalité déclenchée manuellement par Google pour liens artificiels ? Comment éviter une telle pénalité ? Une session du SMX Paris a donné des conseils...

<http://www.journaldunet.com/solutions/seo-referencement/seo-au-smx-paris-sortir-d-une-penalite-manuelle-ou-l-eviter.shtml>

-> Le nettoyage des mauvais liens

Vous-même ou votre prestataire en référencement avez été un peu trop optimiste dans le passé. Votre site a été sanctionné par Google par une pénalité manuelle (liens factices ou artificiels, participation à un système de liens) ou risque de l'être par un filtre...

<http://www.axenet.fr/referencement/nettoyer-vos-liens/>

-> Larry Page : l'histoire inédite du vrai fondateur de Google

Vous connaissez le parcours de Steve Jobs et de Mark Zuckerberg. Mais, connaissez-vous celui de Larry Page, le créateur de Google ?...

<http://www.journaldunet.com/ebusiness/le-net/l-histoire-inedite-de-l-incroyable-retour-de-larry-page.shtml>

-> Backlink via redirection 302 et pénalité manuelle

Attention dans vos analyses de backlinks artificiels, tenez-compte aussi des redirections, même les temporaires (code 302).

<http://www.webrankinfo.com/dossiers/breves/penalite-backlink-redirection-302>

-> Infographie : 6 mythes concernant les moteurs de recherche et le SEO

Quand on ne s'y connaît pas vraiment et qu'on décide petit à petit de s'intéresser aux moteurs de recherche et au référencement naturel, il y a beaucoup de chances de tomber sur des idées fausses ainsi que sur certains mythes qui ont la vie dure dans le domaine. Pour tenter d'y voir plus clair, l'agence spécialisée en webmarketing Veravo a réalisé une infographie où 6 idées fausses sont démystifiées...

<http://www.referenceur.be/infographie-6-mythes-moteurs-de-recherche-seo-13153.html>

-> Google sitelinks : toujours aussi confus

Vous vous êtes sans doute tous habitués aux sitelinks de Google. Les siteslinks, ce sont ces liens supplémentaires qui s'affichent en plus du lien principal pour un site dans les résultats de Google...

<http://blog.axe-net.fr/google-sitelinks-toujours-aussi-confus/>

-> Trouver le juste équilibre

En mai dernier, la Cour de justice de l'Union Européenne a établi un "droit à l'oubli". Aujourd'hui, nous avons publié une tribune de David Drummond, Senior Vice-Président et Directeur Juridique de Google, expliquant notre point de vue sur cet arrêt et notre réponse. Le texte a été publié en France dans Le Figaro, au Royaume-Uni dans le Guardian, en Allemagne dans le Frankfurter Allgemeine Zeitung et en Espagne dans El Pais. Vous pouvez retrouver cette tribune dans son intégralité ci-dessous...

<http://googlefrance.blogspot.fr/2014/07/trouver-le-juste-equilibre.html>

-> Google Quality Rating Guidelines Version 5 – 31 mars 2014

Voici la dernière version du guide Google « Quality Rating Guidelines » fourni par Google à ses Quality Raters. Cette version date du 31 mars 2014 et peut-être téléchargé ici...

<http://www.seoblackout.com/2014/07/10/google-quality-rating-guidelines-2014/>

Revue d'URL en anglais...

-> Everybody Needs Local SEO

If you work in the SEO industry, you need to understand how to do Local SEO. Seriously.. I'm not kidding here... If you're sitting there thinking "Um, no... not really" - then you're exactly the person I'm writing this post for...

<http://moz.com/blog/everybody-needs-local-seo>

-> How Does Google Index Tweets?

In this study we explored the way that Google indexes tweets. The reason we embarked on this was to determine the likelihood that Google might use signals from Twitter for ranking purposes, but we found lots of other interesting information in the process. Spoiler alert...

<http://www.stonetemple.com/how-does-google-index-tweets/>

-> 5 Bad Link Building Habits To Break Now

For those who have been in the SEO industry for a while, you may remember the glory days of link building — you simply submitted your site/article/address/name, etc., and had 100 links within 24 hours. That sure was awesome...

<http://searchengineland.com/5-bad-link-building-habits-break-now-195608>

-> SEO Website Audits: Everything You Need to Know Part II

In a follow-up to an earlier column on what a site audit is in general, today I am going to spend a bit more time helping with the understanding of what a good audit is and why you would get with a tools audit vs. a human one...

