

**Lettre professionnelle
Recherche et Référencement**

Recherche & Référencement

Numéro 183

Juillet-Août 2016



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 & Référencement

Lettre d'actualité professionnelle et mensuelle
sur la recherche d'information et le référencement/SEO

Numéro 183 - Juillet-Août 2016

Au sommaire ce mois-ci :

● Éditorial	3
● Comment utiliser au mieux l'analyse de la recherche dans la Search Console (1ère partie) (12 pages)	4
● Comment prédire quel résultat a le plus de chances d'être cliqué dans les SERP ? (6 pages)	16
● SEO et Prestashop : les fondamentaux (9 pages)	22
● Cocon.se : voir la structure de son site avec les yeux de Google (10 pages)	31
● Curation : comment réussir sa stratégie ? (2ème partie) (12 pages)	41
● Gephi et l'analyse de données pour le SEO (2ème partie) (11 pages)	53
● Le droit d'opposition opposé à Google (4 pages)	64
● Revue d'outils moteurs et SEO (2 pages)	68
● Revue d'URL (7 pages)	70

Lettre parue le **vendredi 15 juillet 2016** (prochaine parution : jeudi 15 septembre 2016)

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

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.

Editorial

Par Olivier Andrieu



Bienvenue dans ce numéro 183 de la lettre "Recherche et Référencement" pour les mois de juillet et août 2016. Comme vous le savez peut-être, aucun numéro n'est publié en août (vacances obligent) et notre prochain rendez-vous sera donc mi-septembre !

Dans ce numéro (volumineux avec 7 articles !), vous trouverez des articles pour tous les goûts à lire sur la plage. :-). De la curation à la représentation graphique des sites web avec Gephi et Cocon.se en passant par la prédiction des clics dans les SERP, Prestashop, l'Analyse de la Recherche dans la Search Console et la notion de "droit d'opposition", vous devriez trouver ce mois-ci de quoi satisfaire vos lectures estivales avec un éclectisme que nous essayons de vous proposer chaque mois. Mais vous avez également le choix de penser à autre chose qu'à Google, la recherche d'information et le SEO. Les vacances, ça sert également à ça pour revenir avec les batteries rechargées à bloc pour la rentrée !

Et nous vous y attendons avec une lettre professionnelle toujours plus complète et des articles toujours plus pertinents. En attendant, passez de bonnes vacances, profitez-en bien et à très bientôt !



Olivier Andrieu
Rédacteur en chef

Comment utiliser au mieux l'analyse de la recherche dans la Search Console (1ère partie)



Par Philippe Yonnet

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

L'outil Analyse de la Recherche (ou Search Analytics en anglais) de la Search Console de Google n'est pas toujours utilisé à bon escient par de nombreux webmasters. Pourtant, il peut s'avérer très intéressant en termes d'analyse de votre trafic SEO en de nombreuses occasions. Cet article en deux parties vous aide donc à mieux appréhender cette fonctionnalité grâce à de nombreux exemples concrets. Exploration...

Cela fait maintenant un peu plus d'un an que Google a déployé dans la Search Console le rapport « search analytics » (« Analyse de la recherche » en français) et qui a remplacé l'ancien rapport « search queries » (« requêtes de recherches »). Plus récemment, le rapport s'est enrichi de nouvelles fonctionnalités. Cet outil fournit des données clé pour optimiser son site web. Mais son utilité est parfois sous-estimée par les webmasters, soit parce qu'ils n'ont pas su découvrir certaines données masquées dans l'interface, soit parce qu'ils ne font pas confiance aux données fournies dans les rapports.

En réalité, les informations obtenues grâce à cet outil sont d'un grand intérêt opérationnel. Pour vous en convaincre, nous allons explorer quelques cas d'utilisation pratique, dans un article découpé en deux parties. Dans cette première partie, nous allons voir comment exploiter les données pour des utilisations basiques. Dans une seconde partie, nous verrons quelques exemples d'utilisation avancée, mais également comment se servir de l'API, et des possibilités d'intégration des données dans d'autres outils de Google comme Google Analytics.

Visite guidée rapide de l'interface

Pour accéder au rapport, il faut se connecter à votre compte « Search Console », et ouvrir le menu « trafic de recherche » : le lien vers « Analyse de recherche » (Search Analytics en anglais) apparaît.

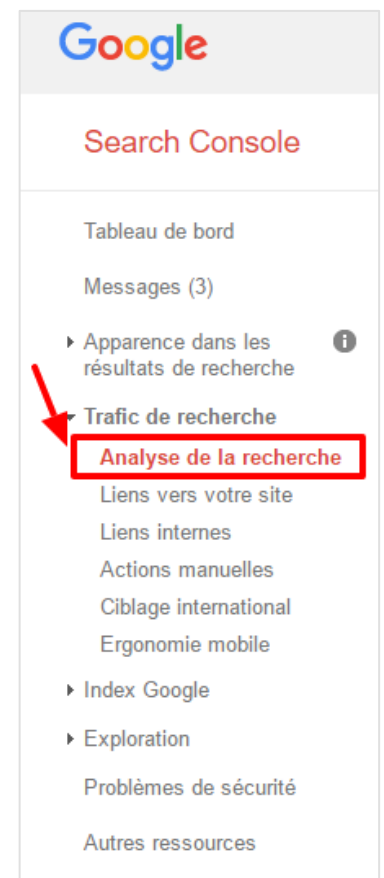


Fig.1. L'outil « analyse de recherche » se trouve dans le menu « trafic de recherche » de la Search Console

En cliquant sur le lien, vous accéderez à l'interface de l'outil « Analyse de recherche ».

Il permet d'afficher 4 données clé (des « métriques ») sur des rapports produits en fonction de 6 dimensions. Les dimensions définissent les lignes des tableaux, et les métriques les colonnes.

Les rapports peuvent ensuite être filtrés en fonctions de plusieurs critères. L'ensemble est donc en mesure de produire un grand nombre de rapports différents.

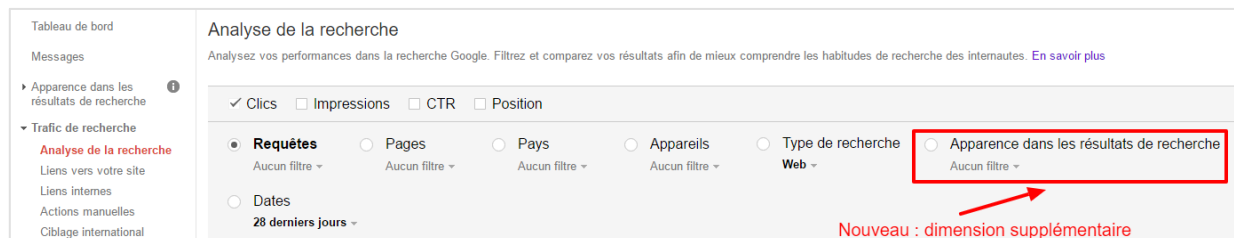


Fig.2. Le « sélecteur » du générateur de rapport « analyse de la recherche »

Les 4 « métriques » disponibles : clics, impressions, CTR, position

L'outil vous donne accès à quatre « métriques » que vous pouvez décider d'afficher ou non à l'aide d'une case à cocher sur les graphiques et les tableaux produits par l'outil :

- Les clics ;
- Les impressions ;
- Le CTR (le ClickThrough Rate, le ratio clics/impressions) ;
- Et la position (moyenne).

Nous reviendrons dans le prochain article sur ce que mesurent précisément ces métriques, qui sont souvent mal comprises. Et nous nous intéresserons en particulier aux raisons pour lesquelles un outil de web analytics mesure un nombre de sessions provenant des moteurs de recherche différent des données de clics affichées dans les rapports search analytics.

6 dimensions + une nouvelle

L'outil propose de générer des rapports segmentés en fonction de 6 dimensions :

- Les requêtes (les « mots clés ») ;
- Les pages (les pages d'atterrissage) ;
- Les pays (depuis quel pays la requête a-t-elle tapée) ;
- Les appareils (la requête a-t-elle été effectuée sur un mobile, une tablette, ou un ordinateur de bureau ?) ;
- Le type de recherche (la requête a-t-elle été tapée sur le moteur normal - web - ou sur Google Images, ou sur Google videos ?) ;
- Les dates.

Une septième dimension apparaît depuis quelques semaines sur les comptes de sites qui ont des données structurées : la dimension « apparence dans les résultats de recherche ». Elle permet d’avoir un rapport sur les pages AMP, ou sur l’affichage des « résultats enrichis ».

Un graphique et un tableau

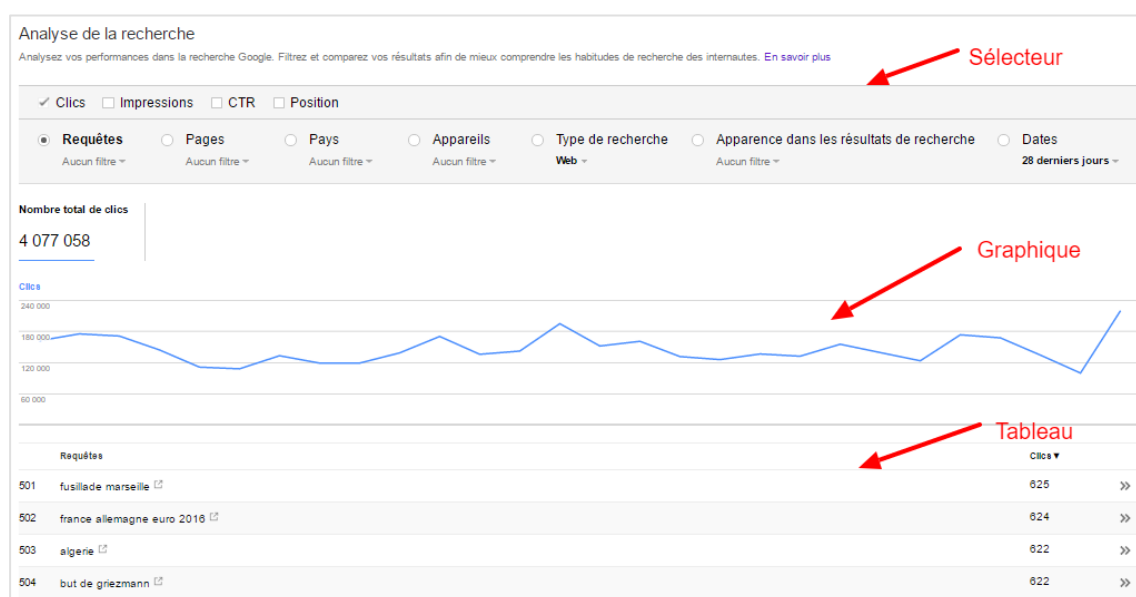
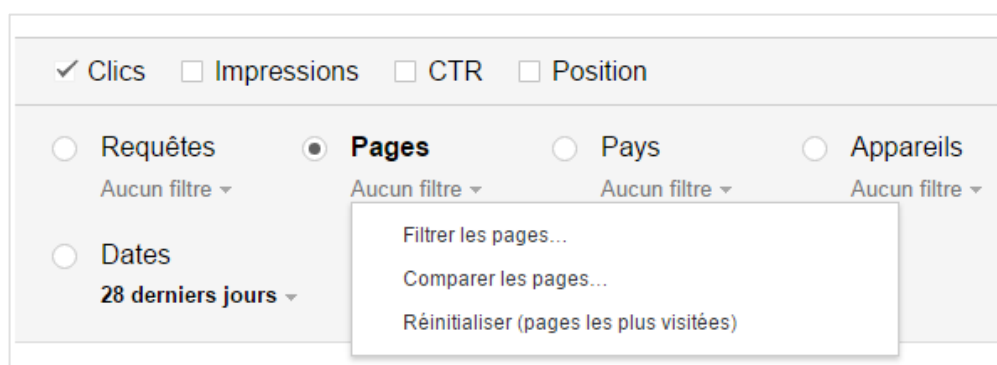


Fig.3. Les 3 zones de l’outil « analyse de la recherche » : sélecteur, graphique, tableau.

L’outil permet d’afficher, une fois la dimension choisie, un graphique avec une courbe pour chaque métrique cochée, et un tableau, avec une colonne pour chaque métrique cochée.

Les filtres et les comparaisons pour les dimensions

Chaque dimension peut être filtrée, on peut donc par exemple créer un rapport pour les pages d’atterrissage commençant par une certaine chaîne de caractères (un répertoire) ou un terme particulier (un nom de produit par exemple).



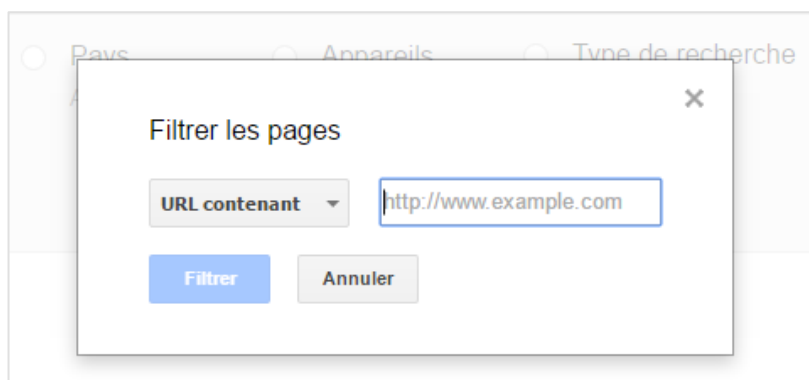


Fig.4. Les options de filtrage d'une dimension : ici filtrage des URL contenant une chaîne de caractères.

On peut aussi choisir de comparer des valeurs de dimension, par exemple les données de deux pages d'atterrissage, ou de deux mots clés (fig. 5).

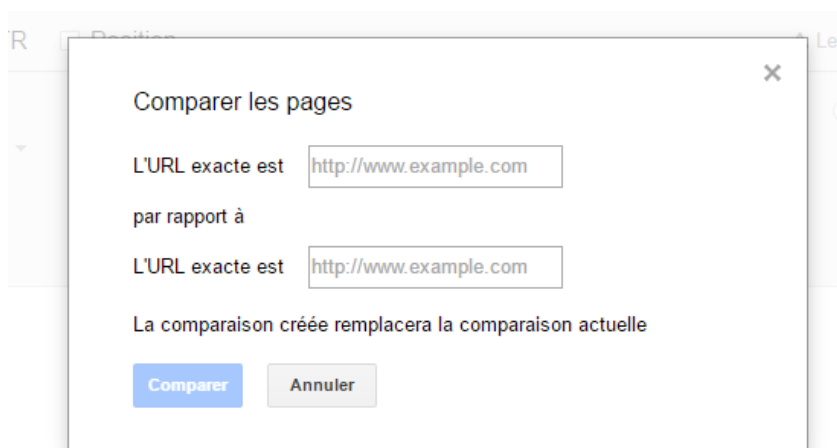


Fig.5. Option de comparaison de deux valeurs de dimension : ici deux URL.

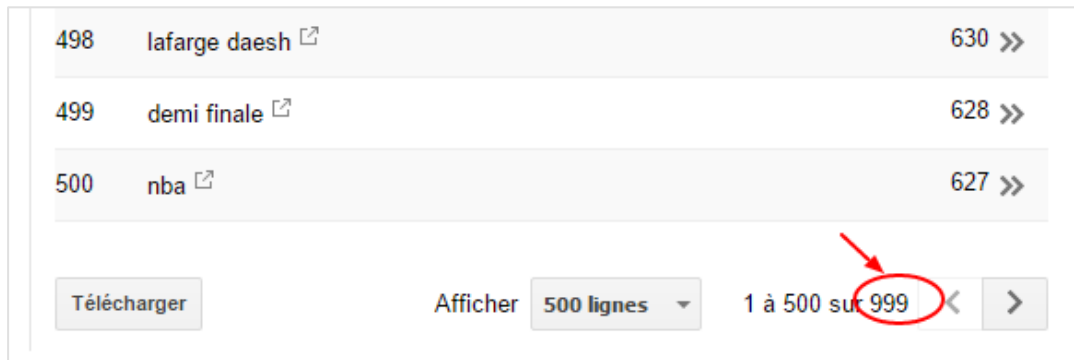
Une fois une valeur de filtre ou de comparaison choisie pour une dimension, ce filtre persiste, même si on choisit d'afficher les données pour une autre dimension. Cette fonctionnalité permet de faire des analyses extrêmement utiles, comme décrit dans l'exemple fourni plus loin.

Pensez à conserver vos données !

Attention : les données les plus récentes fournies par l'outil remontent toujours à un minimum de 3 ou 4 jours. Impossible d'avoir les données de la veille, ce qui interdit un monitoring en temps réel qui doit être effectué par un outil de web analytics.

Par défaut, les rapports présentent les données sur les derniers 28 jours. L'outil permet au maximum de remonter à 90 jours, mais pas un jour de plus. Pour conserver les données sur une période plus longue, il est indispensable de télécharger les rapports au format CSV (ou Google Docs) pour les conserver.

Le bouton de téléchargement se situe en bas à gauche du tableau de données.



498	lafarge daesh	630 >>
499	demi finale	628 >>
500	nba	627 >>

Télécharger Afficher 500 lignes 1 à 500 sur 999 < >

Fig.6. Le requêteur de l'outil « search analytics » est limité à 999 lignes. L'ancien outil (les « requêtes de recherche ») renvoyait beaucoup plus de données sur les mots clés ou les pages d'atterrissage.

Le bouton permet de télécharger les données correspondant à la combinaison segments(dimension)+filtre/comparaison+métriques à afficher définies dans le sélecteur. En revanche, l'outil search analytics limite le nombre de lignes renvoyées à 999 ! Pour avoir plus de données, il faut :

- Soit combiner des rapports filtrés/segmentés différemment (pas toujours pratiques) ;
- Soit passer par l'API qui permet d'obtenir jusqu'à 5000 lignes pour chaque requête (nous décrirons la méthode dans le prochain article).

Utiliser les données de Search Analytics pour améliorer son référencement

L'une des utilisations les plus intéressantes des données de l'outil d'analyse de recherche consiste à découvrir quels sont les performances du site sur une requête donnée d'une part, et d'autre part, quels sont les performances d'une page d'atterrissage donnée.

Quelles pages ont reçu des visites sur un mot clé donné et pourquoi ?

Pour effectuer cette analyse, nous allons commencer par choisir la dimension « dates », pour obtenir les données sur 90 jours et faire une analyse sur le maximum de clics et d'impressions. Puis nous choisissons la dimension « requêtes » pour définir un filtre sur un mot clé.

Nombre total de clics	Nombre total d'impressions	CTR moyen	Position moy.
37 492 177	1 173 322 909	3,2%	5,3

Fig.7. L'étape 1 consiste à sélectionner une période de 90 jours dans les options de dates...

Fig.8. Etape 2 : filtrage des requêtes sur un mot clé.

Ensuite, nous allons nous intéresser aux pages d'atterrissage, en choisissant le segment « pages ».

Fig.9. Etape 3 : sélection de la dimension « pages » pour le rapport. Les filtres persistent sur les dimensions requêtes et dates.

Notons que les filtres sur les autres dimensions (dates, requêtes etc...) sont conservés.

Le résultat fournit les pages du site qui ont été affichées sur cette requête et les métriques associées.

Pages	Clics ▼	Impressions	CTR	Position
1 -direct-brexit	92 192	1 466 409	6,29%	5,1
2 -brexit-bureaux-vote-ouvert-referendum	36 170	615 357	5,88%	5,4
3 -direct-brexit-brexit-souleve-interrogation-toute-planete-estime-francois-hollande	12 191	170 465	7,15%	3,6
4 -brexit-pourquoi-petition-3-millions-signatures-pose-question-va-faire-objet-enquete	8 163	88 844	9,19%	6,3
5 -brexit-consequences-si-royaume-uni-quitte-union-europeenne	6 770	80 534	8,41%	3,9

Fig.10. Le rapport sur la dimension « pages » obtenu.

Ce rapport est utile pour identifier les meilleures pages d'atterrissage pour une requête donnée, et, en comparant les performances des pages, pour identifier ce qui leur permet d'avoir de meilleures performances. Par exemple, un meilleur CTR est-il associé à l'optimisation du contenu de la page, ou le « snippet » est-il plus optimisé ?

Si une requête fait apparaître un grand nombre de pages d'atterrissage différentes, on peut également identifier qu'il serait utile de créer une page de regroupement, ici une page dossier « BREXIT » qui pourrait constituer une meilleure page d'atterrissage que les 142 pages identifiées ici !

Si aucune page ne sort : c'est sans doute le signe qu'il faut en créer une « ad hoc », sauf si c'est juste la conséquence d'un mauvais positionnement de l'ensemble des pages du site.

Cette analyse permet aussi de se rendre compte que Google est susceptible d'afficher, sur une même requête, une variété parfois impressionnante de pages de votre site, certaines optimisées, d'autres non. Stabiliser ce comportement sur des requêtes stratégiques (top tail) peut représenter un chantier intéressant à lancer.

Sur quels mots clés apparaît une page d'atterrissage donnée ?

L'analyse inverse est également riche d'enseignements : toujours sur des données extraites sur une période fixée à 90 jours, on va chercher à obtenir les requêtes sur lesquelles cette page apparaît. On choisit la dimension « Pages » et on paramètre un filtre sur une URL correspondant à un article sur la mort du chanteur Prince.

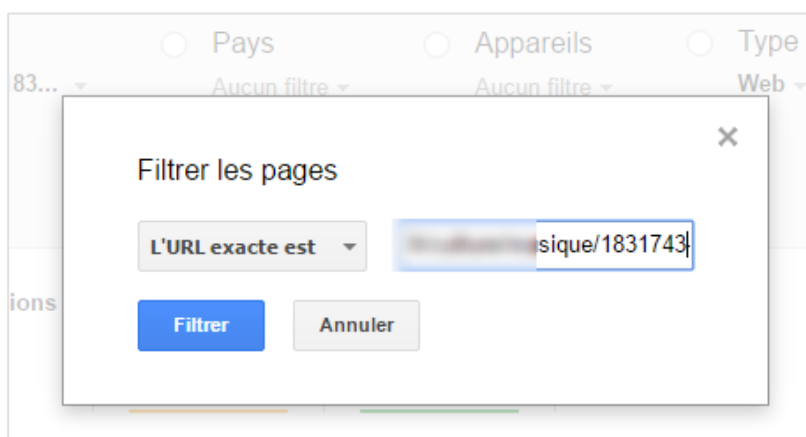


Fig.11. Etape 1 : choix de l'url exacte à analyser

Ensuite, on sélectionne la dimension « Requêtes » et on obtient la liste des requêtes qui a fait apparaître cette page.

	Requêtes	Clics▼	Impressions	CTR	Position
1	prince ↗	57 024	1 090 747	5,23%	6,2
2	prince mort ↗	43 354	331 074	13,09%	3,5
3	mort de prince ↗	10 336	122 762	8,42%	5,4
4	prince est mort ↗	7 966	133 393	5,97%	4,9
5	deces de prince ↗	4 336	32 500	13,34%	3,8
6	prince chanteur ↗	3 412	88 184	3,87%	5,6
7	de quoi est mort le chanteur prince ↗	2 930	54 437	5,38%	2,9
8	chanteur prince ↗	2 619	60 902	4,3%	5,2
9	prince deces ↗	2 171	19 267	11,27%	4,0
10	de quoi est mort prince ↗	2 150	20 055	10,72%	3,4
11	le chanteur prince ↗	2 084	34 780	5,99%	3,8
12	prince mort chanteur ↗	1 967	26 797	7,34%	3,9
13	chanteur prince décédé ↗	1 635	52 877	3,09%	3,1
14	mort prince ↗	1 473	22 563	6,53%	4,9
15	mort de prince chanteur ↗	1 375	12 424	11,07%	4,6

Fig.12. Etape 2 : sélection de la dimension « requêtes » et obtention du rapport donnant les données sur les mots clés pour une page d'atterrissage donnée.

Ce rapport peut être utilisé de multiples façons :

- Les requêtes identifiées peuvent servir de base pour une recherche d'expressions similaires, connexes, associées, proches sémantiquement, à l'aide d'outils (l'outil de suggestion de mots clés adwords par exemple). Ces expressions, une fois qualifiées (il s'agit de trouver celles qui sont les plus tapées) peuvent être ajoutées dans une version réécrite de la page, afin d'augmenter le trafic généré par cette page d'atterrissage.
- D'une façon plus générale, on peut se servir de ces rapports pour suivre l'efficacité de chantiers d'optimisation : notamment en cherchant à positionner les pages sur un plus grand nombre de requêtes, en enrichissant le contenu ou les balises SEO habituelles (title, H1, anchor texts...).
- Plus finement, on peut chercher à comprendre pourquoi, alors que la page se positionne bien en tête des résultats sur une requête donnée, la conversion n'est pas bonne (les internautes ne choisissent pas le résultat).
- On peut également chercher à comprendre pourquoi une page se positionne loin dans les classements sur une requête pour laquelle la page est optimisée.

Le rapport permet de faire des constats, et sert donc de point de départ pour des analyses poussées.

Les autres utilisations basiques des données Search Analytics

Voici quelques cas d'utilisation où les données des rapports d'analyse de recherche peuvent s'avérer d'un grand secours.

Les analyses post-pénalité / changement d'algorithme / baisse de trafic

Dès qu'un évènement de ce type survient, les rapports filtrés sur une comparaison de dates sont d'un grand secours pour mesurer clairement les indicateurs avant et après, mesurer l'impact réel et commencer à comprendre ce qui a changé :

- Quelles pages d'atterrissage ont été impactées ;
- Sur quelles requêtes le site a perdu des positions ;
- La situation est-elle la même sur desktop ou sur mobile ? ;
- Etc.

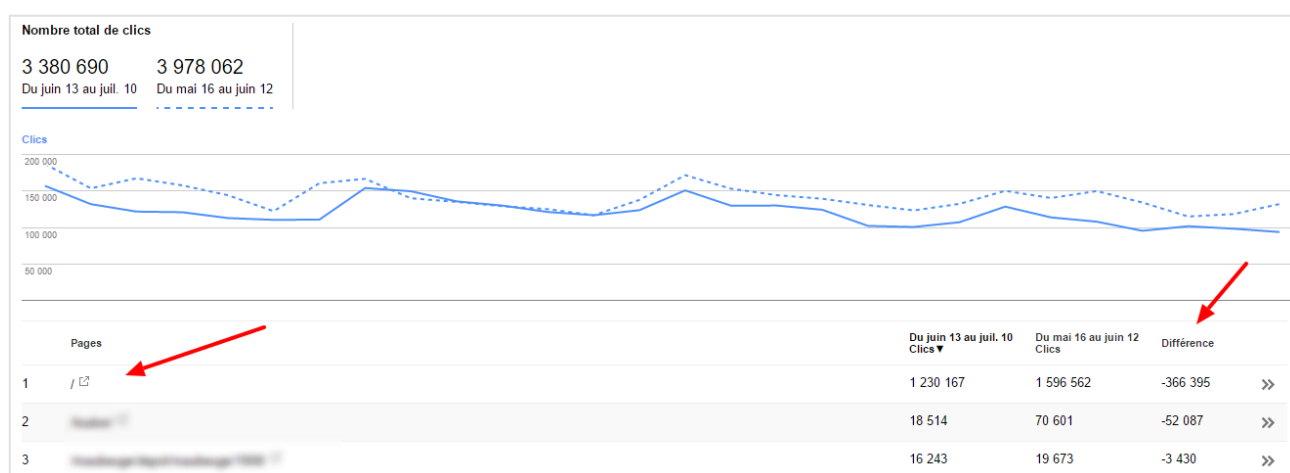


Fig.13. Analyse d'un décrochage de trafic : la comparaison d'un mois sur l'autre montre une chute de trafic significative sur la page d'atterrissage. L'analyse des différents segments montre que la chute est générale sur les mots clés brandés. La vraie cause est la saisonnalité et l'arrêt d'une campagne télé. Notez la colonne « différence » qui n'existait pas dans l'outil précédent (requête de recherches).

L'analyse des positions sur mobile et desktop

Le référencement sur la recherche mobile est devenu stratégique, et les pages de résultats sont devenues très différentes. Il est donc indispensable de bien analyser les différences de performance et de positions entre les deux types d'appareils.

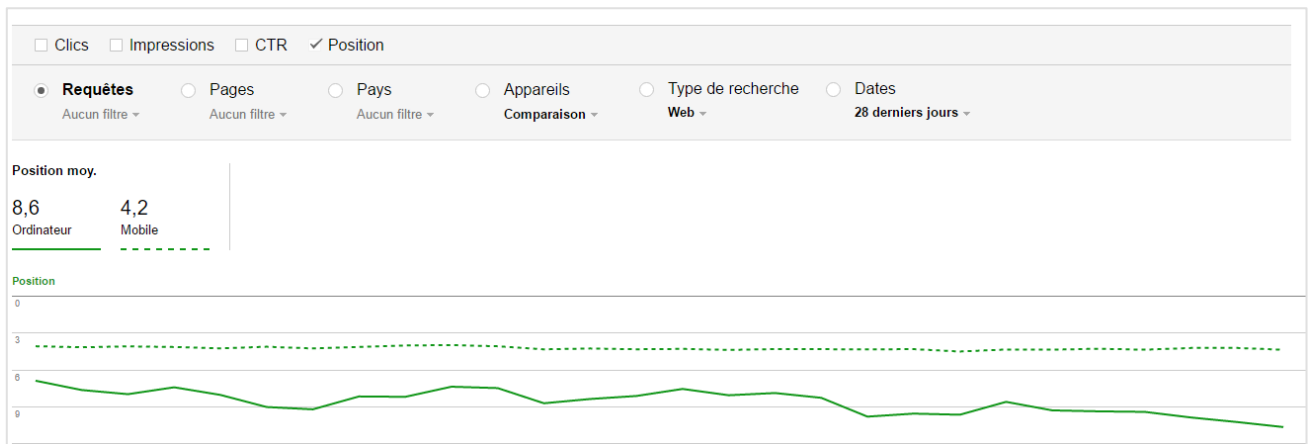


Fig.14. Une comparaison mobile vs desktop sur les positions occupées sur les requêtes met en avant une différence de comportement fondamentale pour ce site : en contexte mobile, les positions occupées sont meilleures et stables. En contexte desktop : les positions occupées sont de plus en plus lointaines, avec une forte influence sur le trafic généré.

Notons qu'il manque encore des données sur les recherches vocales. Mais des officiels chez Google ont laissé entendre que ce rapport était en préparation.

L'analyse du trafic google images (ou video)

Des problèmes de paramétrage ou de taggage empêchent souvent d'isoler correctement le trafic « image » dans les outils de web analytics. Mais il est très facile d'obtenir ces données dans un rapport de search analytics.

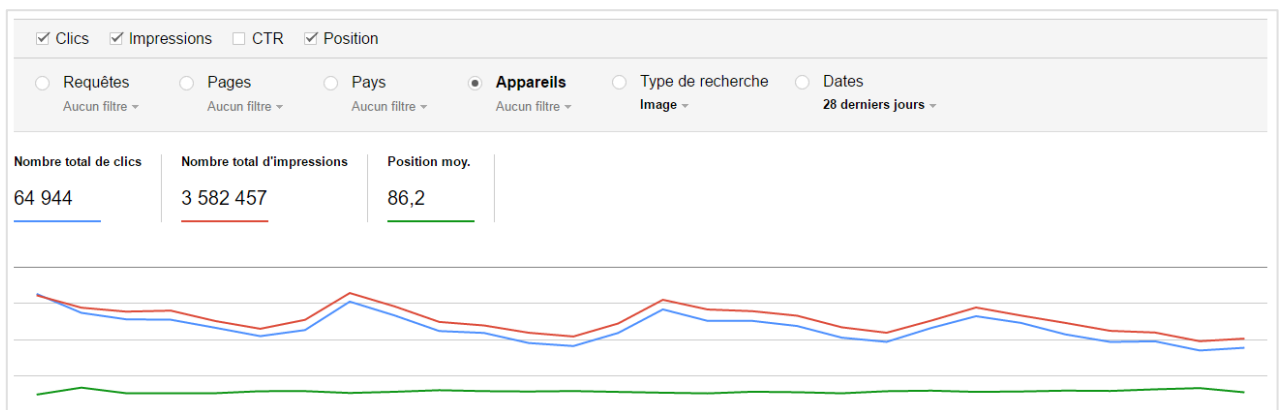


Fig.15. Rapport filtré sur les requêtes sur Google Images

Obtenir les VRAIES positions

Enfin, l'outil Search Analytics est le seul outil à fournir une information fiable sur la visibilité de vos URL, en vous renvoyant la position moyenne occupée par vos pages dans les résultats. En effet, selon le contexte, la géolocalisation, l'historique de recherche, le dispositif utilisé, le moment de la recherche, etc., la position dans les pages de résultat évolue. Les positions fournies par les outils de positionnement ne

sont que des positions instantanées, produites par une requête effectuée à un instant T dans un contexte irréaliste : pas d'historique, en mode déconnecté etc.

Les positions fournies par le rapport d'analyse de la recherche sont donc plus proches de la réalité observée par les internautes.

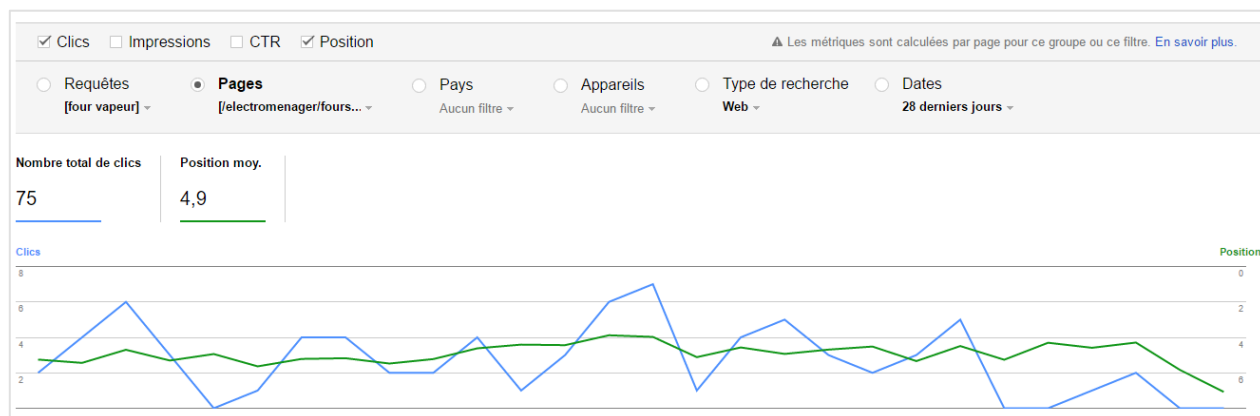


Fig.16. Analyse de l'évolution de la position moyenne sur une requête: le rapport est filtré à la fois sur une requête et sur une page, ce qui permet de voir l'évolution de la position moyenne occupée par une URL sur cette requête.

L'analyse de la recherche : le point de départ de nombreuses analyses SEO

Avec un peu d'habitude, on apprend très vite à utiliser l'interface de l'outil Search Analytics au quotidien pour de nombreux usages basiques :

- Analyser une page d'atterrissage donnée ;
- Travailler une requête ;
- Monitorer les performances SEO ;
- Analyser un changement brutal ;
- Optimiser pour un pays donné, un device donné.

En fait les données de cet outil sont tout bonnement indispensables pour bien optimiser le référencement d'un site. Mais pour exploiter correctement et de manière complète ces données, il faut ensuite apprendre à utiliser des fonctionnalités plus avancées de l'outil. C'est que nous verrons dans l'article du mois prochain qui proposera les informations suivantes :

- Comment regrouper les données de plusieurs comptes Search Console ?
- Comment exploiter les données pour optimiser les performances SEO d'un site marchand ?
- Comment utiliser l'API pour obtenir plus de résultats et créer des outils de monitoring performants ?
- Comment intégrer les données de Search Analytics dans Google Analytics et comment utiliser les données combinées des deux outils pour contourner le problème du not provided ?
- Comment interpréter correctement les données des rapports de Search Analytics ?

Biographie et Liens Utiles

Annonce officielle de lancement du rapport « search analytics » sur le blog pour les webmasters :

<https://webmasters.googleblog.com/2015/05/new-search-analytics.html>

La page support de Google sur le rapport Search Analytics :

<https://support.google.com/webmasters/answer/6155685>

Article de Jennifer Slegg à propos de la nouvelle fonctionnalité de comparaison de mots clés :

<http://www.thesempost.com/googles-search-analytics-now-allows-keyword-comparisons/>



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

Comment prédire quel résultat a le plus de chances d'être cliqué dans les SERP ?



Par Guillaume et
Sylvain Peyronnet

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Bien sûr, le but et l'objectif d'un moteur de recherche est de proposer les résultats les plus pertinents répondant à une requête donnée. Pourtant, si ces liens ne sont pas cliqués, le moteur a échoué dans sa (re)quête. Il est donc important de tester comment une SERP (page de résultats) est analysée par l'internaute et de prendre en compte deux notions importantes : la pertinence et l'attractivité du résultat renvoyé. Cet article décrit des travaux de recherche dans ce domaine, travaux qui ont certainement un impact sur la façon dont nous proposons nos pages à des outils comme Google.

L'article de ce mois présente un résultat scientifique qui a été présenté à la conférence WSDM 2016 (Web Search and Data Mining) qui a eu lieu à San Francisco en février 2016. La précision est importante, car WSDM est, avec la très prestigieuse conférence WWW, l'une des deux rencontres les plus importantes des acteurs du search. Cette année, Jeff Dean (le spécialiste du deep learning chez Google), Yoelle Maarek (VP recherche de Yahoo !) ou encore Lars Backstrom (responsable du fil d'actualités chez Facebook) faisaient partie des orateurs invités.

C'est donc l'endroit pour découvrir les techniques que les moteurs mettront probablement en place dans les prochains mois, et dans cet article, nous allons vous présenter un travail qui a pour objectif de re-ranker, c'est-à-dire de reclasser les SERP en utilisant les caractéristiques des snippets pour prédire ce qui a toutes les chances d'être cliqué par les utilisateurs d'un moteur de recherche. Ce travail est dû à Jiepu Jiang et James Allan, et est décrit dans l'article [1].

Prédire la probabilité d'être cliqué, pour quoi faire ?

On pourrait oublier facilement l'objectif d'un moteur de recherche et penser qu'il ne sert qu'à classer quelques pages web (10 par exemple) dans l'ordre de décroissance de pertinence supposée par rapport à une requête. Hors, il n'en est rien : le but d'un moteur de recherche est de faire en sorte qu'un résultat pertinent soit vu par l'utilisateur. De ce point de vue, si le moteur met en avant une page très pertinente, mais que le lien vers cette page n'est pas cliqué, alors l'objectif n'est pas rempli, c'est un échec pour le moteur. *A contrario*, mettre en avant une page non pertinente via un lien qui sera cliqué est tout autant un échec pour le moteur.

Pour résoudre ce problème de correspondance entre les liens présentés, la pertinence des pages associées et le clic sur ces liens pour accéder à ces pages, plusieurs angles d'attaque peuvent être considérés. On peut par exemple travailler sur l'amélioration de la qualité des snippets, permettant aux utilisateurs de prendre une décision plus informée avant de cliquer ou non. Cette approche est cependant très peu efficace : Yilmaz et ses coauteurs (voir l'article [2]) ont en effet montré que la probabilité de cliquer sur un mauvais résultat restait de 49% dans ce contexte. On peut également éviter le problème en répondant directement à la requête (si il n'y a plus de nécessité de cliquer sur un résultat, on élimine le problème des clics malheureux). Bien sûr, cette astuce ne fonctionne que pour des requêtes directement informationnelles comme par exemple « quelle est la date de l'épreuve de philosophie du bac 2016 ? ».

L'idée des travaux de Jiang et Allan est finalement assez simple : quand deux pages sont de pertinences raisonnablement similaires, on déclassera la page qui a un plus haut taux d'erreur de clics ou de skips (nous verrons un peu plus loin ce que sont ces taux). L'idée est simple, mais la réalisation l'est moins : pour implémenter cette idée, il faut en effet analyser de nombreuses caractéristiques des snippets présentés par un moteur de recherche, et faire une analyse de corrélation entre ces caractéristiques et la satisfaction des utilisateurs.

Les notions clés des travaux de Jiang et Allan

Pour quantifier la qualité d'une page de résultats d'un moteur de recherche, Jiang et Allan utilisent deux notions complémentaires : la pertinence, qui est la notion classique pour exprimer qu'une page satisfait le besoin informationnel sous-jacent à une requête, et l'attractivité, qui est la suspicion de pertinence que peut avoir un utilisateur en voyant le snippet d'un résultat donné. Bien évidemment, pour évaluer la pertinence, il faut avoir cliqué sur le résultat correspondant, et donc une page pertinente a tout intérêt à être également attractive. Au final, la pertinence et l'attractivité sont souvent corrélées, mais peuvent être en désaccord (page pertinente et non attractive, ou page attractive mais non pertinente).

En regard de ces deux notions, les auteurs de l'article identifient deux types d'erreurs totalement complémentaires : l'erreur de clic (« click error » en anglais) et l'erreur de saut (« skip error » en anglais). Une erreur de clic a lieu lorsqu'un utilisateur clique sur un résultat qui va s'avérer non pertinent tandis qu'une erreur de saut arrive lorsqu'un utilisateur ne va pas cliquer sur un résultat qui aurait été pertinent pour lui. Dans la suite, on va noter p_1 la probabilité de cliquer sur la page P_1 , et r_1 la pertinence de cette page.

En étudiant les propriétés mathématiques associées à ces notions, Jiang et Allan proposent plusieurs règles de classement pour obtenir des SERPs de meilleure qualité :

1. Si $p_1 > p_2$ et $r_1 > r_2$ alors on classe P1 avant P2.
C'est le cas le plus facile : si P1 est plus pertinente et plus attractive, alors on va classer P1 en premier, c'est totalement naturel.
2. Si $r_1 > 0$ et $r_2 \simeq 0$ alors on classe P1 avant P2.
Une page pertinente est toujours proposée avant une page non pertinente.
3. Si $r_1 \simeq r_2 \simeq 0$ et $p_1 < p_2$ alors on propose P1 avant P2.
En effet, P1 a moins de chance d'être cliquée, donc on la propose en priorité vu qu'aucune page n'est pertinente.
4. Dans les autres cas, on propose en premier la page qui a la plus grande valeur de $r * p$. Ainsi, si $r_1 * p_1 < r_2 * p_2$ on proposera d'abord p_2.

A la lecture de ces règles, une question doit vous venir aussitôt à l'esprit : comment fait-on pour les vérifier, sachant qu'on ne peut connaître la valeur de la probabilité de clic qu'une fois que les utilisateurs auront cliqué ? C'est le problème de l'œuf et de la poule : pour calculer la probabilité il faut avoir fait un classement visible des utilisateurs, et pour faire le classement on a besoin des probabilités.

Pour pallier ce problème, Jiang et Allan vont remplacer la probabilité de clic par une prédiction de cette probabilité, et c'est là que les statistiques et le machine learning rentrent en jeu.

Prédire la probabilité de clic

Prédire la probabilité de clic (qu'on va souvent appeler CTR chez les SEO) est un problème qui a surtout été étudié pour les annonces sponsorisées (voir par exemple l'article [3]). Ici le contexte n'est pas le même, et la prédiction se fait avec des caractéristiques spécifiques, que nous allons maintenant voir, en même temps que la corrélation avec le taux de clic.

C'est l'information qui va intéresser les lecteurs SEO de cet article, que nous vous délivrons maintenant, et qui va vous permettre de créer des snippets qui seront plus cliqués que d'habitude !

- **Caractéristiques indépendantes de la requête tapée par l'utilisateur.**
Le premier ensemble de caractéristiques étudié par Jiang et Allan est celui des caractéristiques génériques du snippet comme sa longueur, la longueur de l'URL, etc. Ces critères ne dépendant pas de la requête, ils sont généralement faiblement corrélés à la probabilité de clic, que la page soit pertinente ou pas. Cependant, certaines corrélations restent informatives même si faibles. Par exemple, les pages qui ont la plus grande proportion de contenu réel (c'est-à-dire de contenu entre les balises `<p>`) sont les plus cliquées. Techniquement cela veut dire que si vous avez trop de menus, footer, etc. sur vos pages, cela se retranscrira

dans le taux de clic au niveau du moteur. Une autre corrélation intéressante est négative : les pages pertinentes ont un taux de clic qui est inversement proportionnel à la longueur du titre. Plus ce dernier est long moins les pages sont cliquées.

- **Caractéristiques dépendantes de la requête tapée par l'utilisateur.** Le second ensemble étudié est celui des critères dépendants de la requête tapée. On y trouve des éléments très standard comme par exemple le nombre ou la proportion de termes de la requête dans le titre, snippet, l'URL voire dans certaines balises (on pense par exemple au H1 ou H2). On y trouve également des critères plus complexes comme par exemple le nombre de fenêtres de 10 mots consécutifs qui contiennent les termes de la requête, ou encore comme la vraisemblance de la requête, qui est une forme de vraisemblance d'indépendance, c'est-à-dire un score proportionnel à la probabilité qu'un humain lambda tape cette requête. En terme de corrélation, les plus fortes sont liées aux proportions de présence des termes de la requête dans le titre, l'URL puis dans le snippet (les éléments sont dans l'ordre décroissant de corrélation). En résumé, si vous êtes SEO, sans surprise la présence de la requête dans URL, titre et snippet augmente les chances que les utilisateurs du moteur cliquent sur vos liens. Par ailleurs, la corrélation la plus forte est avec le nombre de fenêtres de 10 mots qui contiennent la requête (voir précédemment), mais c'est sans surprise puisque plus ce nombre est grand plus la page a des chances d'être pertinente et donc plus elle va être cliquée de part sa pertinence d'abord.
- **Caractéristiques dépendantes de la session de l'utilisateur.** Il s'agit des caractéristiques portant sur le comportement de l'utilisateur d'une requête à l'autre : a-t-il reformulé la requête précédente en rajoutant, en enlevant et en gardant certains termes ? Si oui, dans quelle proportion ? Les statistiques sont assez claires, la reformulation d'une requête pour l'affiner va permettre d'améliorer la probabilité de cliquer sur la page. Quelle est l'impact de cette corrélation pour les référenceurs ? Il est simple : si vous voulez augmenter la probabilité de clic sur vos résultats dans les SERP, vous devez travailler les requêtes reformulant votre requête principale (« voiture » -> « voiture familiale », etc.).
- **Caractéristiques situationnelles.** Les critères de situation sont ceux qui sont indépendants des requêtes et des résultats, ils dépendent littéralement de la situation de l'utilisateur. On retrouve ici le nombre de clics effectués depuis le début de la session, le nombre de recherches effectuées, le temps passé entre deux clics, etc. L'idée de ces critères est qu'à un certain moment, l'utilisateur va se fatiguer naturellement de l'action de recherche et va cesser de cliquer car il va cesser de chercher, ou en tout cas cesser de chercher de manière détaillée (et donc

le nombre de clics va baisser drastiquement). Ces caractéristiques ne sont pas corrélées à la probabilité de clic pour les pages pertinentes, mais elles le sont pour les pages non pertinentes. C'est une information intéressante mais naturelle : l'éventuelle lassitude de recherche a un impact sur la probabilité de cliquer sur les pages moins pertinentes, mais pas sur la probabilité de cliquer sur des pages pertinentes, qu'il faudra de toute façon voir pour satisfaire son besoin informationnel.

Toutes ces caractéristiques et corrélations sont agrégées au sein d'une formule qui va donner une estimation de la probabilité de clic. Nous renvoyons le lecteur intéressé par les formules exactes à l'article [1].

La vraie vie : la prédiction ce n'est pas facile

De tout ce que nous avons vu précédemment, nous utilisons les probabilités de clic pour faire le classement, et ce qui apparaît implicitement est que la prédiction de ces probabilités se fait à l'aide des caractéristiques les plus corrélées. Or, réaliser cette prédiction à partir des caractéristiques corrélées est extrêmement difficile. Pour pallier ce problème, Jiang et Allan ont utilisé un algorithme de *machine learning* qui utilise des réseaux de neurones pour classer des éléments de manière optimale (voir l'article [4] par Christopher Burges, chercheur chez Microsoft Research).

A l'aide de cet algorithme, ils ont pu estimer les probabilités de clic pour ensuite modifier la position des pages proposées dans un classement déjà existant. Leur outil a de bonnes performances puisque dans le pire cas, l'erreur qu'il commet est de 13% (en absolu, c'est-à-dire que la probabilité réelle est plus ou moins 13% de la prédiction réalisée). L'erreur dans le meilleur cas est de 8%.

Par ailleurs, un autre problème qu'ils ont du prendre en compte est celui des nouvelles requêtes. En effet, pour prédire les probabilités de clic, il faut posséder un historique concernant les sessions de recherche pour cette requête, ce qui n'est pas toujours le cas. Plusieurs solutions sont évoquées pour ce problème, la plus simple (et moins efficace) étant de regrouper les requêtes par thématique.

Enfin, Jiang et Allan ont testé leur méthodologie en reclassant la première page des SERP pour quelques centaines de requêtes. Les résultats montrent une amélioration des taux de clics sur les pages les plus pertinentes, et donc une diminution des sauts de ces mêmes pages.

Conclusion

Ce travail est extrêmement intéressant, car il formalise pour la première fois l'utilisation d'une notion supplémentaire à la popularité (le PageRank) et la pertinence : l'attractivité des résultats présentés par le moteur.

Jiang et Allan ont même une observation assez amusante dans leurs résultats d'expérience : l'amélioration sur les clics est de 10 à 20%, mais la pertinence des pages proposées décroît de 2,1% environ. En utilisant cette méthode de classement, on augmente la satisfaction des utilisateurs en baissant pourtant la qualité moyenne des pages proposées ! Un comble...