<http://searchenginewatch.com/article/2353656/SEO-Website-Audits-Everything-You-Need-to-Know-Part-II>

-> How To Integrate SEO Ranking Tools With Analytics Tools

Search Engine Optimization (SEO) tools and more general web analytics solutions are both critically important for site analysis, but they don't often work well together. This is a how-to guide for integrating SEO ranking tools with analytics tools...

<http://searchengineland.com/integrating-seo-ranking-tools-analytics-tools-195617>

-> Panda 4.0 Help: Google Clarifies Tremors, 3 Analogies for Diagnosing Hits

Panda 4.0 rolled out six weeks ago and we've learned a lot since then. I provided initial P4.0 findings on my blog, which covered the power of the nuclear option, subject matter experts surging, forums benefitting, Phantom victims recovering, and more. Then, after analyzing more companies impacted by Panda 4.0, I wrote a follow-up post here on Search Engine Watch explaining additional findings based on my analysis...

<http://searchenginewatch.com/article/2353139/Panda-4.0-Help-Google-Clarifies-Tremors-3-Analogies-for-Diagnosing-Hits>

-> Breaking Down Page Speed Events For SEO Gain

We all know page speed is important to SEO, but how can we accurately — and quickly — assess if our performance is good enough? Identify where the issues actually lie?...

<http://searchengineland.com/breaking-page-speed-events-seo-gain-195369>

-> Back To Basics: 5 Fundamentals Of Link Building That Will Never Go Away

SEO is a totally different beast today than it was just a few years ago, and many would argue that link building is the aspect that's suffered the biggest beating since then...

<http://searchengineland.com/back-basics-5-fundamentals-link-building-will-never-go-away-195206>

-> Let Google Know About Your Negative SEO

A thread at Google Webmaster Help asks a question that goes unanswered for the most part, "Negative SEO. How can we let the Webmaster Team know?"...

<http://www.seroundtable.com/inform-google-negative-seo-18776.html>

-> Best practices and common mistakes in SEO

For those who missed our #io14 presentation, join us in this office hours hangout for SEO tips on improving your site's appearance in search results. The 3 topics we'll cover are: 1. Mobile 2. Multi-regional and multilingual sites 3. Monitoring your site for spam...

<https://www.youtube.com/watch?v=z51qbs7wLx8&feature=em-uploademail>

-> The 10 Best Ways to Generate Traffic Without Google

Having a mix of trophy and long-tail phrases ranking in Google search engine results pages (SERPs) is the equivalent of manna from heaven to any online business. There's nothing better, if you have it, and nothing worse, if you lose it...

<http://searchenginewatch.com/article/2352408/The-10-Best-Ways-to-Generate-Traffic-Without-Google>

-> Google Begins Removing Search Results to Comply With 'Right to be Forgotten'

Google has started removing search results, following the European Court of Justice landmark "right to be forgotten" ruling in May...

<http://searchenginewatch.com/article/2352388/Google-Begins-Removing-Search-Results-to-Comply-With-Right-to-be-Forgotten>

-> Bing: How We Pick Your Title Tag

Microsoft Bing published a blog post on how they decide on what title tags to publish in their search results...

<http://www.seroundtable.com/bing-title-tags-18762.html>

-> 130 Questions to Ask When Choosing a Local Search Automation Tool

Choosing a local search automation vendor can be a complicated and intimidating decision for companies with hundreds or thousands of brick-and-mortar locations...

<http://searchenginewatch.com/article/2352144/130-Questions-to-Ask-When-Choosing-a-Local-Search-Automation-Tool>

-> 3 Killer Survey Questions to Generate Useful Content For a Link Building Campaign

The quality of the content you create is influenced by the quality of the material on which it is based. However, that material may not be complete – a client's knowledge of their own customers is often far from perfect...

<http://searchenginewatch.com/article/2351870/3-Killer-Survey-Questions-to-Generate-Useful-Content-For-a-Link-Building-Campaign>

-> Analyzing 11,555 Questions Asked by SEOs: The Moz Q&A Meta Study

Sometimes we don't need to travel to exotic linked data sources to discover treasure troves of precious information about our audience's desires, aspirations, fears, and complaints...

<http://moz.com/blog/fearing-and-loathing-and-dreaming-in-mozland>

-> The 10 Worst Link Building Assumptions

Over the past few years, one of the biggest problems I've faced is educating clients about search marketing – about what's reasonable to expect and what isn't in the world of link building...