Références

- [1] Jiepu Jiang and James Allan. 2016. Reducing Click and Skip Errors in Search Result Ranking. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM '16). ACM, New York, NY, USA, 183-192.
https://people.cs.umass.edu/~jpjiang/papers/wsdm16_rel_click.pdf
- [2] Emine Yilmaz, Milad Shokouhi, Nick Craswell, and Stephen Robertson. 2010. Expected browsing utility for web search evaluation. In Proceedings of the 19th ACM international conference on Information and knowledge management (CIKM '10). ACM, New York, NY, USA, 1561-1564.
- [3] Ramakrishnan Srikant, Sugato Basu, Ni Wang, and Daryl Pregibon. 2010. User browsing models: relevance versus examination. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '10). ACM, New York, NY, USA, 223-232.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.361.2986&rep=rep1&type=pdf>
- [4] BURGESS, Christopher JC. From ranknet to lambdarank to lambdamart: An overview. Learning, 2010, vol. 11, p. 23-581.
<https://www.microsoft.com/en-us/research/publication/from-ranknet-to-lambdarank-to-lambdamart-an-overview/>



Guillaume Peyronnet est gérant de Nalrem Médias. **Sylvain Peyronnet** est co-fondateur et responsable des ix-labs, un laboratoire de recherche privé. Ensemble, ils font des formations, pour en savoir plus :

<http://www.peyronnet.eu/blog/>

SEO et Prestashop : les fondamentaux



Par Yoann Boidin

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Prestashop est aujourd'hui un des incontournables du monde des CMS e-commerce. Et comme pour toute plateforme de création de site web, il doit être bien configuré afin de bâtir des sites qui soient réactifs par rapport aux critères de pertinence des moteurs de recherche. Voici, dans un premier temps, les fondamentaux proposés par Prestashop en termes de SEO ou comment bien démarrer avec cet outil, avant d'aborder dans un futur article des conseils plus pointus. Mais il est avant tout important de partir sur de bonnes bases...

Contexte marché

Dans la catégorie des CMS e-commerce, nous allons nous intéresser ce mois-ci à Prestashop et plus particulièrement aux fonctionnalités SEO de cet outil.

Solution Open Source créée en 2005, Prestashop représente aujourd'hui 6% de part de marché dans le monde. En France, c'est plus de 40% des sites e-commerce qui seraient sous Prestashop selon les sites Datanize (<https://www.datanyze.com/market-share/e-commerce-platforms/France/prestashop-market-share>) et e-komercio.fr (<http://www.e-komercio.fr/solutions-e-commerce.php>), confirmant ainsi bien la position de leader qu'occupe actuellement cette plateforme dans le paysage du CMS E-Commerce Français. Avec de tels chiffres, force est de constater que la solution doit avoir des avantages certains, notamment en SEO. Mais, ne nous limitons pas à une analyse en surface et rentrons sous le capot de Prestashop pour voir ce qu'il en est vraiment.

A noter : Prestashop va prochainement sortir une nouvelle version, numérotée 1.7 (<https://www.prestashop.com/blog/fr/les-nouveautes-prestashop-1-7/>) et annonce des changements majeurs, notamment un développement sous le framework Symfony (<https://symfony.com/>) (transition Smarty vers Twig). Il faudra donc surveiller de très près cette évolution et son impact en matière de gestion SEO.

Prestashop, sous le capot

Le CMS, dans sa version actuelle 1.6.1.6, fonctionne sur un serveur web sur lequel PHP 5.2 ou plus et MySQL 5.0 ou plus sont installés. Le moteur de template est Smarty. Il est recommandé d'installer Prestashop sur un serveur Unix avec Apache 1.3 ou plus et de disposer au minimum de 64 Mo de RAM. Il est important de ne pas opter pour un hébergement mutualisé trop faible en ressources sous peine de constater certaines

lenteurs qui pourraient impacter négativement l'expérience utilisateur et l'exploration du site par les robots, ce qui en SEO est un facteur important.

L'éditeur propose des options facultatives mais recommandées comme MemCached pour améliorer les performances de Prestashop côté serveur (<https://www.prestashop.com/fr/configuration-requise>). La solution Prestashop est également alimentée en modules complémentaires (gratuits et payant) afin de répondre à des besoins de nouvelles fonctionnalités.

Fonctionnalités SEO dans Prestashop

Prestashop intègre nativement des fonctionnalités SEO. Voici lesquelles...

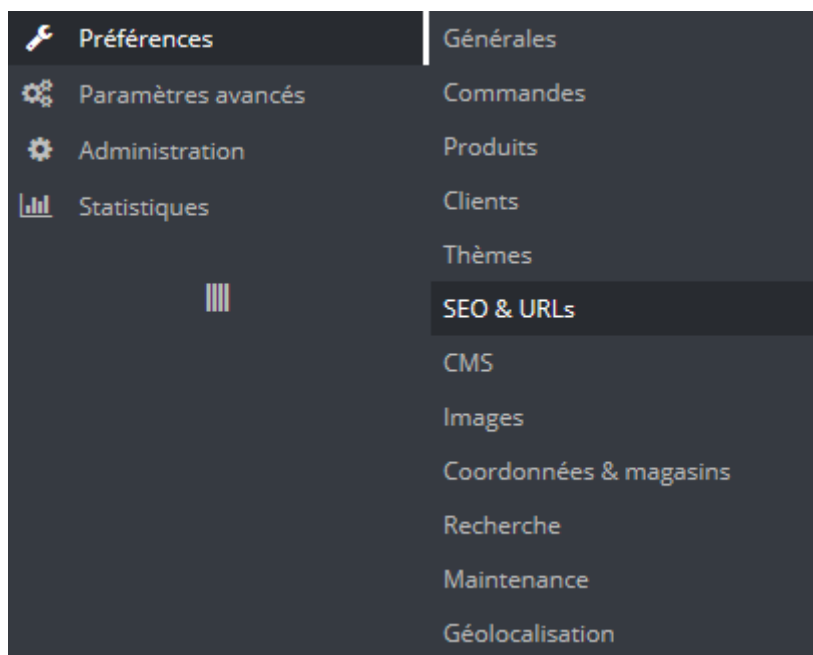


Fig.1. Vue depuis le back-office

La première partie est consacrée à la liste des pages template par défaut : **page d'accueil, page promotions, page marque...** Chaque ligne est administrable pour configurer une URL.

ID ▼▲	Page ▼▲	Titre de la page ▼▲	URL simplifiée ▼▲	
--				Rechercher
	1	pagenotfound	Erreur 404	page-introuvable
	2	best-sales	Meilleures ventes	meilleures-ventes
	3	contact	Nous contacter	nous-contacter
	4	index		
	5	manufacturer	Fabricants	fabricants
	6	new-products	Nouveaux produits	nouveaux-produits
	7	password	Mot de passe oublié	recuperation-mot-de-passe
	8	prices-drop	Promotions	promotions
	9	sitemap	Plan du site	plan-site

Fig.2. Back-office – préférences > SEO & URLs

On peut remarquer les colonnes « Titre » et « URL simplifiée », éléments majeurs pour le SEO, qu'il est possible, voire même recommandé, de modifier afin d'éviter l'indexation d'informations issues de champs renseignés par défaut. Cela évite le genre de résultats dans les SERPs illustré par la figure 3.

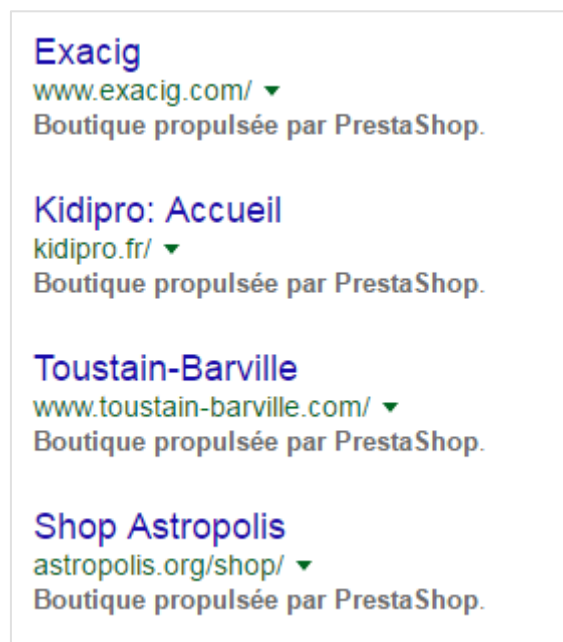


Fig.3. Exemple de balises meta non optimisées et indexées dans Google

Il est important de modifier et optimiser les balises meta et URL de ces pages en raison de leur poids SEO. En effet, ces pages sont souvent liées depuis le footer ou le header et sont donc très explorées par les robots.

A PROPOS	AIDE	SERVICES
Qui sommes nous ?	Livraison	Cartes Cadeaux
Plan du site	Paieement	Magazine
Mentions légales	Echange et retour	
Confidentialité	FAQ	
Boutique Smallable Paris 6ème	CGV	
Contactez-nous	Guide des tailles	
	Guide des pointures	

Fig.4. Exemple d'un footer Prestashop

Sur un site e-commerce, il est bien-sûr difficile de faire l'impasse sur ce type de page. La recommandation SEO est d'optimiser les meta et les URL afin de se différencier des innombrables boutiques en ligne sous Prestashop et qui ont toutes la même structure par défaut.

BALISE MÉTA MOT-CLÉS

* Page index

Titre de la page

Meta description Boutique propulsée par PrestaShop

Meta mots-clés Ajouter un mot-clé

* URL réécrite

Fig.5. Modifier la page d'accueil : champs titre et meta description. La meta mots-clés n'est pas ou plus pris en compte par Google (<https://www.youtube.com/watch?v=jK7IPbnmvVU>), Bing et Yahoo mais peut (encore !) servir le SEO sur Yandex (<https://help.yandex.com/webmaster/controlling-robot/html.xml>) (Russie).

BALISE MÉTA MOT-CLÉS

* Page prices-drop

Titre de la page Promotions

Meta description Nos promotions

Meta mots-clés Ajouter un mot-clé

* URL réécrite promotions

Fig.6. Autre exemple avec la page Promotions : champs titre et meta description + modification de l'URL pour se différencier.

Configuration des URL

Fig.7. Configuration des URL de la boutique

Ici, 2 options sont importantes à configurer. La première concerne l'activation de la réécriture des URL. Par défaut Prestashop génère des URL au format :

www.monsite.com/index.php?id_product=3&controller=product

Ce qui correspond dans cet exemple à une URL produit. Afin de rendre plus « seo friendly » les URL pour les robots d'exploration, il est vivement recommandé d'activer l'option avant la mise en production de la boutique. La fonction *URL simplifiée* fonctionnera uniquement si `rewrite_module` sous Apache (https://httpd.apache.org/docs/current/fr/mod/mod_rewrite.html) est activé.

La deuxième option concerne la redirection vers des URL canoniques. Cette option est incontournable en SEO pour gérer la multi-catégorisation des produits et la navigation à facette qui génère de nombreuses URL en duplicate content.

Format des URL

Chaque champ correspond à un format d'URL de template (catégorie, produit...). Vous pouvez ici modifier les variables. On peut constater que Prestashop utilise une clé unique par URL, représentée par `{id}`. L'ID est obligatoire dans les URL Prestashop donc inutile d'essayer de le supprimer cela va faire planter votre site.

Fig.8. Personnalisation des formats d'URL

La question de l'ID dans les URL a souvent fait débat sur le forum de la communauté Prestashop (<https://www.prestashop.com/fr/communaute>). A défaut de le supprimer

pour obtenir l'URL parfaite, il est tout à fait possible de basculer l'ID à la fin. Cela permet de prioriser à l'exploration un chemin d'URL vers un mot-clé et non un nombre.

Génération du robots.txt

La dernière option disponible est de pouvoir générer le fichier robots.txt.



Fig.9. Bouton pour générer le robots.txt

Le fichier robots.txt est plutôt bien configuré par défaut. En revanche, dès lors qu'une boutique Prestashop est customisée (développement de nouveaux templates de pages par exemple), il est important de surveiller l'exploration de votre fichier via Google Search Console (<https://www.google.com/webmasters/tools/home?hl=fr>) et d'utiliser l'outil de test de robots.txt afin d'y amener des correctifs éventuels. Si vous modifiez manuellement le robots.txt, n'oubliez pas que si vous le re-générez depuis le back-office, il sera écrasé par le nouveau.

Optimisation SEO – catalogue

De la même manière que les pages template de la boutique, Prestashop propose un mode édition pour chaque fiche produit, catégorie... qui permet de remplir des informations SEO.

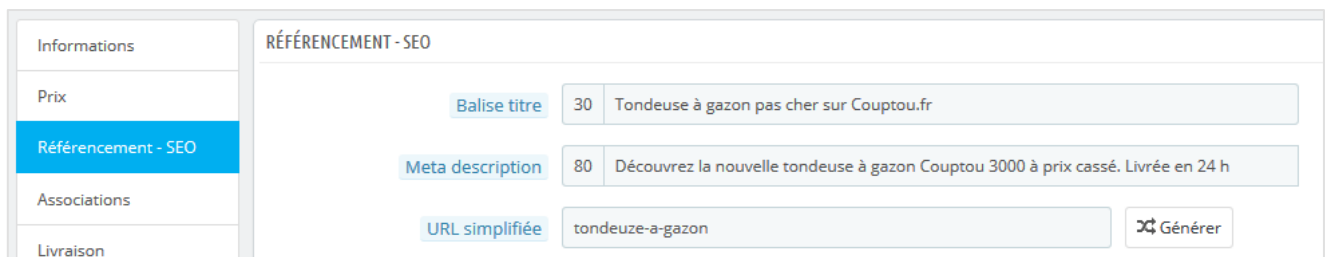
L'image montre l'interface d'administration Prestashop pour l'édition des balises meta d'un produit. À gauche, une barre latérale contient des liens : Informations, Prix, Référencement - SEO (surligné en bleu), Associations, et Livraison. La zone principale est intitulée "RÉFÉRENCIEMENT - SEO". Elle contient trois champs de saisie : "Balise titre" avec la valeur "30" et le texte "Tondeuse à gazon pas cher sur Couptou.fr"; "Meta description" avec la valeur "80" et le texte "Découvrez la nouvelle tondeuse à gazon Couptou 3000 à prix cassé. Livrée en 24 h"; et "URL simplifiée" avec la valeur "tondeuse-a-gazon". À droite de ces champs se trouve un bouton "Générer" avec une icône de document.

Fig.10. Edition des balises meta d'un produit

Prestashop et Performance

Au-delà des fonctionnalités purement SEO, il existe d'autres éléments que vous pouvez piloter à partir du back-office, notamment la performance du site. Pour cela, il suffit d'aller dans « Paramètres avancés > Performance > Smarty » et « Paramètres avancés > Performance > CCC (CONCATÉNATION, COMPRESSION ET MISE EN CACHE) ».

Fig.11. Réglages Smarty

Les réglages par défaut conviennent pour une boutique en production. Si vous effectuez des modifications majeures de template et que vous souhaitez servir aux internautes la dernière version de votre site, il est recommandé de forcer la compilation et de la maintenir quelques heures.

Maintenir également l'option « vider le cache chaque fois qu'il y a une modification », ainsi vos pages seront rafraîchies dès qu'une nouvelle modification est détectée.

Fig.12. Fonctionnalités désactivables

Prestashop propose des options par défaut, lesquelles si elles ne sont pas utilisées pour gérer votre catalogue ou vos clients peuvent être désactivées.

Fig.13. Concaténation, compression et mise en cache

La figure 13 propose des paramétrages qui vous permettent d'optimiser davantage les performances de votre boutique : système de cache pour les CSS, javascripts, compression de scripts... Attention de vérifier que la boutique fonctionne bien et que l'affichage du template est conforme. Il est recommandé de le faire pas à pas.

En bas de la page des réglages, une option « Cache » est désactivée par défaut. Il est recommandé de l'activer.

Prestashop : modules SEO et thèmes

Nativement, et comme vous avez pu le constater ci-dessus, la solution Prestashop intègre des fonctionnalités de base en SEO. Elles restent finalement assez sommaires.

Pour aller plus loin en SEO Prestashop, il est nécessaire de se pencher sur d'autres aspects que les simples fonctionnalités du CMS, notamment les modules SEO et le thème.

Comme de nombreux CMS e-commerce, une place de marché propose une multitude de modules complémentaires dont une section orientée optimisation SEO. Plus de 200 modules dédiés au référencement naturel (<http://addons.prestashop.com/fr/491-seo-referencement-naturel>) sont disponibles et permettent, entre autres, de travailler le maillage interne de sa boutique, de faciliter l'intégration de Rich Snippets, la gestion des contenus, ou bien encore l'édition auto des balises meta...

Les modules sont développés soit par l'éditeur soit par des développeurs indépendants certifiés ou des agences partenaires. Le forum Prestashop est le meilleur endroit pour sourcer de l'information sur un module spécifique. A noter que des modules, gratuits pour certains, sont proposés en dehors de la plate-forme mais sans la garantie de conformité que l'éditeur impose dans sa place de marché.

Avant de vous lancer tête la première dans l'installation d'un de ces modules, il est recommandé de bien comprendre son fonctionnement et de le confronter à vos objectifs SEO. Les modules promettent trop souvent des automatisations de tâches SEO associées à des résultats rapides ! Ne vous faites pas bernier. N'oubliez pas que les modules sont avant tout un business model additionnel pour Prestashop...

Modules SEO populaires :

- Easy SEO (<http://addons.prestashop.com/fr/20640-optimisation-du-referencement-easy-seo.html>), permet l'édition en masse des balises meta et attributs Alt des images.

- SEO internal linking (<http://addons.prestashop.com/fr/4982-seo-internal-linking.html>) qui, comme son nom l'indique, est orienté autour de l'optimisation de votre maillage interne.
- Module SEO 404/301 (<http://addons.prestashop.com/fr/11258-seo-404-301-produits-et-categories-supprime-desactive.html>) pour gérer au mieux l'obsolescence de votre catalogue (produits non disponibles).
- Advanced Search 4 (<http://addons.prestashop.com/fr/2778-advanced-search-4.html>) pour gérer efficacement la navigation à facette.

Cette liste n'est bien sûr qu'un aperçu des innombrables modules disponibles.

Le thème Prestashop que vous allez choisir (un thème prêt à l'emploi ou un thème customisé) ne doit bien sûr pas être négligé. Temps de chargement, caractère Mobile-Friendly, il impactera également de manière directe vos performances SEO.

Conclusion

Dans l'ensemble, Prestashop fournit les fondamentaux du SEO. En respectant rigoureusement la procédure d'installation et de configuration par défaut, une boutique sous Prestashop a toute capacité à bien s'indexer.

En fonction de vos objectifs, de la complexité de votre site, de la richesse et de la profondeur de votre catalogue produits, vous aurez des choix à faire qui impacteront nécessairement le SEO. La mise en place d'une navigation à facettes, l'intégration de schema.org, la gestion des images sont autant de sujets qui soulèveront des questionnements : développement spécifique, achat d'un module prêt à l'emploi... Il n'y a pas de solution miracle.

Sachez au final que la customisation d'un CMS (intégration, développement spécifiques, pages personnalisées non natives...) impacte le fonctionnement de base et peut être à l'origine de dysfonctionnements SEO. Afin de sécuriser au mieux les performances de votre site, il est recommandé de bien évaluer les changements à mener, surtout pour une boutique existante qui dispose déjà d'un positionnement dans les moteurs de recherche.



Yoann Boidin, *Consultant Expert Full Traffic, Synodiance* (<http://www.synodiance.com>)

Cocon.se : voir la structure de son site avec les yeux de Google



Par Sylvain Deauré

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

La représentation graphique de la structure et de l'arborescence d'un site devient une étape indispensable lorsqu'on désire auditer l'optimisation d'une source d'information pour Google. Comprendre comment le contenu est articulé, comment le visite et le voit un moteur de recherche, permet de détecter rapidement et visuellement d'éventuelles erreurs et de les corriger dans la foulée. L'outil Cocon.se permet ainsi d'effectuer ce travail et de voir votre site « avec les yeux du moteur ». Explications par le concepteur de l'outil...

En SEO, on utilise la plupart du temps deux grandes familles d'optimisations : le « off site » et le « on site ». Le « off site » s'oriente principalement autour du linking. Il s'agit sans doute du levier le plus souvent utilisé, mais également le plus facilement pénalisé par Google, notamment avec ses filtres Penguin.

Les leviers « on site » sont moins bien exploités et semblent parfois accessoires, moins importants. Pourtant, il n'en n'est rien. En levier « on site », on peut bien sûr optimiser le contenu, mais il ne faut pas oublier le **maillage interne**, la façon dont les pages de son site sont reliées entre elles. Google, dans ses guidelines (<https://support.google.com/webmasters/answer/35769?hl=fr>), insiste d'ailleurs sur cet aspect. Voici quelques extraits choisis : « Accéder à toutes les pages du site depuis un lien situé sur une autre page accessible » « Limitez le nombre de liens par page » « Empêcher l'exploration d'espaces infinis » « Évitez de faire appel à des ID de session ou à des paramètres d'URL » « Veillez à ce que tous les liens redirigent vers des pages Web en ligne » et surtout, « **Créez une arborescence de pages claire et conceptuelle.** »

Il semble évident que pour que Google puisse correctement parcourir un site, l'indexer, et décider de quoi celui-ci parle, de la thématique à laquelle il appartient, du besoin auquel il répond, etc., la « cartographie » du site, les chemins qu'il propose sont un élément clé.

Dans cet article, nous allons vous présenter le site Cocon.Se (<http://cocon.se/>), un outil qui a pour vocation de vous montrer votre site comme le voit un robot, au travers de son arborescence et de ses liens. L'objectif de cet outil est d'obtenir une représentation graphique du site, sur laquelle se baser pour détecter des erreurs de structure et de conception.

La genèse de l'outil

Tout est parti comme souvent d'un besoin personnel, et d'une intuition. Il y a deux ans de cela, on pouvait voir quelques visualisations de sites, mais cela ressemblait plus à un gadget à l'époque. On pouvait considérer ces visualisations comme expérimentales. Ou on les utilisait pour mettre en évidence un souci qu'on avait déjà identifié, ce qui limitait leur intérêt.

Nous avons l'intuition qu'on pouvait faire bien plus que cela, et qu'avec une bonne visualisation, une bonne spatialisation, on pouvait tirer des analyses concrètes de la représentation graphique d'un site.

L'idée a fait son chemin depuis, et d'autres personnes (en France comme à l'international) commencent à utiliser des visualisations graphiques pour montrer au client ses problèmes de structure, voir l'arborescence comme un cocon, pour diagnostiquer des soucis techniques ou conceptuels, plus facilement qu'avec des indicateurs purement numériques.

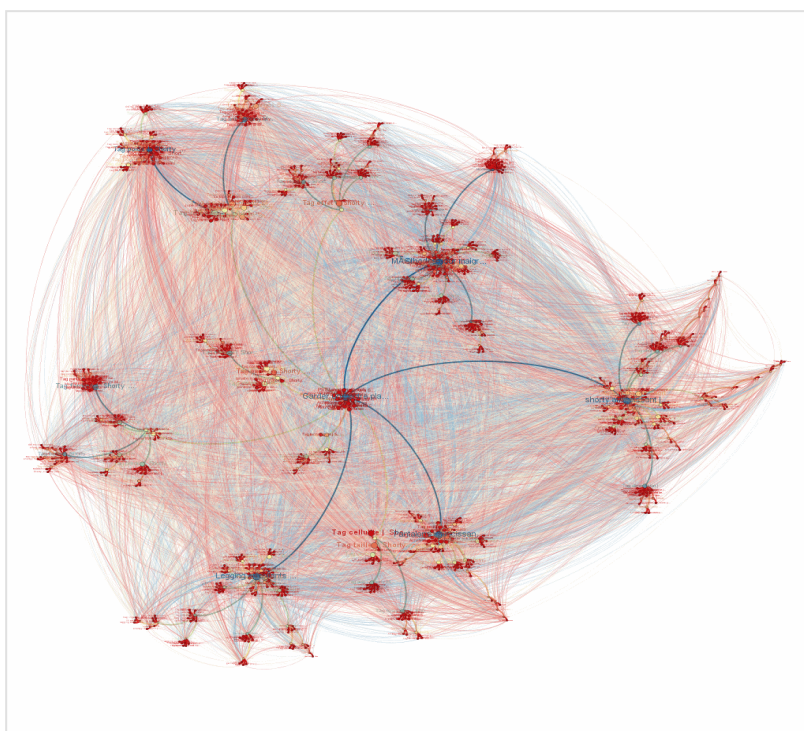


Fig. 1 : Une des premières vues expérimentales, avec Gephi (voir article à son sujet par ailleurs dans la lettre).

Au début, pour prototyper, expérimenter, nous utilisions Gephi. Il permet assez facilement d'obtenir des vues globales, macroscopiques. Mais on est vite confronté à des problèmes de lenteur, de ressources, de manipulations longues si on veut aller plus loin avec cet outil. C'était trop laborieux et une fois le principe validé, nous avons basculé sur du développement personnalisé.

Comment ça marche, en quelques mots...

Le fonctionnement de l'outil Cocon.se est le suivant :

- On crawle le site à la façon de Google, en suivant les liens depuis une page de départ (le plus souvent la page d'accueil) ;
- Au passage, on détecte d'éventuelles erreurs de base ;
- On calcule des indicateurs de structure locaux et globaux, dont le PageRank interne ;
- On calcule des organisations spatiales des pages du site, qui rendent compte de sa hiérarchie telle que perçue par un moteur ;
- On affiche pages et liens sous une forme structurée qui met en évidence l'arborescence interne du site.

Le crawl

Le Crawl est le premier contact d'un moteur avec votre site. C'est lors du crawl que le moteur se fait une idée de votre site, des pages importantes. Avant même d'avoir tout crawlé, Google a déjà décidé de classer en premier les pages « tag » ou « catégorie » de votre site, par exemple, selon le maillage interne partiel qu'il a rencontré. C'est un élément vraiment critique, essentiel, et qu'il est pourtant facile (et courant) de faire de travers.

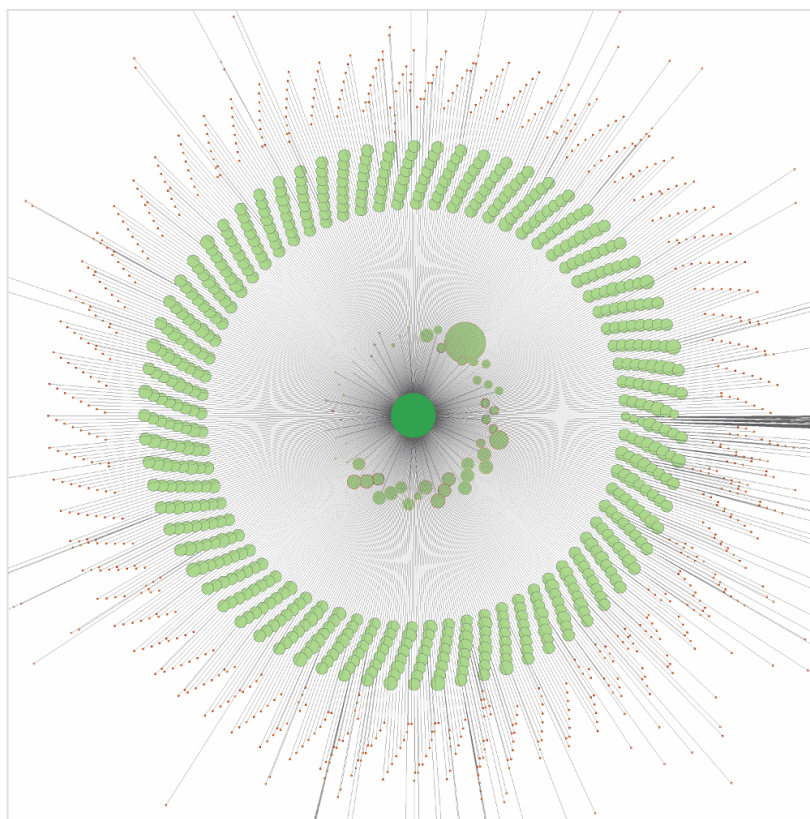


Fig. 1 : Un exemple de site surmaillé, avec un énorme mega-menu (ronds verts dans le cercle autour de la page d'accueil)

Le fait d'avoir développé notre propre crawler est un atout important : cela nous a permis de nous rendre compte du nombre de points essentiels qui peuvent être mal codés, ou qui posent des problèmes à un robot. Corriger les problèmes de crawl permet de se recentrer sur l'essentiel, de donner tout de suite toutes les billes à Google, toutes les infos au bon endroit, dans le bon ordre, et sans ambiguïté.

Même si Google utilise des tonnes de rustines pour corriger des problèmes de crawl et de structure (*canonical*, *redirects*, *no follow*, *robots.txt*, *rel prev/next...*) cela ne veut pas dire qu'il faut les utiliser à tout bout de champ.

Un site bien conçu dès le départ, qui n'aura pas besoin de toutes ces rustines, aura toujours une longueur d'avance sur un site rafistolé à coup de « verrues ».

In fine, notre crawler a pour objectif d'explorer votre site de la même manière que Googlebot, afin de voir les mêmes informations locales que Google, et fournir donc une représentation similaire à ce qu'en voit le moteur de recherche.

Les algorithmes et calculs

Nous n'allons pas nous étendre ici sur les aspects purement techniques de notre infrastructure. Disons juste qu'elle a été pensée dès le début pour être performante et « scalable ».

Nous nous basons sur des algorithmes connus (parcours de graphe, arbre recouvrant, surfeur aléatoire / PageRank...) et avons également toutes une série d'algorithmes *maison* pour extraire les données importantes lors du crawl, optimiser les calculs, les traiter de manière répartie, et surtout calculer une organisation efficace des pages du site pour la visualisation.

Coté langages, nous utilisons, selon les composants, les technologies PHP, Python, Javascript, C#, Node.js, etc.

La visualisation du site

Une visualisation du maillage interne permet de lever le doute, d'avoir une vraie réponse à la question « mon site est-il vu par Google comme je l'ai pensé/conçu ? ». On y trouve souvent une bonne marge de manœuvre, qui donne un levier concurrentiel énorme, sans le risque associé à des actions de linking plus ou moins factice.

On pourrait penser que la restitution des datas est une brique « simple », et que plus on a de données, plus on a d'information. En fait, c'est le contraire. Plus on a de datas, plus on est perdu ! La difficulté n'est pas d'avoir beaucoup de données, mais bien de trouver celles qui sont significatives, qui vont structurer le reste et le rendre intelligible et actionnable.

Avec notre crawler, on a l'avantage de pouvoir collecter, filtrer, la data comme on veut, pour la présenter, on l'espère, sous une forme utilisable.

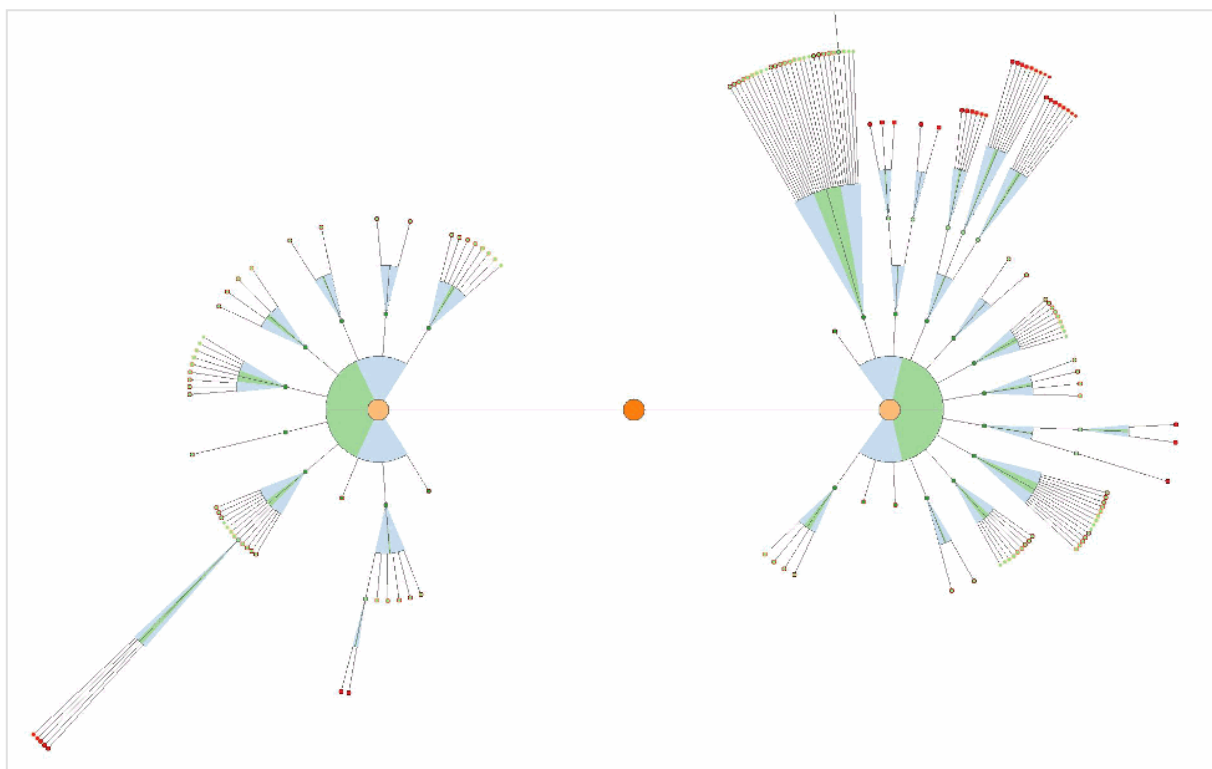


Fig. 3 : Exemple de visualisation de 2 catégories (silos) d'un site.

Lorsqu'on gère un site web, on pense parfois que le CMS que l'on utilise est « optimisé seo » ou que le plugin « super opti seo » que l'on a installé est fait pour, que si il nous affiche une coche verte dans son administration, tout se passera parfaitement bien. Ce n'est malheureusement jamais aussi simple. Un blog, avec ses catégories, ses articles chronologiques, son nuage de tags en vrac, n'est pas optimisé pour le Seo. Un CMS E-Commerce, de type Prestashop, par défaut, n'est pas optimisé pour le SEO.

Quand on fait une visualisation de ce type de site, on se rend tout de suite compte de la différence énorme entre l'organisation « pensée » du site, et celle que voit le moteur, qui est bien bien plus brouillonne et complexe !

En pratique, comment voir mon site avec Cocon.Se ?

Voici la procédure à suivre pour visionner votre site avec notre outil :

- Inscrivez-vous sur <https://self.cocon.se> ;
- Vous aurez accès à quelques sites déjà cartographiés si vous voulez prendre l'interface en main ;
- Renseignez votre profil (important pour la gestion de la TVA et des factures) et choisissez un type de compte. Le compte « pay as you go » est gratuit et permet d'acheter des crawls à la demande. Les abonnements sont sans engagement de durée et plus économiques ;
- Cliquez sur « Nouveau crawl », copiez/collez l'adresse de votre site depuis la barre d'adresse de votre navigateur vers la zone « start url » ;
- Vérifiez et lancez le crawl ;

- Quelques minutes plus tard (selon la taille et vitesse de votre site), vous avez une première photo de votre site disponible dans votre espace d'administration ;
- Vous pouvez ensuite créer d'autres vues, selon d'autres facettes ou depuis d'autres pages de départ.

Comment interpréter ce que je vois ?

Les codes suivants sont utilisés par l'outil pour visualiser graphiquement votre arborescence :

- Les pages du site sont représentées par des disques colorés ;
- La taille du disque rend compte de l'importance relative de la page (Pagerank interne) ;
- La couleur du disque représente le nombre de « clic » depuis la page d'accueil ;
- Les liens entre les pages ne sont, volontairement, pas tous représentés. Vous voyez les liens qui constituent le plus court chemin vers chaque page, depuis l'accueil (la vue interactive, elle, permet d'afficher tous les liens). Ceci permet de mieux mettre en évidence la hiérarchie de vos pages ;
- Les pages sont numérotées, vous avez la correspondance avec l'URL via le menu 'top pages', ou directement au survol de la souris, en passant en vue interactive.

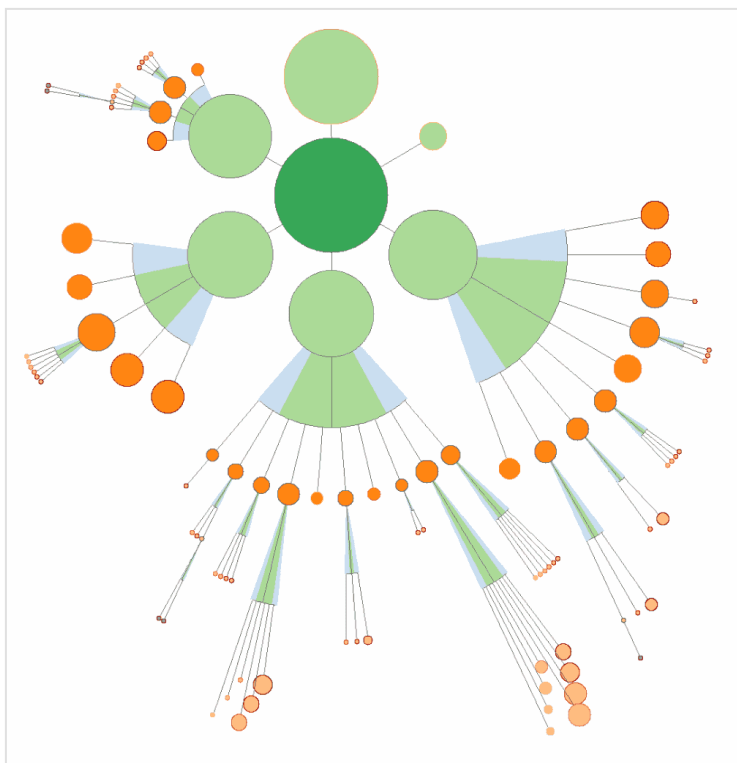


Fig. 4 : Un exemple typique de visualisation d'une arborescence en silos.

Si votre site est de taille conséquente, vous ne verrez clairement que les premiers niveaux. Vous pouvez alors créer d'autres vues centrées sur des catégories de votre site, en indiquant le N° de la page de départ, pour entrer dans le détail, section par section.

Quelques exemples de vues et d'analyses simples

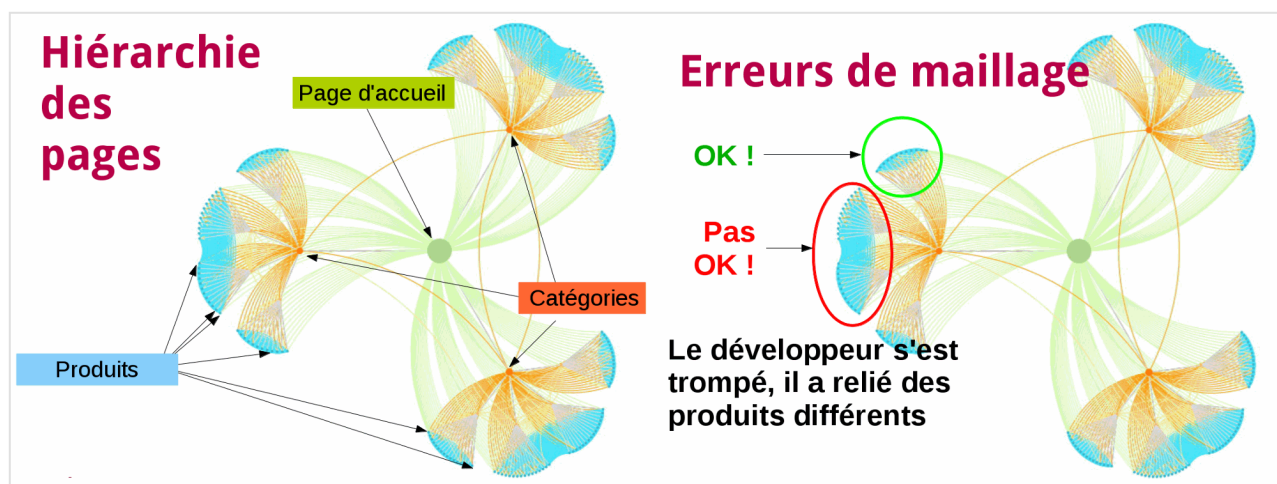


Fig. 5 : Exemples d'analyse et détection d'erreur de maillage

Sur la cartographie du site Fig. 5, on voit plusieurs choses : d'une part, l'arborescence des pages chère à Google est bien respectée. On distingue une hiérarchie claire avec une page centrale, en vert, suivie de pages « catégories » en orange (plus petites, donc secondaires en terme de Pagerank interne) et enfin des pages « produits » en bleu, encore plus petites, reliées en grappes, chacune dans son silo (sa catégorie).

Chacune dans son silo ? Pas tout à fait ! On voit des liens bleus qui passent d'une branche (catégorie) à une autre : ce sont des produits d'un silo qui font des liens vers le silo voisin. Ce n'était pas voulu, et cela aurait été fort difficile à détecter autrement que sous cette forme graphique.

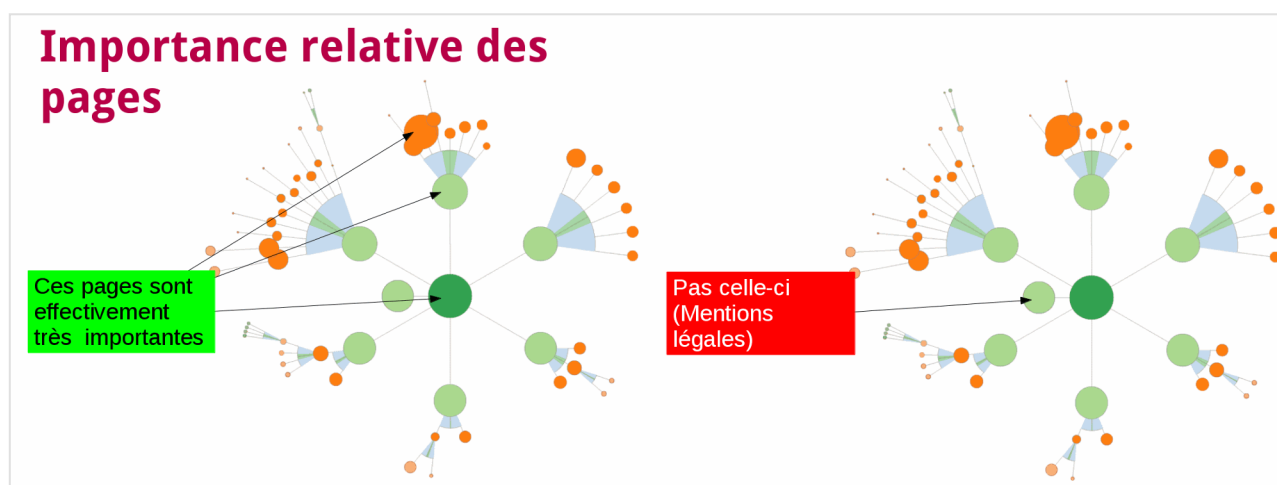


Fig. 6 : Exemples d'analyse : Importance relative des pages

Pour le site de la Fig. 6, on s'est concentré sur les pages importantes en termes de Pagerank interne, donc celles qui le mieux placées pour se positionner.

La page d'accueil est effectivement importante, ainsi que les grosses catégories du site (disques vert clair autour). Une page orange ressort également : C'est une page qui a un objectif et de positionnement et de conversion, c'est donc normal.

La page « Mentions légales » est elle aussi indiquée comme importante. Pourtant on ne cherche pas à me positionner sur ce mot clé, et il est dommage de « gaspiller » du Pagerank sur cette page.

En ajustant le maillage du site, il sera possible de redistribuer le PR plus efficacement.

Comme la visualisation se fait dans les mêmes conditions, on peut facilement faire un comparatif avant/après : que ce soit lors d'une refonte complète ou des mises à jour incrémentales, on voit nettement la différence.

Dans la rubrique « Historique » de l'outil, on voit les effets d'actions de maillage, d'ajout de lien sitewide, d'ajout de contenu, de suppression de pages parasites. On peut ainsi voir ce qui a changé, ce qu'il est difficile de faire de manière fiable avec Gephi.

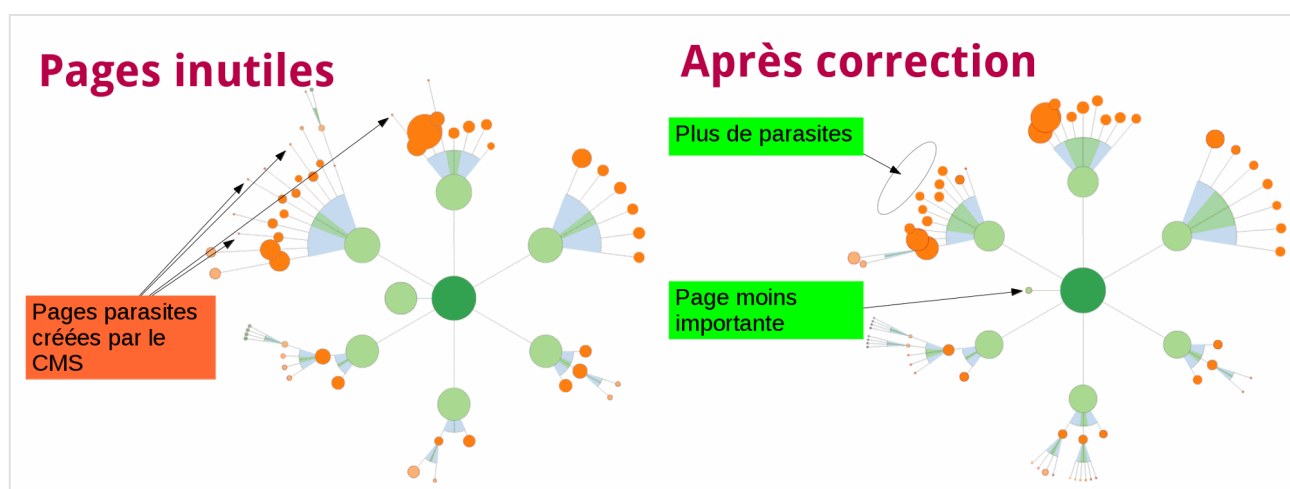


Fig. 7 : Avant – Après : Correction du maillage et des pages inutiles

Sur la Fig. 7, on voit le même site (il s'agit de cocon.se) avant et après correction. On voit que la page « mentions légales » a une importance moindre après correction. De même, les pages parasites créées à tort par le CMS ont pu être supprimées. On remarque également de nouvelles pages, en orange clair, qui sont apparues en bas de la cartographie, sans pour autant chambouler toute l'organisation. On peut comparer avant/après.

C'est cette stabilité et cette facilité dans les visualisations qui permet d'agir sur la structure, d'observer les changements et d'améliorer la situation par touches successives pour aller vers un site « parfait ».

Typologie des utilisateurs de Cocon.se

Au départ, nous pensions que les clients allaient surtout être des agences web. En effet, ils connaissent les sites de leur clients, leurs objectifs, et ont les pré-requis techniques pour analyser et corriger les sites. En fait, à l'usage nous avons également comme utilisateurs de l'outil beaucoup de webmasters et de clients finaux qui travaillent eux-même et en direct sur la structure de leurs sites.

Avantages pour les agences :

- Une photo du site, construite toujours de la même manière ;
- Pas de paramétrage compliqué, pas besoin de grosse machine ni de temps de calcul, pas de tâtonnements pour obtenir une vue ;
- Les erreurs graves sont visibles immédiatement ;
- Identification de gros leviers d'amélioration ;
- Une base concrète pour expliquer au client ce qui ne fonctionne pas. Le client est convaincu beaucoup plus facilement de la nécessité de mise à jour, d'action, quand il a vu, quand on lui a expliqué image à l'appui ce qui ne va pas ;
- Le travail effectué lors d'une refonte est visible. Sur le site lui-même, ce n'est pas forcément évident. Sur la vue, c'est flagrant : il est plus facile de valoriser l'impact de la refonte... ;)
- Plus on a l'habitude des vues, et plus vite on détecte les problèmes, plus facilement on les lit et on en extrait des infos actionnables et donc de plus en plus rentables au fil du temps ;
- Effet « whaouh » pour le client qui peut enfin « voir » son site, qui a une représentation concrète de quelque chose qui sinon, reste assez abstrait pour lui.

Avantages pour les webmasters :

Nous avons également de très bons retours d'expérience à ce niveau. L'outil permet de ne plus travailler en aveugle. Il permet des tests beaucoup plus rapides et efficaces. Au lieu d'avancer à tâtons, on peut en *pre-prod* simuler un changement de menu, de structure, ajouter un silo, et voir, réellement, l'impact que ce changement a sur le site, sur les pages qui sont boostées (ou pas), sur la topologie du site. On sait quelle image on présente à Google, et quelle est la différence avec la version précédente.

Pour analyser ensuite les changements de positionnement, cela n'a pas de prix. On en déduit plus facilement des corrélations. Bien sur, cela suppose qu'on ait réfléchi en amont à la structure de son site, c'est là qu'on voit de suite si « ça colle » ou si on découvre une grosse divergence entre ce qu'on imagine et ce que voit un robot.

Il est conseillé de commencer à utiliser l'outil sur petits sites, pour bien appréhender la signification des vues, le parallèle entre la structure du site et comment elle transparaît dans la vue.

Donnez du sens à votre structure de site

On nous demande souvent : « Voilà la vue de mon site, est ce que c'est bien, comment je peux améliorer ? » C'est hélas impossible à dire comme ça... Il n'existe pas une

structure magique qui fonctionne pour tout ! Ce qui fonctionne, c'est une structure, une hiérarchie logique, qui a un sens pour l'utilisateur, qui a un sens d'un point de vue conversion, et qui se retrouve dans le site tel qu'exploré par le robot. Là on a tout bon.

Ce qu'on peut voir de immédiatement sur une visualisation Cocon.Se, ce sont les gros problèmes structurels, et il certains sont très fréquents :

- Site sur maillé ;
- Mega menus ;
- Liens site-wide vers des pages inutiles (sans conversion) ;
- Tunnels infinis ;
- Erreurs de canonical, de http/https, voire de redirections ;
- Etc.

Pour le reste, il n'y a pas de réponse à l'emporte pièce : il faut avoir eu une réflexion préalable sur l'objectif du site, son organisation, et on cherchera ensuite à voir si cette organisation se retrouve effectivement dans la vue.