<http://searchengineland.com/10-worst-link-building-assumptions-194353>

-> How Content Quality Analysis Works With SEO

The Periodic Table of SEO Ranking Factors panel at the recent SMX Advanced show stimulated some fresh thinking for me on the role of content analysis and SEO...

<http://searchengineland.com/content-quality-analysis-seo-194314>

-> Google Panda 4.0 Hit Sites Blocking CSS & JavaScript

Joost published a blog post showing how some of the sites he hit by Google Panda 4.0 and in one case, unblocking the CSS & JavaScript resulted in the site returning to its normal rankings...

<http://www.seroundtable.com/google-panda-4-block-css-js-18741.html>

-> Are You Worried About Negative SEO? Here Is How To Take Action

Not only have webmasters claimed negative SEO works, they are saying it is much easier to implement these days than in the past...

<http://www.seroundtable.com/negative-seo-google-download-latest-links-18739.html>

-> How Responsive Design Increases the Results of Your Online Marketing

We all hear about responsive design. But as a marketer, sometimes we are only focused on the specific tasks, such as creating content, developing email campaigns, building links, and updating social media networks. The fact of the matter is, if you or your clients do not have a responsive design in place, then all of the online marketing strategies we implement will not bring the full potential of results...

<http://searchenginewatch.com/article/2351176/How-Responsive-Design-Increases-the-Results-of-Your-Online-Marketing>

-> Google Working On Super Secret hreflang Webmaster Feature

Google seems to be working up a new webmaster feature around rel-alternate-hreflang. rel-alternate-hreflang is the attribute to communicate to Google if there are alternative language based web pages for the specific web page the attribute is placed on...

<http://www.seroundtable.com/hreflang-google-test-18734.html>

-> Does Google Automatically Flag Sites Using The noscript Tag?

Last week at SMX Advanced, Maile Ohye of Google warned SMX attendees of using the noscript tag on their web site. She said something like it is an "automatic" issue of some sorts...

<http://www.seroundtable.com/google-noscript-tag-18729.html>

-> Google: Penalized Sites Don't Get Crawled Less Often

Yesterday in a Google Webmaster Hangout with Google's John Mueller I asked John if Google crawls penalized sites less frequently and/or slower. He responded that Google typically does not slow the crawl for penalized sites...

<http://www.seroundtable.com/google-crawl-speed-penalty-18720.html>

-> A Step-By-Step Guide To Link Building For Boring Industries Or Products

Within the current search landscape, most search engine optimisation campaigns have become increasingly content-led. The majority of the campaigns that I work on revolve largely around the production and placement of high quality content on top-tier sites...

<http://searchengineland.com/link-building-approach-boring-b2b-industries-193619>

-> 5 Reasons a Site Hit by Google Penguin Won't Recover

When Google launched the Penguin algorithm April 24, 2012, many sites who had relied strongly on low-quality link building were severely affected. I regularly hear stories from business owners who were hit hard over two years ago and haven't been able to recover...

<http://searchenginewatch.com/article/2350646/5-Reasons-a-Site-Hit-by-Google-Penguin-Wont-Recover>

-> Feeding the Hummingbird: Structured Markup Isn't the Only Way to Talk to Google

I used to laugh at the idea of Hummingbird optimization. In a recent poll, Moz asked nearly 300 marketers which Google update affected their traffic the most. Penguin and Panda were first and second, followed by Hummingbird in a distant third...

<http://moz.com/blog/feeding-the-hummingbird>

-> More On Google Sandbox 2.0

A couple weeks ago we brought up some conversation on the black hat forums around a new Google Sandbox...

<http://www.seroundtable.com/google-sandbox-again-18833.html>

-> Google Panda: Understanding the Technical Inner Workings

The concept of a search engine is very simple. Database the location of pages and return the most relevant results. Primitive search engines only understood words on the pages in their database and returned pages that included those searched words. Google essentially

came along with the ability to reorder those search results based on how many other pages on the Internet linked to them...

<http://searchenginewatch.com/article/2354803/Google-Panda-Understanding-the-Technical-Inner-Workings>

-> Google Updates Quality Rating Guidelines - EAT

Jennifer Slegg reports that Google has completely revamped the Google Quality Rating Guidelines, which is used for Google contractors to manually review search results and send that feedback to Google's team to improve the overall algorithm...

<http://www.seroundtable.com/google-updates-quality-rating-guidelines-18830.html>

Rubrique proposée par Olivier Andrieu, éditeur du site Abondance.com.

Merci pour votre lecture...

Pour toute suggestion au sujet de cette lettre professionnelle : olivier@abondance.com