- J'ai voulu isoler des sections, faire des silos cohérents, ces silos sont-ils détectés ? Sont-ils bien isolés ou fuient-ils ?

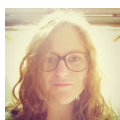
- Je veux « pousser » cette page sur telle thématique, le PageRank interne circule-t-il bien comme je le pense, cette page est-elle effectivement bien mise en valeur ?

Bref, cocon.se est un outil qui met en évidence bon nombre de points, mais qui ne se substitue pas à une réflexion de fond sur la finalité et l'organisation de son site. C'est un outil qui a besoin d'un humain et d'un cerveau aux commandes !



Sylvain Deauré est ingénieur de formation, et développeur à son compte depuis 1997. Il édite plusieurs services grand public et s'intéresse principalement aux défis techniques ainsi qu'aux applications pratiques de la technologie. Pour en savoir plus : <https://twitter.com/SylvainDeaure>.

Curation : comment réussir sa stratégie ? (2ème partie)



Par Anne-Victoire de Boisgelin

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

La curation de contenu est une part importante d'une stratégie éditoriale web. Elle ne doit cependant pas être LA stratégie de votre site/blog à elle seule. Elle doit être une brique à l'édifice. Qu'est-ce que la curation de contenu ? Quelle stratégie adopter ? Quels outils et quels KPI suivre ? Autant d'éléments que nous étudions dans cet article en deux parties pour vous aider à mettre en place une stratégie de curation efficace !

Dans la première partie de cet article, le mois dernier, était défini ce qu'était la curation et les avantages à mettre en place une telle stratégie : gain de temps sur la production de contenu, une diversité de type d'articles pour gagner en visibilité et en notoriété tout en améliorant le positionnement de son site. Nous y abordions également les toutes premières questions à se poser avant sa mise en place : pour quelle audience et avec quel(s) support(s) à disposition.

Une fois ces éléments clairement identifiés, il ne suffit pas de publier pour publier. Cinq autres étapes viennent s'ajouter à l'élaboration d'une stratégie de contenu : de la recherche des sources et des sujets à l'analyse de l'impact de votre stratégie sur votre marque. Chacune d'entre elles a son importance et sera détaillée ci-dessous. Elles pourront également être facilitées par l'utilisation d'outils et de sites web listés en fin d'article.

Comment mettre en place une stratégie de curation ?

Etape 1 : Identifiez et recherchez des sujets

o Utilisez le « mind mapping »

Pour identifier quelles thématiques vous pourriez aborder via la curation, l'utilisation d'un outil de « mind-mapping » peut s'avérer très utile. Le « mind mapping » permet d'identifier des sujets ou thématiques qui gravitent autour d'un concept ou d'une idée.

Vous verrez que le travail de recherche de sujets en sera facilité. Vous irez également plus loin dans la recherche et la découverte de nouveaux sujets.

Si on prend l'exemple d'un acteur dans la santé et le bien-être, la figure 1 montre à quoi pourrait ressembler un mind mapping dans ce domaine.



Figure 1 – Exemple de mind mapping sur le secteur de la santé

Ici, sont identifiées de nombreuses pistes de thématiques à explorer dans une curation de contenu.

Pour chercher de nouvelles pistes, il est recommandé d'utiliser des outils sémantiques pour savoir ce qui est le plus recherché par les internautes.

Les réseaux sociaux peuvent également venir compléter ce travail. La figure 2 montre un exemple de suggestions de sujets par Pinterest qu'il pourrait être intéressant d'aborder dans cette curation.

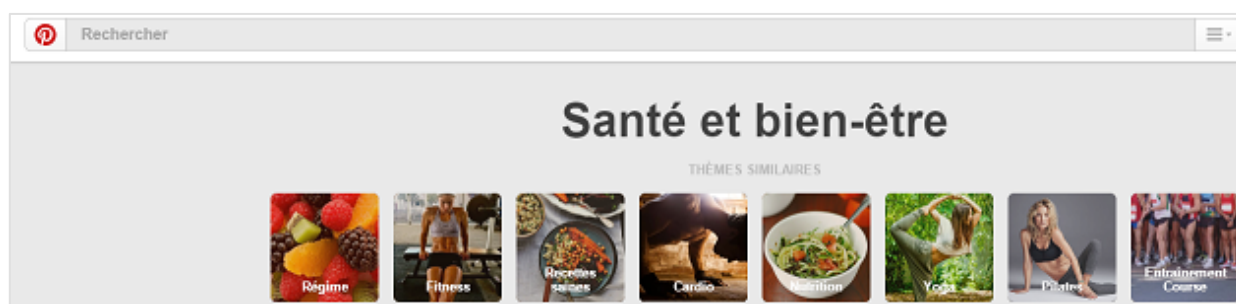


Figure 2 – Exemple de suggestions de thématiques de recherches liées à la santé – Pinterest

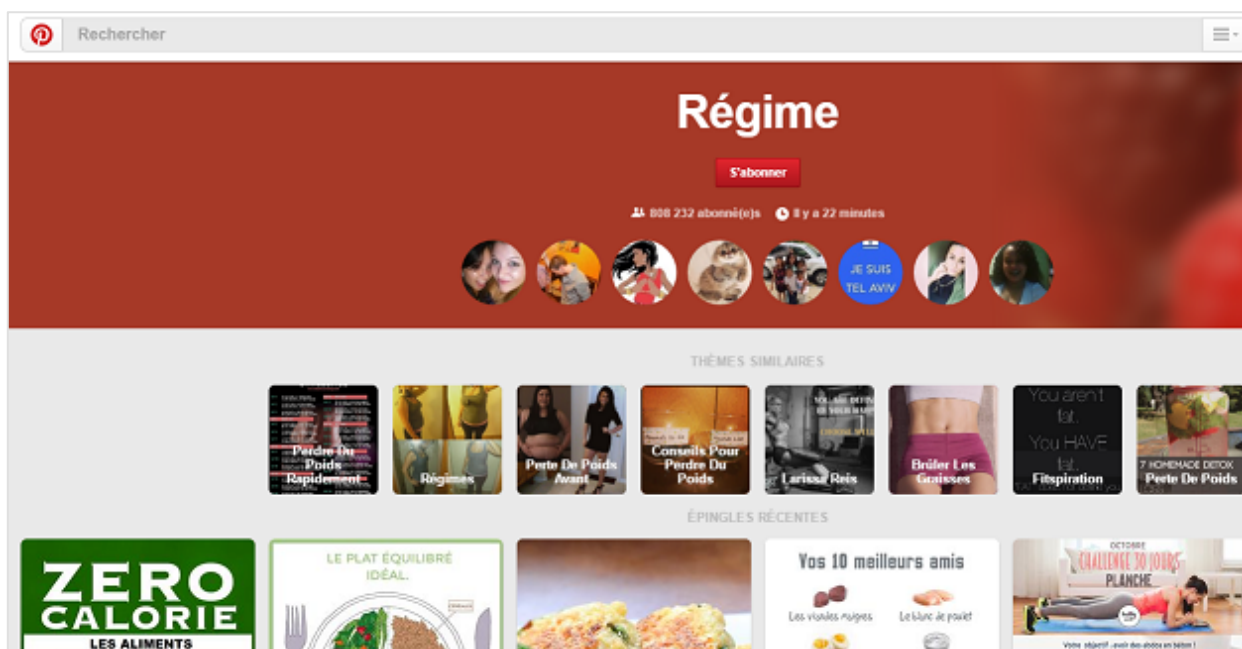


Figure 3 – Focus sur une sous-thématique – Pinterest

La mise en place de ce process peut prendre un peu de temps, ce qui est normal. Il est important de se rappeler que le fait de **se créer une base solide dès le départ**, représente autant de temps gagné par la suite.

Etapes 2 : Identifiez les influenceurs de votre secteur et vos sources d'information

Que ce soit directement sur des plateformes sociales ou via des outils, **identifiez une première base de personnes à suivre** sur les réseaux sociaux. Regardez quels influenceurs ou quels sites ressortent dans Google Actualité. N'oubliez pas non plus les nouveaux arrivants qui commencent à se démarquer. Cette base, au fil de votre curation et de l'actualité, pourra venir s'enrichir.

Identifiez sur quelles plateformes ces influenceurs communiquent et ce qu'ils partagent. Abonnez-vous à leur newsletter et nouez des relations avec eux.

Multipliez et diversifiez les sources d'information. Cela vous permettra d'élargir autant que possible le type de sujets relayés ainsi que les sites web mentionnés dans vos articles.

Intégrez les flux RSS, des alertes Google dans un outil de gestion de contenus type Feedly, Scoop.it ou autre. Bien souvent, sur ce type de plateforme, vous pourrez intégrer différents flux d'informations. Vous pourrez également les catégoriser en différentes thématiques, ajouter des tags, etc. pour faciliter votre travail. Enfin, ces outils pourront vous suggérer de nouvelles sources susceptibles de vous intéresser.

N'hésitez pas à **intégrer vos concurrents à vos outils**. C'est une bonne façon d'être rapidement au courant de ce qu'ils font de leur côté.

Ne négligez pas non plus **ce qui se fait à l'international**. Ce peut être une source d'information précieuse sur de nouvelles tendances pour être précurseur du relais de cette information. Identifiez quelques acteurs de référence sur le marché pour avoir un premier aperçu de ce qui se dit et se fait.

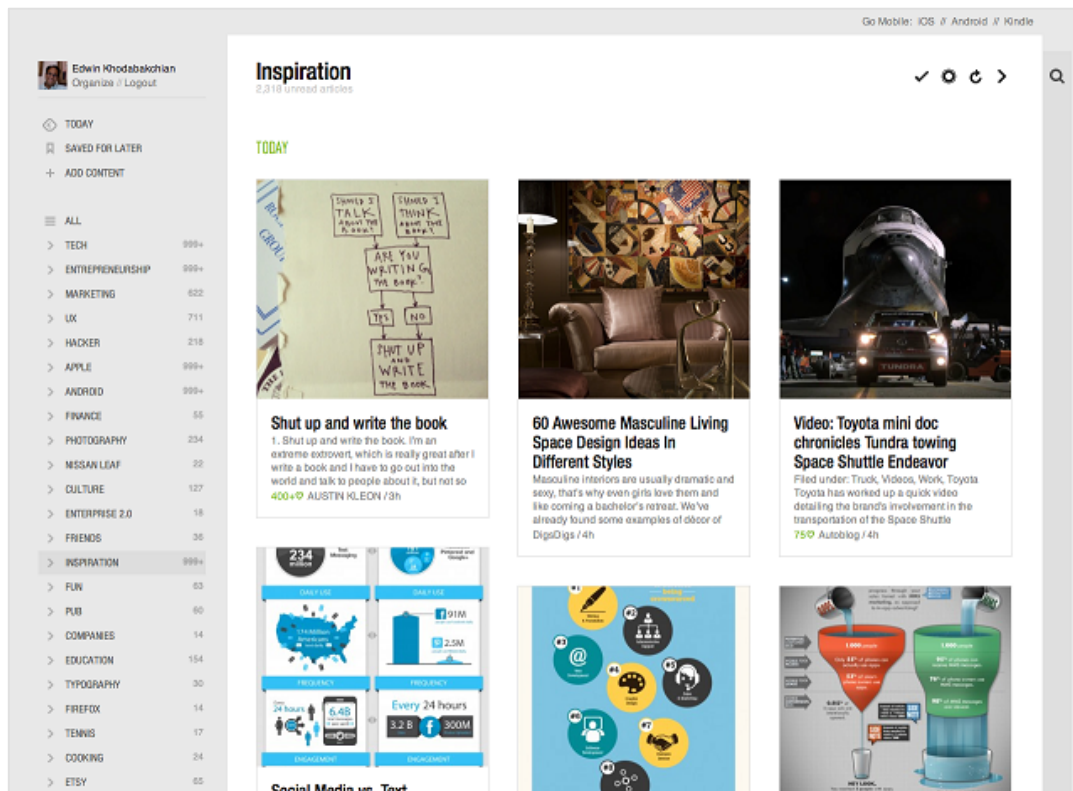


Figure 4 – Exemple de Dashboard Feedly

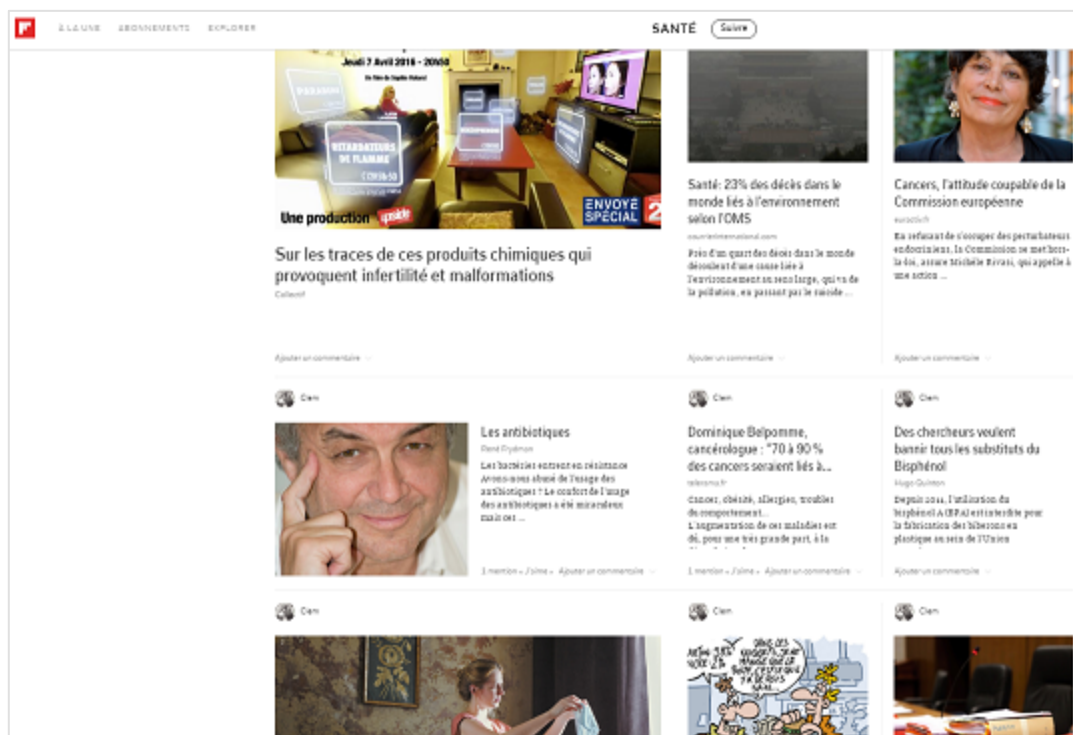


Figure 5 – Exemple de flux de contenus - Flipboard

Définissez un volume d'articles à publier par mois. **Trouvez un ratio raisonnable entre le volume d'articles créés et le volume de contenus curés.** Dans certains secteurs, cela peut-être du 60/40. Il faut trouver un juste milieu et, pourquoi pas, l'ajuster dans le temps.

Créez-vous un calendrier éditorial pour suivre les événements importants de l'année pour ne pas passer à côté d'un sujet chaud qui pourrait faire l'objet d'un article ou de ce qu'on appelle une « *newsjacking* » (stratégie marketing ciblée sur l'actualité).

o *Identifiez les contenus tendances et populaires sur un sujet précis*

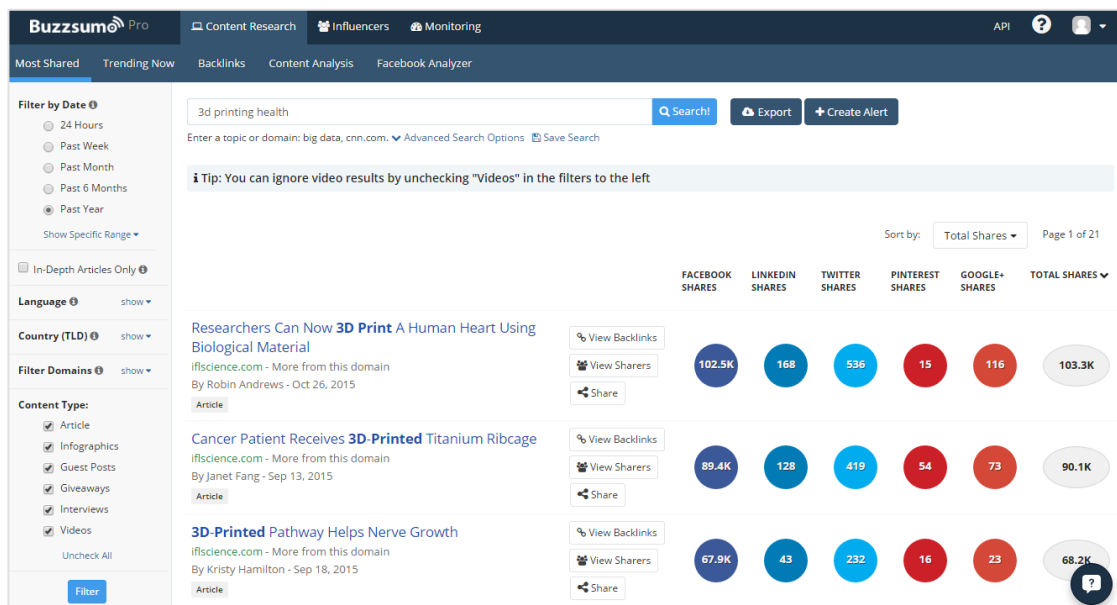


Figure 6 – Exemple de sujets les plus partagés sur l'impression 3D dans le secteur de la santé – Buzzsumo

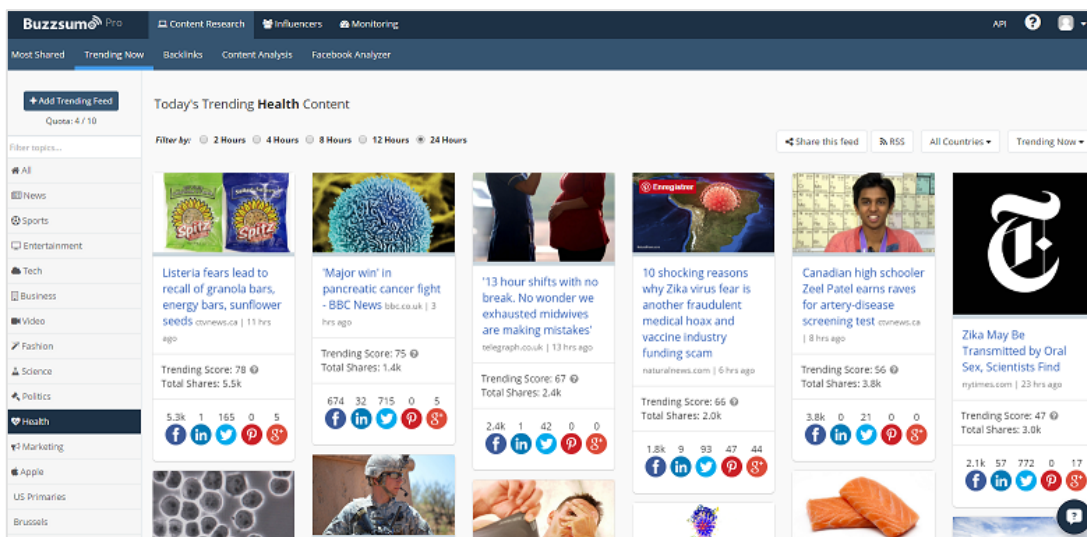


Figure 7 – Exemple de sujets tendances dans la Santé, ici sur les dernières 24 heures – Buzzsumo

Identifiez les sujets les plus pertinents, **soyez réactifs pour être les premiers à relayer l'information.**

Parcourez les réseaux sociaux pour voir les tendances du moment, ce qui se partage et affiner votre recherche.

Etape 3 : Réécrivez les contenus & démarquez-vous des autres

Ecrivez un contenu unique. La curation, même si vous relayer une information issue d'un site tiers, doit faire **l'objet d'une réécriture**. Pour chaque sujet, il faut un contenu unique et ce, bien entendu, pour éviter le duplicate content mais aussi se démarquer des autres articles traitant du même sujet.

Pour chaque contenu créé, ajoutez-y votre expertise : introduisez avec une explication du contexte, insérez par la suite vos commentaires et/ou une brève analyse. D'un point de vue SEO, comme pour tout contenu, optimisez et structurez votre page avec des mots clés recherchés.

Travaillez votre titre. N'hésitez pas à y consacrer un peu plus de temps. Choisissez un titre qui puisse vous démarquer des autres articles déjà publiés sur le même sujet.

Ne soyez pas radins, citez la source d'origine ! Vous appréciez sûrement qu'on vous cite lorsque vous êtes la source d'une information, alors pourquoi feriez-vous autrement ? Identifiez ainsi clairement vos sources. D'un point de vue netlinking, il est toujours bien vu d'avoir des liens sortants de son site vers des sites de confiance qui font autorité dans leur domaine. Une raison supplémentaire de ne pas hésiter à faire de liens. Privilégiez des liens en dofollow qui attestent de la qualité et de la confiance que vous portez au site vers qui vous faite un lien.

Maillez vos contenus que vous auriez pu avoir déjà rédigé et qui pourraient être un complément de lecture pour vos internautes. Exemple si vous avez un site dans le secteur auto-moto, que vous relayer les photos de la dernière moto imprimée en 3D et que vous avez déjà écrit sur ce sujet, invitez les internautes à poursuivre la lecture sur les articles en question.

Ajoutez des « call to action » pour générer des leads. Assurez-vous qu'une fois qu'un internaute est sur votre article, il puisse par la suite prolonger son parcours sur votre site. Cela peut prendre plusieurs formes : des liens vers des landing pages stratégiques, une inscription à une newsletter, un lien vers un formulaire de contact, un bouton pour une demande de démo, un livre blanc à télécharger, un formulaire de souscription à un outil ou des liens vers des fiches produits etc.



Figure 8 – 83% des marketers listent la génération de leads comme étant un des objectifs clés de leur stratégie de contenus en 2015 - Content Marketing Institute & Marketing Profs (1)

Soyez engagé avec votre audience. Assurez-vous que tous vos contenus puissent être facilement partageables via des boutons ou des *embed codes* lorsqu'il s'agit de partage de données/infographies. Des éléments simples et rapides à mettre en place pour optimiser sa diffusion.

Prenez le temps de répondre aux questions, échangez avec ceux qui prennent le temps de commenter et réagir avec votre contenu. Votre participation sera d'autant plus appréciée qu'elle permettra d'aller plus loin et de créer et fidéliser votre communauté. Ces échanges participent à animer une expertise.

Enfin, soyez régulier dans la production de contenu.

Etape 4 : Diffusez vos contenus de curation, partagez !

- o *Publiez vos contenus et faites la promotion de vos articles sur différents supports*

Certains outils vous permettront d'automatiser une partie de la publication. Pensez notamment à l'utilisation de plateforme comme Buffer ou encore aux recettes IFTTT pour planifier vos relais et gagner en temps.

Multipliez les relais sociaux et choisissez votre plateforme de diffusion selon votre contenu. Pensez à soigner également votre message selon l'audience et

le réseau social utilisé. Utilisez les bons hashtags sur Twitter. Pour cela, aidez-vous d'un outil comme hashtagify.me.

On ciblera davantage une audience BtoB en communiquant sur LinkedIn. Le type de contenu à relayer peut varier également selon les plateformes : Pinterest ou Instagram seront parfaits pour partager des visuels selon qu'ils s'adressent à un public large ou à des professionnels.

Augmentez le reach de votre audience *via* des campagnes publicitaires sur les réseaux sociaux.

o Vous avez la chance d'avoir une audience déjà existante ?

Tant mieux, ils seront les premiers auprès desquels il faut communiquer. Ils sont déjà sensibilisés à votre univers, votre marque et seront sans doute plus enclins à réagir à votre contenu et à le relayer via leurs réseaux.

Appuyez-vous sur votre communauté, mais ne vous limitez pas pour autant à celle-ci. Il faut en effet savoir qu'en janvier 2016, Locowise a mesuré qu'une publication Facebook était vue par seulement 10,3% des fans d'une page Facebook (2).

Développez donc votre audience en communiquant auprès de nouvelles cibles ayant des intérêts similaires à votre secteur.

o Vous n'avez pas encore de communauté ? Pas de panique, l'exercice prendra certainement plus de temps mais la tâche n'est pas impossible. A partir de vos profils sociaux (Facebook, Twitter, Instagram ou LinkedIn), créez quelques campagnes publicitaires en ciblant une audience ayant des centres d'intérêts proche du type de contenu que vous relayez et des services/produits que vous proposez à vos clients.

Avec un budget qui peut commencer à quelques dizaines d'euros par contenu, vous pourrez commencer à lancer des campagnes d'acquisition de fans ou amplifier la visibilité de vos contenus.

Sur Facebook, vous pourrez par exemple vous aider d'outils comme les statistiques de votre Page ou de Facebook Audience Insight pour définir une audience ciblée.

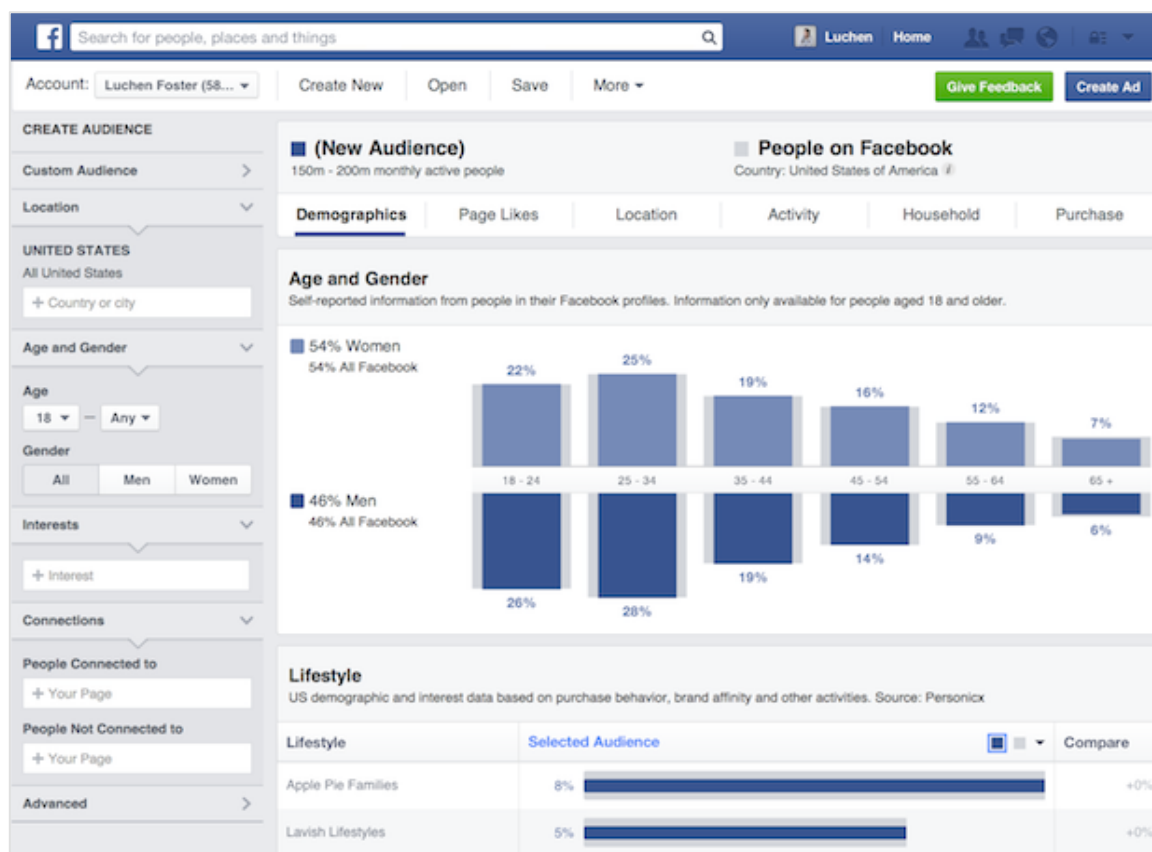


Figure 9 – Aperçu de l'outil Facebook Audience Insight
 (<https://www.facebook.com/business/news/audience-insights>)

Utilisez tous les autres leviers traditionnels à votre disposition pour promouvoir votre curation et lui donner un maximum de visibilité. Intégrez la promotion de votre curation via une newsletter par exemple.

o *Pour aller plus loin : la curation comme partie intégrante d'une stratégie netlinking ?*

La curation est, en fin de compte, une façon de générer du contenu ayant une réelle valeur ajoutée. Même si l'objectif principal de la curation n'est pas de générer des liens vers votre site, si votre contenu est réputé pour sa pertinence et son expertise dans un domaine, **les internautes n'hésiteront pas à le relayer sur des sites tiers.**

C'est également **un moyen de prendre contact et de nouer des relations** avec les auteurs des sites référencés dans votre curation et les influenceurs de votre secteur.

Etape 5 – Analysez vos contenus : quels KPI pour la curation ?

o *Learning : apprendre de ce qui marche et ce qui ne marche pas*

Une fois partagé, que se passe-t-il ? Votre contenu a-t-il bien fonctionné ? Pouvez-vous l'adapter à d'autres plateformes ? Ou encore utiliser une des 5 autres formes de curation listées le mois dernier pour retraiter une information partagée ?

Le rythme de publication est-il le bon ? Pouvez-vous améliorer vos relais sur les réseaux sociaux en ciblant mieux les plages horaires de diffusion ?

Autant de questions dont les réponses peuvent vous donner des indications de **ce qu'il faudra ajuster dans votre stratégie de curation.**

Analysez donc les résultats de chaque article de curation. Créez-vous un reporting pour pouvoir comparer vos données d'un article à l'autre. Pour cela munissez-vous d'outils d'analyse en fonction des KPI définis.

o Exemples de KPI à suivre

Suivant les objectifs que vous aurez définis en amont de votre stratégie de curation, les KPI peuvent être les suivants :

- Trafic
 - o Par article ;
 - o Part du trafic issu des articles de curation.
- Mesure d'engagement de l'audience
 - o Temps passé sur la page ;
 - o Nombre de pages visitées ;
 - o Taux de rebond ;
 - o % de nouveaux visiteurs.
- Social
 - o Likes/émotions, tweets, épingles ;
 - o Partages, commentaires ;
 - o Nombre de nouveaux fans acquis.
- Liens acquis
- Génération de leads & ventes

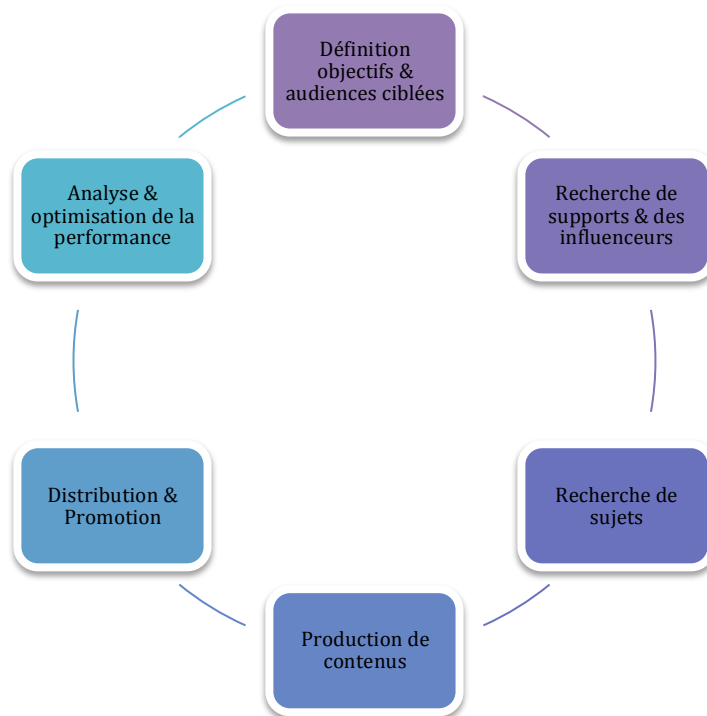


Figure 10. Une stratégie de curation en résumé

Quels outils de curation choisir ?

Automatisez autant que possible. Voici une sélection d'outils pour vous aider dans la curation. Notre conseil : testez-en plusieurs pour voir lesquels vous conviendront le mieux, en termes de résultats de recherche mais aussi en terme d'ergonomie et de fonctionnalités.

Cette liste n'est pas exhaustive, mais elle donne déjà une bonne base pour commencer voire optimiser sa curation :

MindMeister : <https://www.mindmeister.com/> (outil de mindmapping) ;

Coggle.it : <https://coggle.it/> (outil de mindmapping) ;

Text2mindmap : <https://www.text2mindmap.com/> (outil de mindmapping) ;

Google Alerts : <https://www.google.fr/alerts> (être alerté de la publication de contenus susceptible de vous intéresser) ;

Feedly : <https://feedly.com/> (agrégateur de fils RSS et de contenus par thématique) ;

Digg Reader : <http://digg.com/reader> (agrégateur de fils RSS et de contenus par thématique) ;

Flipboard : <https://flipboard.com/> (agrégateur de contenus) ;

Paper.li : <http://paper.li/> (agrégateur de contenus) ;

HubPages : <http://hubpages.com/> (agrégateur de contenus) ;

Delicious : <http://del.icio.us/> (service de bookmarking, classement et partage de contenus) ;

Pearltrees : <http://www.pearltrees.com/> (explorer du contenu selon des centres d'intérêts) ;

Stumbleupon : <http://www.stumbleupon.com/> (outils de recommandation de contenu par intérêt) ;

Buzzsumo : <http://buzzsumo.com/> (outil de stratégie de contenu et de curation) ;

Curata : <http://www.curata.com/> (plateforme de curation) ;

Scoop.it : <http://www.scoop.it/> (plateforme de gestion de contenu de curation) ;

Storify : <https://storify.com/> (plateforme de gestion de contenu de curation) ;

Pocket : <https://getpocket.com/> (outil de gestion de contenu) ;

Evernote : <https://evernote.com/intl/fr/> (permet de partager des notes en mode public) ;

Ahrefs : <https://ahrefs.com/> (possède une partie dédiée à l'exploration de contenus) ;

IFTTT : <https://ifttt.com/> (outil pour automatiser des tâches) ;

Mais aussi, [Sharetally.co](http://shareally.co) et [Sharedcount.com](http://sharedcount.com) pour analyser le nombre d'interactions pour une ou plusieurs URL données.

Enfin, nous pourrions ajouter également à cette liste les réseaux sociaux et diverses plateformes dans lesquelles vous pouvez faire des recherches avancées pour trouver des sources de contenus pertinents : Pinterest, Facebook, Twitter, LinkedIn, Instagram, Vimeo, Youtube, etc.

Sources

- (1) http://contentmarketinginstitute.com/wp-content/uploads/2014/10/2015_B2B_Research.pdf
- (2) <http://locowise.com/blog/facebook-reach-hits-9-month-high-with-the-latest-algorithm-tweak>



Anne-Victoire de Boisgelin, *Consultante SEO, Resoneo*

(<https://twitter.com/AnneVictoiredeB>)

Gephi et l'analyse de données pour le SEO (2ème partie)



Par Daniel Roch

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Gephi est un logiciel dont on entend parler régulièrement pour analyser la structure d'un site web. En réalité, Gephi est un logiciel d'analyse capable de traiter et de visualiser n'importe quel type de données, permettant ainsi d'approfondir notre analyse en référencement naturel. Nous allons donc voir dans cet article en deux parties comment ce logiciel fonctionne, ses qualités (mais également ses défauts) et comment on peut en tirer profit dans le cadre d'une stratégie SEO. Le mois dernier, nous avons abordé l'utilisation du logiciel Gephi. Puissant mais complexe, ce dernier permet d'analyser et surtout de visualiser des données. Dans cette seconde partie, nous allons donc aborder des cas concrets d'utilisation du logiciel dans une optique de référencement naturel.

Analyser les silos de son site

Il s'agit sûrement ici de la première utilisation pertinente que l'on peut faire avec Gephi. L'idée est de pouvoir visualiser si, oui ou non, le maillage interne est optimisé et si l'on a réussi à créer des sections cohérentes.

Pour faire cela, le premier moyen est de regarder le rendu initial que l'on a réalisé pour voir si des grappes de nœuds se démarquent ou non. La problématique est que cela est souvent assez difficile à visualiser.

Nous utiliserons donc la méthode dite de la modularité. Cet algorithme sert à déterminer et mettre en avant des groupes dans un tableau de données. Théoriquement, si les silos sont bien conçus, l'algorithme va les détecter automatiquement.

1^{ère} étape : calculez la modularité. Pour cela, utilisez dans la colonne de droite l'onglet « Statistiques », puis cliquez sur « Exécuter » pour la ligne « Modularité ».

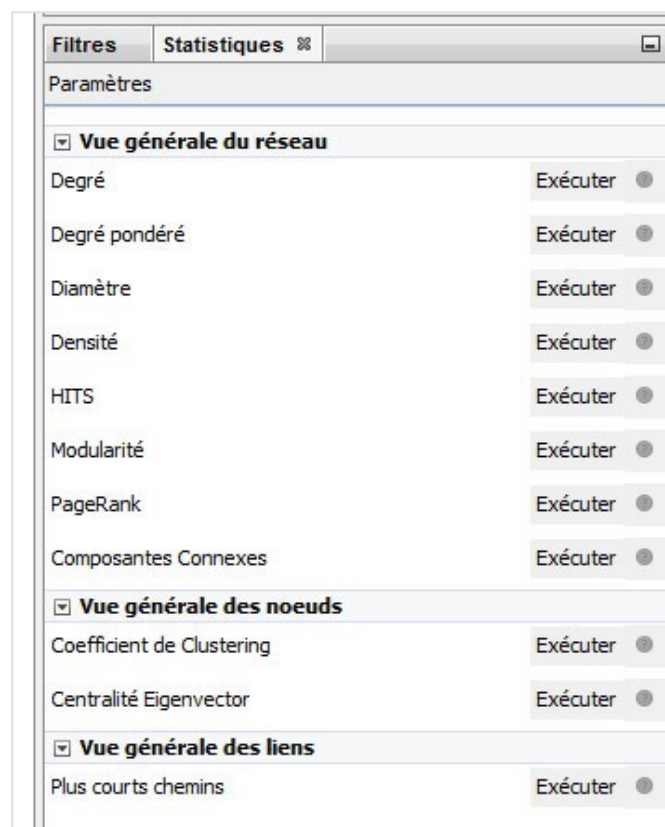


Fig. 1. L'onglet de statistiques

Dans la fenêtre qui s'ouvre, cochez « Aléatoire » et « Utiliser le poids ».

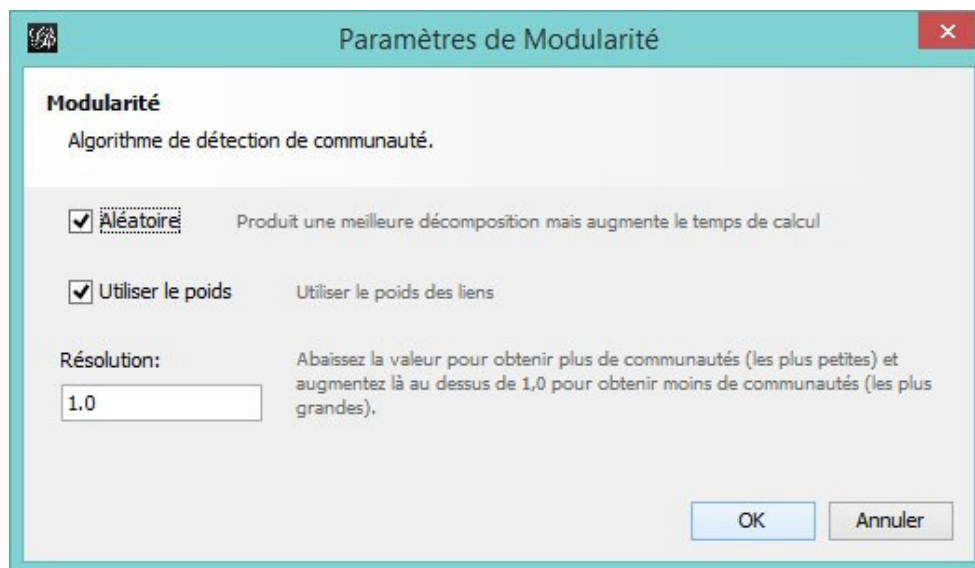


Fig. 2. Le paramétrage de la modularité

Une fois cette étape réalisée, la donnée « Modularité » est prête à l'emploi. Dans la colonne de gauche, sélectionnez dans « Aspect » :

1. l'onglet « Nœuds » ;
2. L'icône de couleur ;

3. Dans la liste déroulante, l'attribut « Modularité ».

Cliquez enfin sur « Appliquer ». Si les silos sont mal conçus, le graphique ressemblera à un sapin de Noël. Si à l'inverse ils sont bien créés, le logiciel Gephi va colorer plus ou moins parfaitement vos différentes sections de manière homogène, comme sur la figure 3.

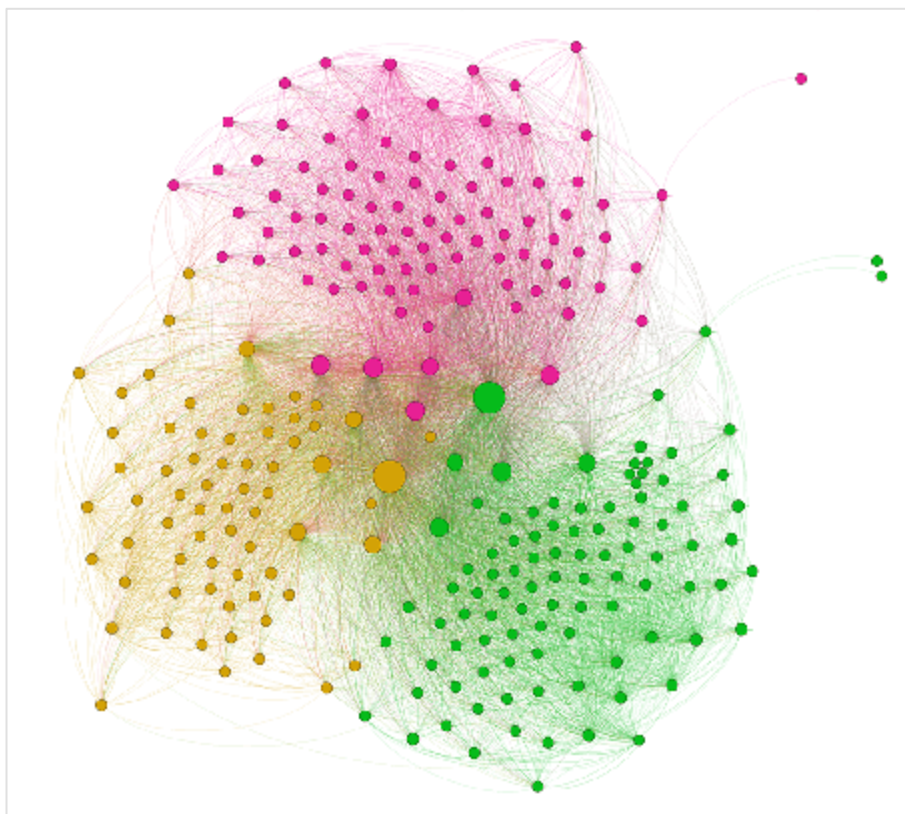


Fig. 3. Un exemple de rendu d'un site ayant 3 silos.

Silos et liens réciproques pertinents

Tout silo est améliorable. Si on prend l'exemple précédent, le site est bien conçu car trois sections sont visibles. Le souci est que dans ces sections, l'information n'est pas hiérarchisée (page principale, pages secondaires, pages de troisième niveau...). On peut donc améliorer notre travail.

Pour cela, on peut partir du dernier graphique (fig.3) pour chercher les endroits où notre hiérarchie est mauvaise, mais ce n'est pas forcément le plus efficace. Nous allons donc utiliser une autre analyse : les liens réciproques entre les pages (A faisant un lien vers B qui fait également un lien vers A).

Normalement, chaque silo doit contenir un grand nombre de liens réciproques internes, et peu de liens réciproques avec les autres silos. Et dans chaque silo, le nœud principal doit être celui faisant le plus de liens réciproques. Pour cette analyse, nous utiliserons cette fois-ci les filtres de la colonne de droite. On peut ainsi filtrer visuellement tous les nœuds

et les liens selon de très nombreux paramètres dans la partie « Bibliothèque », et les filtres actuellement appliqués seront visibles dans la section « Requêtes ».

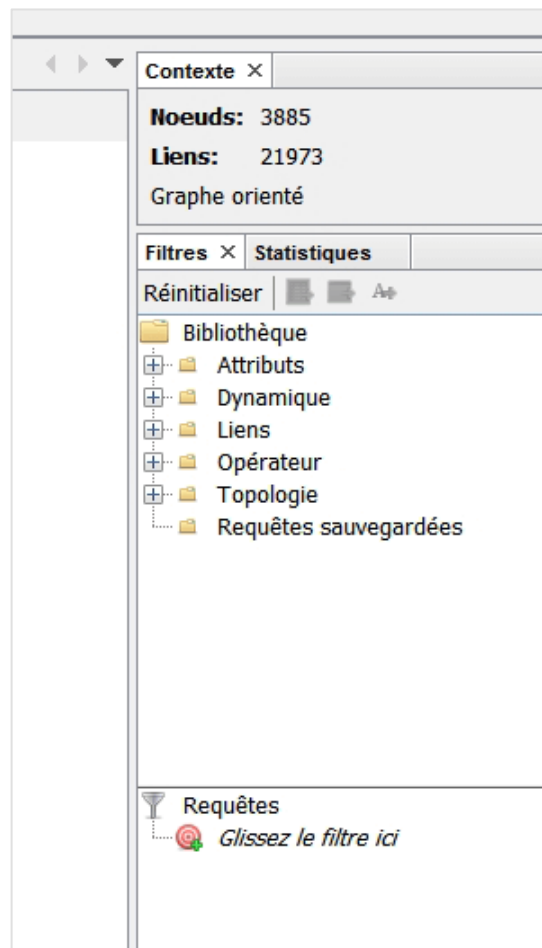


Fig 4. Les filtres dans Gephi

Dans « Liens », faites un double clic sur « Mutual Edge », et cliquez sur « Filtrer » en bas à droite. Cela peut ainsi vous donner ce type de rendu où l'on peut voir tous les liens réciproques, dont :

- Ceux dans un silo (que l'on doit conserver) ;
- Ceux entre les silos (que l'on doit réduire).

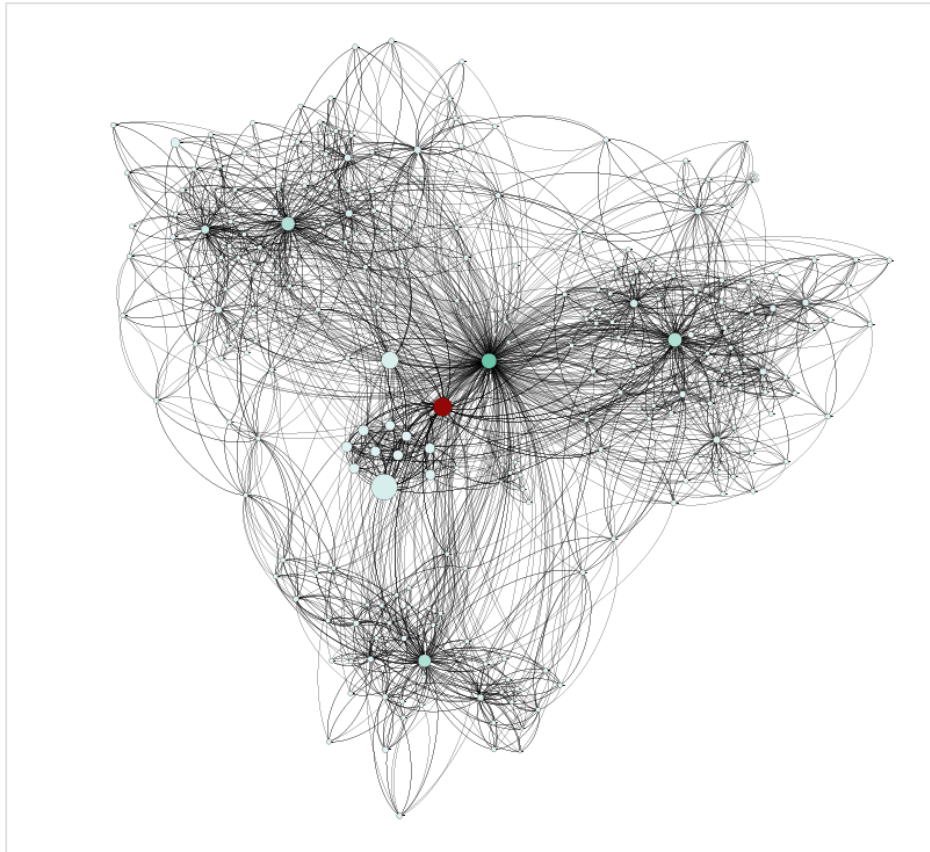


Fig. 5. La visualisation des liens mutuels

On voit d'ailleurs que la page située au centre est la page en couleur est celle avec le plus de liens réciproques. Il s'agit de la page « Plan du site », preuve d'un potentiel maillage trop important vers cette URL.

Analyser la répartition des pages populaires

Il est également possible de vouloir faire ressortir le nombre d'impressions de chaque nœud dans les moteurs de recherche. Pour rappel, cette donnée a été obtenue via la connexion à l'API Search Console dans Screaming Frog Spider SEO.

Dans cet exemple, nous avons coloré les nœuds selon le nombre d'impressions, tandis que la taille du nœud correspond à son degré interne. Ainsi, si le rond n'est pas coloré (peu d'impressions) mais que ce dernier est grand, cela veut dire que cette URL a une place importante dans la structure du site, mais qu'elle n'est pas visible dans les moteurs de recherche. Ce serait donc une URL importante à travailler en priorité au niveau de son contenu, ou alors on devrait réduire le nombre de liens vers cette URL.

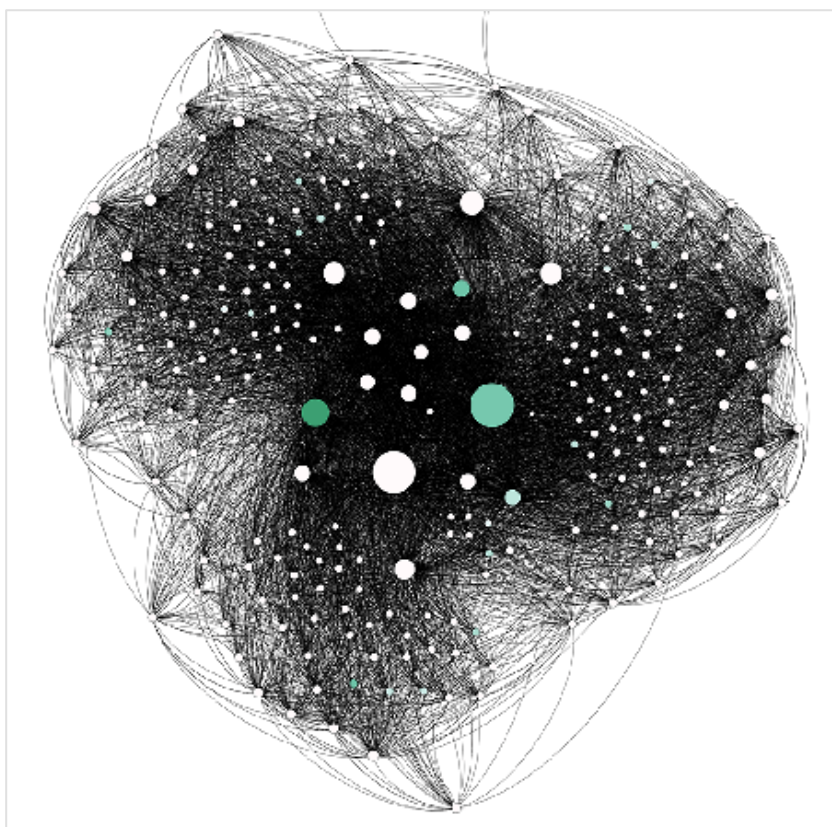


Fig. 6. Le croisement entre l'importance du nœud dans la structure et sa visibilité

D'ailleurs, concernant la taille des nœuds, sachez que vous avez plusieurs solutions possibles. Gephi vous propose en effet d'utiliser différents attributs :

- Le degré, c'est-à-dire un calcul brut du nombre de liens entre les nœuds ;
- Le PageRank, qui reprend le même concept que le degré en donnant plus de poids à la place du nœud dans la structure ;
- Eventuellement le Trust Flow et le Citation Flow de l'outil Majestic.

Pour les deux derniers, l'information vient de l'import du fichier CSV. Pour le PageRank, faites attention car il ne s'agit pas du PageRank de Google, mais d'une formule algorithmique différente. D'ailleurs, pour l'utiliser, vous devez d'abord calculer ce PageRank dans la colonne de droite.

Faire ressortir les contenus inutiles ou en erreur

Autre exemple dans la figure 7 : faire ressortir les pages en erreur ou inutiles (erreur 404, redirections 301...). On peut ainsi voir où se situent les URL qui polluent le référencement global du site analysé :

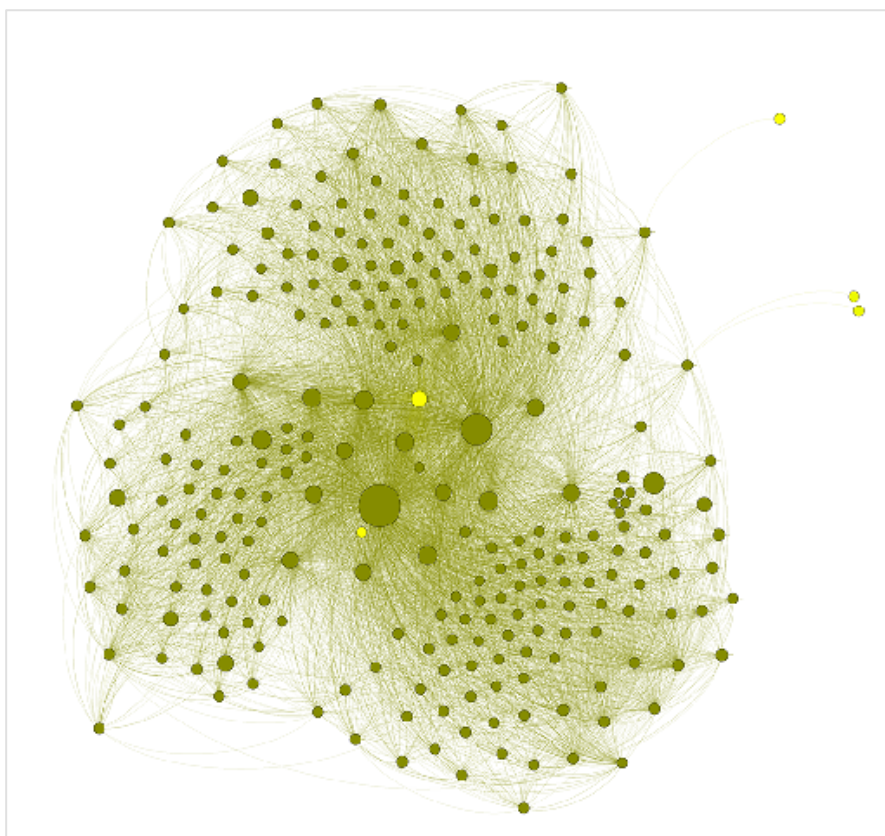


Fig. 7. Les URL malsaines apparaissent en jaune.

Faire des test SEO

Gephi peut également être utile pour faire des tests de référencement naturel, et notamment pour voir des corrélations ou non entre certaines données. Prenons un exemple :

Le nombre de mots de mes contenus a-t-il une influence sur les backlinks obtenus, et donc sur le Trust Flow ?

Pour répondre à la question, nous donnons donc à chaque rond une taille proportionnelle au Trust Flow, et nous les colorons de plus en plus foncé en fonction du nombre de mots présents :

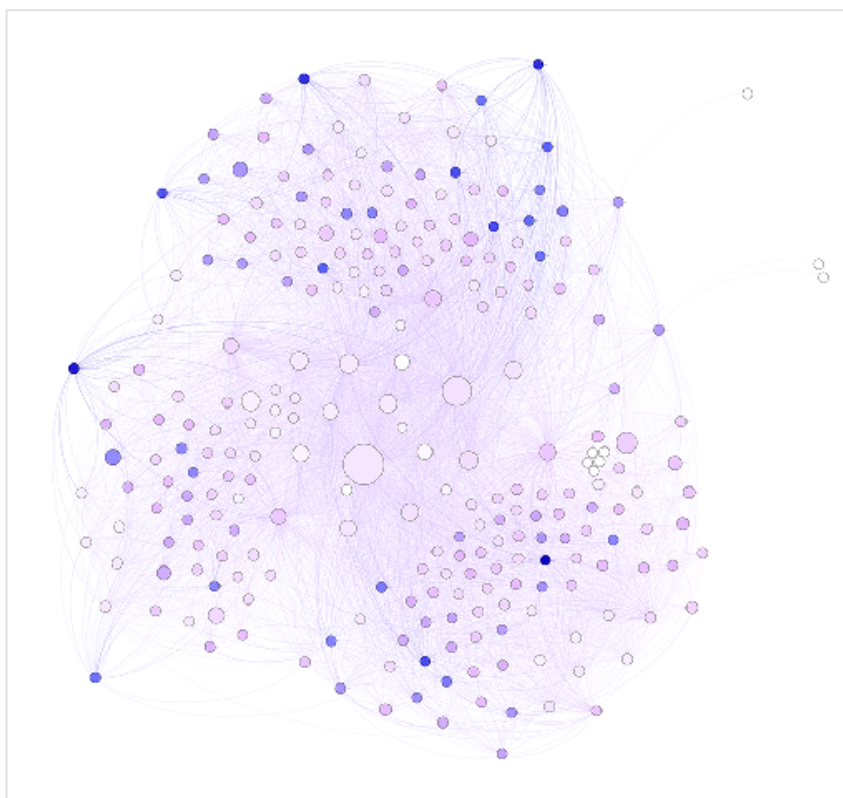


Fig. 8. Y a-t-il corrélation entre Trust Flow et nombre de mots dans le contenu ?

Le résultat est simple : il ne semble pas y avoir de corrélation entre ces deux données car dans le cas contraire, nous aurions des ronds foncés de grande taille.

Visualiser les domaines référents de chaque URL de mon site

Autre exemple : nous pourrions vouloir visualiser les pages de notre site ayant peu de domaines référents afin de mieux rediriger notre travail de netlinking. Pour cette analyse, nous devons utiliser une fois de plus nos filtres, avec cette fois-ci les « Partitions ».

Une fois la partition cochée sur le paramètre « Nombre de domaines référents » (ou « RefDomains » en fonction de comment vous l'avez appelé dans le fichier CSV), nous allons pouvoir cocher ou décocher les chiffres qui nous intéressent. En bas à droite, juste avant le bouton « filtrer », vous verrez en effet des cases pour n'afficher que les nœuds ayant un nombre déterminé de RefDomain. Dans notre exemple, nous avons coché toutes les cases ayant un RefDomain supérieur ou égal à 10.

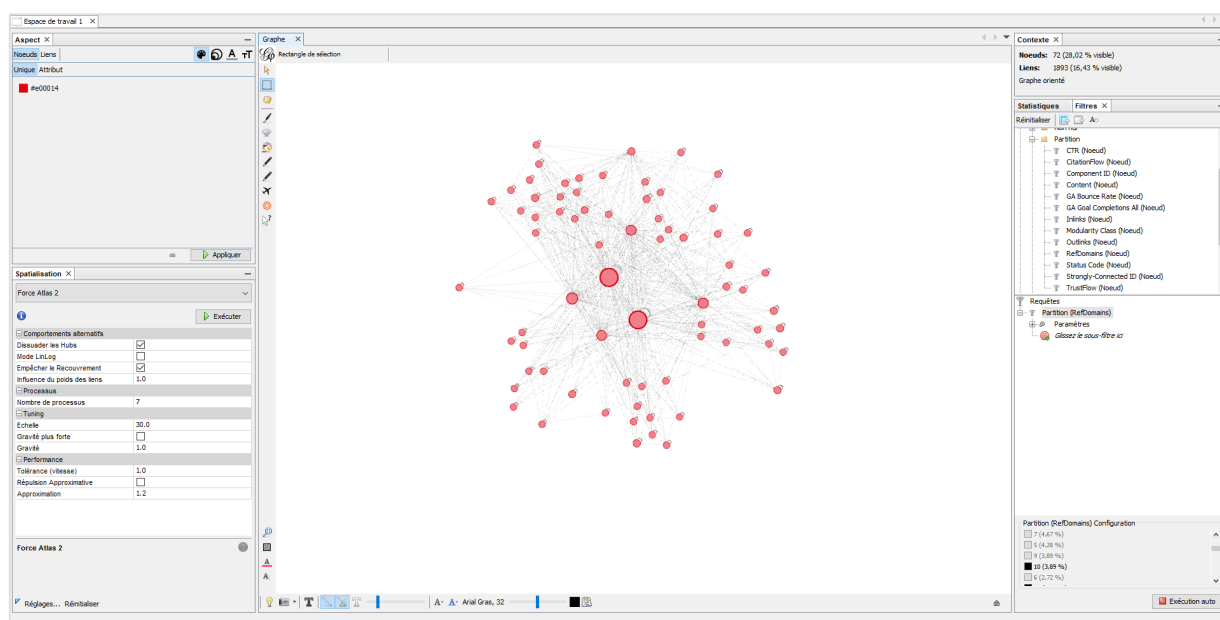


Fig. 9. Un exemple d'utilisation de filtres

Vous pouvez alors sélectionner tous les nœuds dans la fenêtre centrale, puis à gauche dans « Aspect » vous pouvez mettre en place une coloration unique de tous les nœuds. Vous pourrez ensuite cocher « Exécution auto » en bas à droite pour désactiver le filtre, tout en ayant mis en couleur les ronds qui correspondaient à votre filtre.

Dans la figure 10, nous colorons ainsi en rouge foncé les URL ayant un grand nombre de domaines référents.

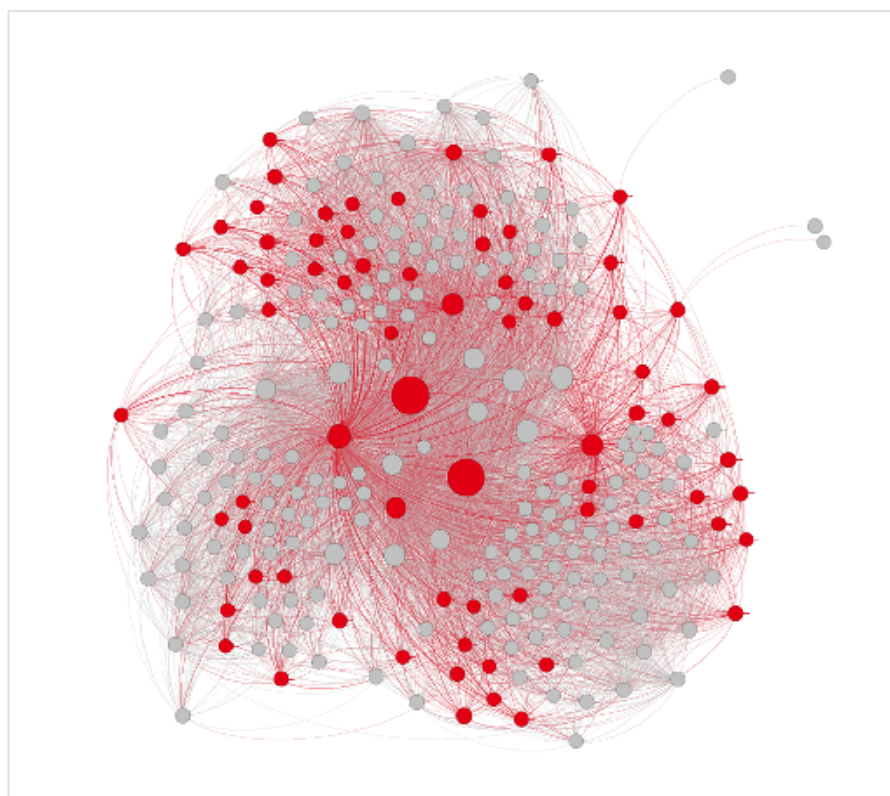


Figure 10. Les nœuds en rouge foncé reçoivent suffisamment de domaines référents

Visualiser les liens de chaque URL de mon site

Il est également possible d'analyser et de mieux visualiser le nombre de domaines référents pour chaque URL, et ainsi savoir vers quelles pages nous devrions créer des liens depuis de nouveaux domaines.

Pour cela, une nouvelle étape d'importation a été nécessaire. Sur Majestic SEO, nous allons exporter la liste des liens vers le domaine analysé. A partir de ce fichier, nous allons donc créer deux nouveaux fichiers :

- Un fichier de nœuds avec les sources de chaque backlink (ID, Label) ;
- Un fichier de liens (Source, Target, Label).

On va également garder notre premier fichier de nœuds, celui de toutes nos URL de notre site. Commencez un nouveau projet, puis :

- Importez le fichier de nœuds initial, c'est-à-dire tous vos URL internes ;
- Importez le fichier de nœuds de vos backlinks ;
- Importez le fichier de liens de vos backlinks.

Une fois l'importation réalisée, on peut alors réutiliser nos outils de spatialisation comme au départ. La problématique que vous allez rencontrer est que l'on ne peut différencier nœud interne (URL de mon site) et nœud externe (source des liens). Vous pouvez donc soit procéder avec un filtre comme dans l'exemple précédent, soit utiliser une autre méthode de coloration plus simple. Pour cette deuxième solution :

- Rendez-vous dans le tableau de données ;
- Classez le tableau par ordre alphabétique d'URL ;
- Sélectionnez toutes vos URL internes ;
- Clic-droit puis « Sélectionner dans la vue d'ensemble » ;
- Puis coloration manuelle unique, ici en rouge dans notre exemple.

Dans la figure 11, le nombre de domaines référents de chacune de mes URL en rouge est bien plus parlant.

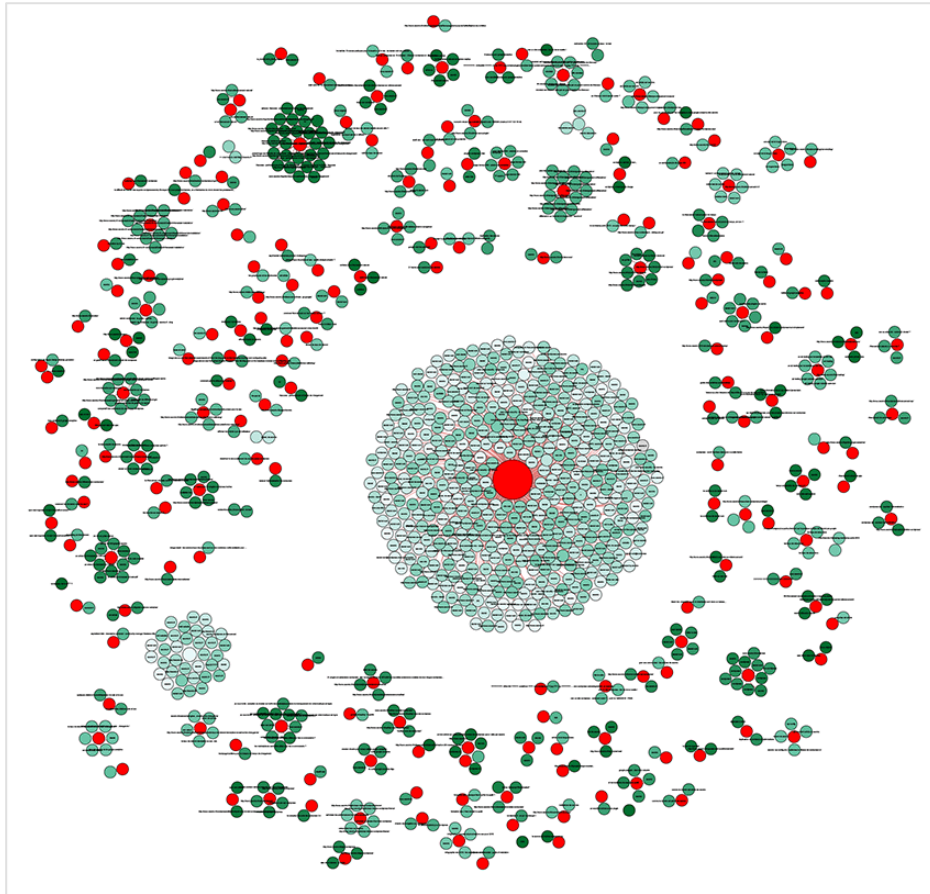


Fig. 11. Les nœuds en rouge sont les pages de mon site. Tous les autres sont les domaines référents associés.

Conclusion

Les exemples donnés ici ne sont justement que des exemples. On pourrait faire bien plus. Il est possible d'utiliser Gephi dans une analyse Webmarketing au sens large, par exemple avec l'analyse des partages sociaux ou des conversions réalisées.

Bien entendu, il n'est pas conseillé de systématiquement recourir à ce type de logiciel, car l'analyse peut prendre du temps. Et surtout, on peut parfois avoir des données tout aussi lisibles dans un fichier Excel.

Malgré cela, Gephi est un logiciel très utile et puissant, notamment lors d'audits complexes de référencement naturel. Seul réel défaut, et non des moindres : Gephi reste un logiciel complexe et peu stable, ce qui peut parfois nuire à votre travail d'analyse.



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

Le droit d'opposition opposé à Google



Par Alexandre Diehl

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Un récent cas jugé par le Tribunal de Grande Instance de Paris a mis en lumière la notion de droit d'opposition pour faire supprimer un contenu des index de Google. Mais, au-delà de ça, ce jugement nous permet de nous interroger sur le futur de nos lois et sur l'américanisation à venir, dans les années futures, de notre vision de la justice. Est-ce un bien ou un mal ?...

Notre bonne vieille loi « informatique et libertés » connaît de nombreux principes bien utiles pour les consommateurs et autres utilisateurs de moteurs de recherche. Parmi ceux-ci, le droit d'opposition octroyé légalement à toute personne physique pour s'opposer à un traitement automatisé de ses données personnelles. Ce droit, qui peut paraître extrêmement large et à portée illimitée, est en fait encadré et finalement, peu utilisé. Google a pourtant pu récemment en faire l'expérience en se voyant imposer, par le Président du Tribunal de Grande Instance de Paris, la suppression d'un référencement naturel sur ce fondement.

L'affaire

Une personne bien en vue a été accusée d'être impliquée dans un scandale sexuel (<http://www.lemondedudroit.fr/droit-a-entreprises/technologies-de-linformation/218915-google-suppression-dun-referencement-sur-le-moteur-de-recherche-en-vertu-du-droit-dopposition.html>). S'estimant accusée à tort, cette personne a porté plainte, saisi le webmaster des sites relayant ces informations (qui ont refusé de statuer sur la demande de l'intéressé), saisi le site www.internet-signalement.gouv.fr (ce qui n'a, évidemment, servi à rien), puis la CNIL (qui est contrainte par des délais incompressibles), puis, enfin, a demandé à Google (vous pouvez le faire par mail : removals@google.com et via une page dédiée : <https://support.google.com/legal/answer/3110420?rd=1>) de déréférencer les contenus « diffamatoires ».

Google a donné une réponse que l'on obtient très fréquemment (voire systématiquement), à savoir que les sites concernés « contiennent des informations [...] concernant [Monsieur X] qui sont pertinentes et à jour », estimant, en conséquence, « que la référence à ce contenu dans [ses] résultats de recherche est justifiée par l'intérêt du grand public à y avoir accès ». Cette notion d'intérêt du grand public est très souvent mise en avant par Google dans ce type de cas.

La personne concernée a alors saisi, en procédure de référé (procédure d'urgence), le Président du Tribunal de Grande Instance de Paris pour que celui-ci ordonne à Google de supprimer les référencements naturels des contenus présumés diffamatoires. Google

Inc, qui est comme tout le monde le sait, seul exploitant du moteur de recherche, a demandé (puisque l'avocat du plaignant a « oublié » de l'assigner) à intervenir au procès.

Dans le cadre de l'instance, le juge devait trancher entre la protection de la vie personnelle / données personnelles de la personne concernée et la protection de l'information légitime du public. La question n'était donc pas de savoir si l'information était vraie ou pas, mais uniquement de savoir lequel de ces deux intérêts doit-on le plus défendre :

- D'un côté, la loi « informatique et libertés » précise que toute personne physique a le droit de s'opposer pour des motifs légitimes, à ce que des données à caractère personnel la concernant fassent l'objet d'un traitement.

La CJUE a également rappelé, dans un arrêt Google Espagne du 14 mai 2014 (<http://curia.europa.eu/juris/liste.jsf?num=C-131/12&language=FR>), que les données personnelles traitées doivent être nécessaires au regard des finalités pour lesquelles elles sont collectées ou traitées. La CJUE avait poursuivi en précisant que Google ne peut plus traiter des données personnelles « *lorsqu'elles apparaissent inadéquates, qu'elles ne sont pas ou plus pertinentes ou sont excessives au regard de ces finalités et du temps qui s'est écoulé* » ;

- D'un autre côté, les droits fondamentaux à la liberté d'expression et d'information sont énoncés à l'article 10 de la Convention des droits de l'homme et à l'article 11 de la Charte des droits fondamentaux de l'Union européenne. Ce droit comprend la liberté d'opinion et la liberté de recevoir ou de communiquer des informations ou des idées sans qu'il puisse y avoir ingérence d'autorités publiques et sans considération de frontières.

Dans le cadre de sa demande de refus de retirer le lien, Google Inc a soutenu que l'auteur du contenu litigieux est susceptible d'avoir agi comme un « lanceur d'alerte » et que l'information querellée, dont la véracité n'est pas démontrée, a été véhiculée dans l'intention légitime d'alerter le public.

Pour rejeter l'argumentaire de Google, le juge a relevé que les termes utilisés par le contenu est « *gravement attentatoire à la réputation du requérant* », notamment en mettant en rapport la situation professionnelle précise de Monsieur X. avec le scandale sexuel sur mineur dénoncé, « *sans que l'évocation de sa situation professionnelle, dépourvue de lien avec le « scandale » dénoncé, apparaisse motivée par une autre intention que celle de satisfaire une vindicte personnelle contre lui* ».

En conséquence, le juge a accepté la demande de Monsieur X en concluant que « *le référencement du lien [concerné] a directement porté atteinte au droit à la protection de ses données personnelles, sans que cette atteinte soit légitimée par le droit à l'information légitime du public* ».

Le droit d'opposition et son sort prochain

Les termes de l'article 38 de la loi « informatique et libertés » disposent : « Toute personne physique a le droit de s'opposer, pour des motifs légitimes, à ce que des données à caractère personnel la concernant fassent l'objet d'un traitement ». La notion de motifs légitimes renvoie à une appréciation au cas par cas du juge (ou de la personne devant prendre la décision finale). Le cas présenté est un excellent exemple de son application.

Le nouveau Règlement (UE) 2016/679 du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données entrera en vigueur le 25 mai 2018. Après d'âpres discussions (surtout avec les représentants des intérêts Américains d'ailleurs), un texte reprenant les grands thèmes de la directive a finalement (et heureusement) été proposé. En l'état, les droits des personnes concernées ont été réaffirmés. Le droit d'opposition a été même renforcé. Ainsi, l'article 21 dispose :

« 1. La personne concernée a le droit de s'opposer à tout moment, pour des raisons tenant à sa situation particulière, à un traitement des données à caractère personnel la concernant fondé sur l'article 6, paragraphe 1, point e) [traitement nécessaire à l'exécution d'une mission d'intérêt public] ou f) [traitement nécessaire aux fins des intérêts légitimes poursuivis par le responsable du traitement], y compris un profilage fondé sur ces dispositions. Le responsable du traitement ne traite plus les données à caractère personnel, à moins qu'il ne démontre qu'il existe des motifs légitimes et impérieux pour le traitement qui prévalent sur les intérêts et les droits et libertés de la personne concernée, ou pour la constatation, l'exercice ou la défense de droits en justice.

2. Lorsque les données à caractère personnel sont traitées à des fins de prospection, la personne concernée a le droit de s'opposer à tout moment au traitement des données à caractère personnel la concernant à de telles fins de prospection, y compris au profilage dans la mesure où il est lié à une telle prospection.

3. Lorsque la personne concernée s'oppose au traitement à des fins de prospection, les données à caractère personnel ne sont plus traitées à ces fins.

(...) »

Les limites du système du droit français

Malgré tout, force est de constater que, sauf exception, Google et plus largement les moteurs sont très souvent en position de juges. En effet, lorsqu'une personne s'estime accusée ou atteinte dans ses droits, elle contacte les moteurs en vue de déréférencement en avançant des arguments afin que sa demande soit honorée. En pratique, c'est donc le moteur qui va devoir trancher si oui ou non, un contenu sera déréférencé.

Or, Google, Microsoft ou autre Qwant ne sont pas des juges professionnels qui ont reçu un enseignement et des instructions, conformément à des règles de droit et une

politique pénale d'un Etat. Sans entrer dans de grands débats philosophiques, il est important de rappeler que depuis l'Antiquité, en passant par Montesquieu pour aller jusqu'aux théories libérales de la moitié du 20ème siècle, le juge a un rôle central dans la gestion des problèmes car il statue au nom du peuple et conformément aux intérêts communs de la Nation (ou d'un Etat pour les dictatures). Or, Google ne représente aucun peuple et certainement pas les intérêts d'une Nation, mais bien uniquement les siens. Et, nous pensons que l'intérêt premier de Google est bien de gagner de l'argent et non pas de respecter certaines philosophies d'une règle juridique...

En conséquence, lorsque le législateur a décidé que les hébergeurs devenaient juges, il a transféré de facto ce pouvoir à ceux-ci et a changé profondément la protection des intérêts. Par exemple, l'article 6-I-2 de la loi « Confiance en l'économie numérique » du 21 juin 2004 précise que les hébergeurs peuvent voir leur responsabilité engagée s'ils n'ont pas promptement mis hors ligne un contenu « illicite » lorsqu'on leur signale. Dans une démocratie, c'est le juge qui détermine, au cas par cas, ce qui est illicite ou non (d'après la fameuse séparation des pouvoirs), et non pas à Google. Pourtant, en écrivant ce principe dans la loi de 2004, le législateur a mis de manière évidente, la détermination du contenu « illicite » à la charge des moteurs. Ainsi, aujourd'hui, si on veut faire retirer un contenu « illicite » (en tout cas, qu'on estime illicite), on ne va pas voir un juge, on demande à Google et ce n'est que si Google ne répond pas favorablement qu'on va se plaindre à un juge du fait que Google a refusé. **En d'autres termes, la justice est devenue une sorte de cour d'appel de Google.**

Ce principe d'accroissement des pouvoirs des moteurs, et plus largement des intermédiaires, est venu des Etats-Unis qui, dès 1997, l'appliquaient dans leur droit interne. L'idée était simple : le marché (dominé par les principaux acteurs privés) se régule lui-même, y compris pour des questions juridiques. Ce principe philosophique est typiquement américain, dans la droite ligne des principes directeurs de l'Ecole de Chicago, et surtout, parfaitement accepté sur le continent : dans cet environnement ultra-libéral, il est logique et normal que ce soit le marché qui régule, l'Etat (et notamment la justice) n'est là qu'en tant que simple gendarme extérieur au marché pour sanctionner si les acteurs (par exemple Google) vont trop loin. Aucun Américain ne sera choqué que ce soit Google qui décide si oui ou non on met hors ligne un lien vers un article.

Les principes à venir, les futures lois sont de cette ampleur et de cette philosophie : les acteurs seront de plus en plus responsables de dire le droit et de jouer les juges sur leur marché, la justice ne sera là que pour contrôler si ces acteurs ne sont pas trop mauvais comme juges. L'environnement juridique s'américanisant à très grande vitesse (et encore beaucoup plus quand le TAFTA entrera en vigueur d'ici 2020 ou 2021), il est à craindre qu'il faille s'habituer et vite apprendre et vivre « à l'Américaine ». En d'autres termes, il faudra vite apprendre à maîtriser « la loi Google » plutôt que la loi française et à comprendre que ce sont les intérêts des intermédiaires (type Google) qui prévaudront sur ceux des peuples et Nations...



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

Revue d'outils moteurs et SEO



Par Olivier Andrieu

Cette rubrique signale chaque mois 5 outils (site web, logiciel, plug-ins Firefox ou Chrome...) 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 : **Personal Blocklist**

URL :

<https://chrome.google.com/webstore/detail/personal-blocklist-by-goo/nolijnfnkgaikbjbdaogikpmpbdcdef>

Domaine : Recherche / SEO

Type : Extension chrome

Descriptif : Une petite extension Chrome qui permet de bloquer certains résultats dans les SERP (votre site, ceux de vos concurrents, etc.). Les résultats Google seront les mêmes que d'habitude, mais sans les sites que vous avez listés au préalable.



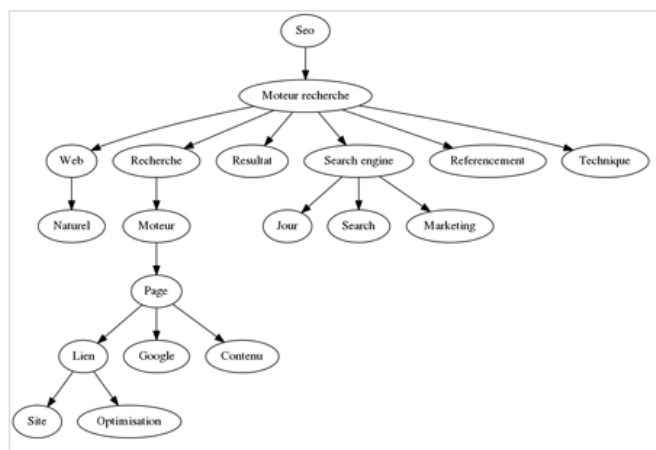
Nom : **SEO Quantum**

URL : <http://www.seoquantum.com/>

Domaine : Recherche / SEO

Type : Outil en ligne

Descriptif : Un outil qui va vous aider à créer vos corpus lexicaux pour proposer des textes plus riches sémantiquement à vos lecteurs et... à Google ! Vous tapez un mot clé ou une expression et l'outil vous en propose d'autres, connexes.



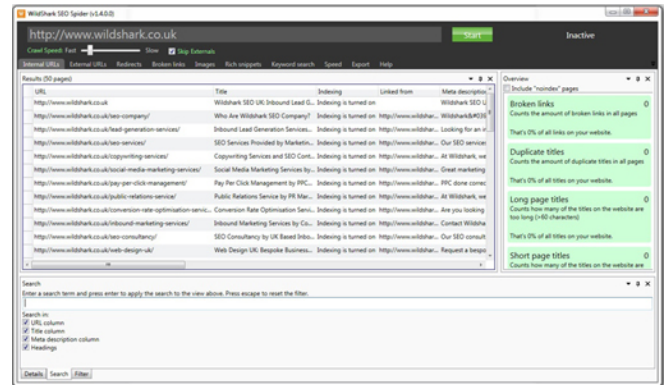
Nom : WildShark SEO Spider Tool

URL : <https://www.wildshark.co.uk/seo-spider/>

Domaine : Recherche / SEO

Type : Logiciel

Descriptif : Un outil de type "spider / crawler" (similaire à Screaming Frog, Botify ou autres Xenu Link Sleuth) qui va "aspérer" et explorer votre site et vous fournir bon nombre d'informations sur ce qui se passe bien. Ou pas... Gratuit. Disponible uniquement sur PC.



Nom : SiteImprove

URL : <http://siteimprove.com/fr/>

Domaine : Recherche / SEO

Type : Outil en ligne

Descriptif : L'outil se présente comme un système de correction de site web : fautes d'orthographe, accessibilité, normes du Web, SEO, performances, etc. Seul inconvénient : le site web qui le présente... est truffé de fautes d'orthographe et de mauvaises traductions :-)
Voir par exemple : <http://siteimprove.com/fr/products/>.



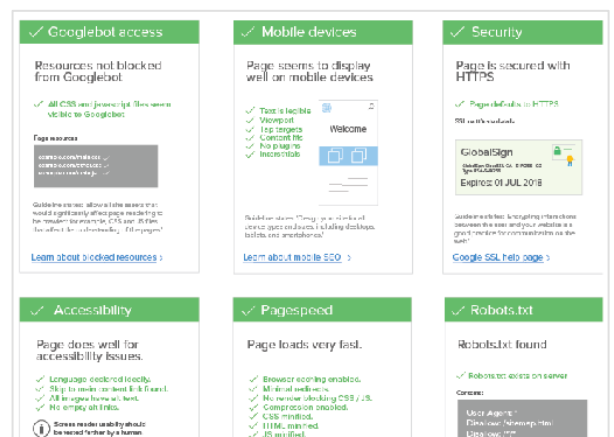
Nom : Varvy

URL : <https://varvy.com/>

Domaine : SEO

Type : Outil en ligne

Descriptif : Un petit outil en ligne qui vérifie que l'URL que vous lui fournissez suit bien les consignes ("guidelines") Google : HTTPS, temps de chargement, compatibilité mobile, ressources bloquées, etc. Bref, êtes-vous "un bon élève Google ?". Ce site vous le dira... :-)



Revue d'URL



Par Olivier Andrieu

Les articles de notre veille sur les moteurs de recherche et le référencement en français et en anglais, ainsi que le dossier du mois (lorsque l'actualité fait ressortir un fait marquant)...

Revue d'URL en français... ■ ■ ■

- Des fondamentaux du « Search » au SEO de RankBrain !

J'ai invité Françoise Halper, ma co-speaker au SMX, pour partager les réflexions que lui ont inspirées cette conférence, notamment sur l'avenir du métier de SEO revu par RankBrain, l'intelligence artificielle de Google. Attention, remise en cause assurée sur les stratégies de contenus... Le clavier est maintenant à Françoise...

<http://blog.axe-net.fr/fondamentaux-search-seo-rankbrain/>

- L'avenir de la recherche sur Internet

L'avenir du search passe par la recherche vocale et les assistants personnels virtuels. Un nouveau terrain de jeu pour les professionnels du marketing...

<http://www.journaldunet.com/ebusiness/le-net/1177829-l-avenir-de-la-recherche-sur-internet/>

- L'art et la manière d'obtenir des backlinks facilement

Avant toute chose, vous devez vous demander pourquoi avoir choisi une image un peu provocante et sexy pour illustrer mon article ? c'est simple et c'est pour 2 raisons...

<http://www.webinfoconseils.com/des-backlinks-pour-le-referencement/>

- Le référencement par l'expérience client – Livre Blanc Brioude Internet

Envie de connaître les étapes clés pour booster l'expérience client et améliorer votre référencement ? L'expérience client c'est l'atout fidélisation. Au travers du parcours d'achat apprenez à mieux connaître vos visiteurs. D'où viennent-ils ? qui sont-ils ? qu'attendent-ils ? SEO, Google Analytics, Marketing automation et bien d'autres, découvrez les outils et techniques qui vous permettront d'optimiser l'expérience client...

<http://www.brioude-internet.fr/le-referencement-par-l-experience-client-livre-blanc-brioude-internet.html>

- 6 moyens de booster votre référencement local

Commençons par quelques chiffres pour vous permettre de vous rendre compte de l'importance d'être visible localement sur la toile (en espérant que vous n'ayez pas le vertige)...

<http://www.e-alchimie.fr/referencement-local/>

- 10 façons de promouvoir votre contenu

Certains ont des blogs dans lesquels ils ont mis une énergie folle. Ils y publient constamment du contenu avec des articles de grande valeur et y disent des choses intéressantes pour leur audience. Et pourtant, même après cinq semaines, ils n'ont aucune visite si ce n'est au plus, celui de leur cousin, leur grand-mère ou leur conjoint...

<http://www.webmarketing-com.com/2016/06/22/47928-10-petites-facons-de-promouvoir-contenu>

- Techniques de Negative SEO : comment se protéger contre des pirates !

Petit rappel : Le Negative SEO est un concept anglophone qui désigne des techniques visant à nuire au positionnement d'un concurrent sur les moteurs de recherche (Google, Bing,...)...

<https://www.zoneref.com/negative-seo/>

- Du mot clé à l'intention de recherche

La base de données de 100 millions de mots-clés de YOODA INSIGHT, représentant 6,3 milliards de recherches mensuelles sur Google, nous permet une plongée dans les comportements de recherche des internautes...

<http://blog.yooda.com/seo/4998-mot-cle-intention-recherche/>

- Entretien avec le directeur de Google Research Europe

Emmanuel Mogenet est le directeur du Google Research Europe situé à Zurich en Suisse. Il a accueilli plusieurs journalistes, dont celui (ou celle) de l'hebdomadaire Le Point, pour l'inauguration du centre de recherche...

<http://moteurzine.org/2016/06/27/entretien-directeur-de-google-research-europe/>

- Votre trafic SEO a progressé et vous n'avez rien vu !

Vous l'aurez peut-être remarqué lors de l'analyse de vos leviers d'acquisition du mois de Mai 2016: une part significative du trafic provient maintenant d'un « site référent » nommé « com.google.android.googlequicksearchbox ». De quoi s'agit-il réellement ? Quelle source de trafic se trouve remplacée par celle-ci ? Comment requalifier ces visites dans le bon levier, si jamais tel est votre souhait ?...

<http://blog.resoneo.com/2016/06/votre-traffic-seo-a-progresse-et-vous-n-avez-rien-vu/>

- .htaccess : mettre en place une page 404 personnalisée

Voici plusieurs techniques pour afficher sa propre page d'erreur en passant par le fichier .htaccess...

<http://www.journaldunet.com/solutions/pratique/google-analytics-et-seo/18987/htaccess-mettre-en-place-une-page-404-personnalisee.html>

- Un test sur « cocon sémantique » montre les limites de RankBrain

Le site de Patrice Krysztofak, à savoir sitepenalise.fr, a lancé un test pour tenter de remonter sur l'expression « cocon sémantique » très en vogue dans le milieu du SEO depuis plusieurs années. Démocratisée par Laurent Bourrelly il y a déjà plusieurs années, le « cocon sémantique » est une stratégie SEO on-site optimisant le contenu et le maillage avec des objectifs de positionnement précis...

<http://www.digitalkeys.fr/test-cocon-semantic-montre-limites-de-google-rankbrain/>

- Comment faire de l'analyse de fichiers de logs gratuitement ?

L'analyse de fichiers de logs vous aide à comprendre la manière dont les moteurs de recherche crawlent un site et leur impact sur le SEO. Les informations qui en ressortent sont d'une grande aide pour améliorer votre habilité à être crawlé et vos performances SEO...

<http://fr.oncrawl.com/2016/comment-faire-de-analyse-de-fichiers-de-logs-gratuitement/>

- Dites non aux pratiques douteuses en matière de SEO

On ne le répétera jamais assez : si vous désirez positionner votre site web sur Google de manière durable, bannissez les pratiques se rapportant au « black hat ». S'il est parfois tentant de s'y frotter, les retombées négatives conséquentes doivent vous en dissuader...

<https://www.referenceur.be/dites-non-pratiques-douteuses-seo.html>

- Comprendre Google RankBrain et son impact pour le SEO

Système d'intelligence artificiel, RankBrain semble devenir un élément central de l'algorithme de Google pour le classement des pages de résultat. Si lors de son lancement officiel en 2015 l'impact de RankBrain ne concernait que 15% des requêtes que le moteur n'a jamais vu, Google semble très confiant en cette technologie aux capacités mystérieuses puisqu'elle serait désormais utilisée pour toutes les requêtes...

<http://blog.moov-up.fr/google-rankbrain-impact-seo/>

- Passer du HTTP au HTTPS fait-il chuter le revenu publicitaire du site ?

Décidément, passer d'un hébergement normal via HTTP à un hébergement sécurisé via HTTPS ne serait pas un long fleuve tranquille. Non seulement la procédure de migration est un peu complexe, mais un impact négatif pourrait aussi se faire sentir sur les revenus publicitaires via Adsense générés par les sites...

<http://www.arobasenet.com/2016/06/monetisation-site-adsense-https-3013.html>

- Comment améliorer votre SEO local ?

Avec la croissance de la recherche mobile et vocale, les requêtes locales n'ont cessé d'augmenter ces dernières années: 88% des utilisateurs mobiles en effectuent. En 2016, le SEO local semble plus exigeant que jamais, au vu des changements réguliers dans les résultats de recherche. Aujourd'hui, pour se faire remarquer des utilisateurs, votre SEO local doit être au top...

<http://fr.oncrawl.com/2016/comment-ameliorer-votre-seo-local/>

- Arrêtez d'utiliser Google Trends n'importe comment

Google Trends est un outil extrêmement intéressant, puisqu'il fournit des données sur l'utilisation de Google en temps réel. Google est devenu la barre de navigation de l'Internet : si vous avez besoin d'une information sur le Championnat d'Europe de football, il n'y a qu'à taper « Euro » dans votre navigateur, et vous obtiendrez immédiatement tous les détails sur les matches disputés jusque-là, ainsi que le calendrier des matches à venir...

<http://motherboard.vice.com/fr/read/arretez-dutiliser-google-trends-nimporte-comment>

- Le content spinning

On fait souvent des rencontres intéressantes dans le milieu du SEO, en discutant avec des professionnels qui sont particulièrement pointus dans un domaine par exemple. C'est le cas de Florence Le Vot, la directrice de RaDiance Conseil. Florence fuit les projecteurs mais certains font appel à son expertise pour des services très spécifiques comme la mise en place de réseaux de sites, le déploiement d'opérations mixant presse/influenceurs/SEO, la génération de contenu par traitement automatique du langage et le content-spinning, dont nous allons plus particulièrement parler dans cet article...

<http://blog.axe-net.fr/le-content-spinning/>

- Cocon multicoeur intentionnel, avec les metamots

Je commencerai ce billet par un grand merci pour l'intérêt que vous avez porté à l'article « Réaliser un cocon sémantique avec les meta-mots ». Aujourd'hui, je vous propose d'aller un peu plus loin. Les meta-mots vont vous parler...

<http://www.referencement-naturel-white-hat.fr/cocon-multicoeur-intentionnel-metamots/>

- SEO : à la conquête de la position #0

Lors du SMX Paris, un retour d'expérience SEO original a été partagé : celui de Total, qui, aidé de l'agence WAM, a réussi à apparaître dans la "Quick Answer Box". Une position encore plus visible que le premier résultat naturel de Google...

<http://www.journaldunet.com/solutions/seo-referencement/1180380-seo-a-la-conquete-de-la-position-0/>

- eBay a déployé le format AMP

eBay n'a pas attendu que Google fasse remonter des pages AMP d'e-commerçants dans ses résultats sur mobile pour adopter le nouveau format imaginé par Mountain View...

<http://www.journaldunet.com/solutions/seo-referencement/1181455-ebay-a-deploye-le-format-amp/>

- Sitemap et SEO : le duo gagnant ?

Dans la continuité des précédents dossiers consacrés au SEO, j'ai décidé de traiter un sujet de fond : le sitemap. Vous savez, ce petit fichier qui partage le web et suscite mine de rien beaucoup de questionnements autour de son rôle pour le SEO ?

<http://www.digitude.fr/dossiers-seo/sitemap-et-seo-le-duo-gagnant/>

- Phantom 4, nouvel update Google de juin 2016 (SEO)

Courant juin 2016, Google a procédé à des mises à jour de son algorithme ayant eu pour certains sites un gros impact. Rien n'a été officialisé, on parle donc de l'update "Phantom 4"...

<http://www.webrankinfo.com/dossiers/conseils/phantom-4>

- SEO : une nouvelle mise à jour Phantom lancée par Google en juin ?

Même si Google n'a rien annoncé, plusieurs spécialistes ont observé des mouvements inhabituels dans les résultats remontés par le moteur le mois dernier...

<http://www.journaldunet.com/solutions/seo-referencement/1181684-seo-une-nouvelle-mise-a-jour-phantom-lancee-par-google-en-juin/>

- Comment agir face à un problème en SEO ?

Vu que je suis en ce moment en vacances, j'ai laissé le clavier à Benjamin Descamps sur un sujet que j'affectionne, la patience et le SEO...

<http://blog.axe-net.fr/comment-agir-face-a-un-probleme-en-seo/>

- Comment passer son site WordPress en HTTPS

Depuis mai 2016, la Marmite est passée au HTTPS. Cela signifie que toutes les pages du site ont une URL débutant par https:// au lieu du classique http://. Vous allez me dire : « Très bien, mais qu'est-ce que ça change concrètement ? Pourquoi as-tu fait ce choix ? En quoi est-ce important ? ». Eh bien c'est ce que nous allons voir dans ce nouvel article...

<https://wpmarmite.com/wordpress-https/>

- Euro 2016 du SEO : quels sites ont triomphé dans les résultats de Google ?

Pour le JDN, l'outil spécialisé de MyPoseo a suivi la visibilité SEO des sites web sur des requêtes importantes lors de cet Euro 2016. Voici ceux qui ont brillé...

<http://www.journaldunet.com/solutions/seo-referencement/1181936-euro-2016-du-seo-quels-sites-ont-triomphe-dans-les-resultats-de-google/>

Revue d'URL en anglais...

- How Google May Map A Query To An Entity For Suggestions

A website by the name of SourceFed produced a video that claimed that Google was intentionally manipulating search results to make Hillary Clinton look good, because it wasn't showing results tied to her name that SourceFed insisted Google should be showing...

<http://www.seobythesea.com/2016/06/how-google-may-map-a-query-to-an-entity-for-suggestions/>

- Google: Here's How to Optimize for Google's RankBrain

At Big Digital Adelaide, I had a fireside chat with Gary Illyes from Google, and RankBrain was one of the hot topics of the afternoon. While we have seen many articles published on "RankBrain optimization" and "RankBrain SEO", many are simply based on speculation and assumptions. After all, RankBrain is the newest algo buzzword so SEOs naturally want to do what they can to optimize for it in order to get that edge up. So I asked Illyes if SEOs could optimize for RankBrain and how they could optimize for it...

<http://www.thesempost.com/google-how-to-optimize-for-googles-rankbrain/>

- Do Google's Featured Snippets Prefer Older or Fresh Content?

If you are trying to land a featured snippet, are you better serving up a brand new page to take advantage of freshness? Or are you better to tweak a currently ranked but older page to make it more appealing as a featured snippet option?...

<http://www.thesempost.com/featured-snippets-favor-fresh-older-content/>

- The voice search explosion and how it will change local search

Voice search usage is seeing unprecedented growth, with personal assistant devices leading the way. Columnist Wesley Young explores why this new medium is taking off, how it differs from keyword searches, and the challenges for local businesses to compete on yet another platform...

<http://searchengineland.com/voice-search-explosion-will-change-local-search-251776>

- How to improve your CTR using Google Search Console

Last month I wrote a comprehensive guide on how to use Google Search Console, covering every aspect of what is essentially a giant toolshed full of useful stuff for all webmasters to use...

<https://searchenginewatch.com/2016/06/21/how-to-improve-your-ctr-using-google-search-console/>

- How to speak 'Search Engine'

The challenge of how to 'speak' search engine and tell it how to surface our content is what Search Engine Optimisation is all about. But are we doing it as well as we could?...

<https://searchenginewatch.com/2016/06/24/how-to-speak-search-engine/>

- Machine Learning Inside Google

When I was in high school, one of the required classes I had to take was a shop class. I had been taking mostly what the school called "enriched" courses, or what were mostly academic classes that featured primarily reading, writing, and arithmetic. A shop class had more of a trade focus. I was surprised when the first lesson on the first day of my shop class was a richer academic experience than any of the enriched classes I had taken...

<http://www.seobythesea.com/2016/06/machine-learning-inside-google/>

- How do people view search engine results pages?

The F-shaped pattern has been the commonly understood way in which web users browse sites and search results...

<https://searchenginewatch.com/2016/06/28/how-do-people-view-search-engine-results-pages/>

- Google Displaying Data from Tables in Regular Search Results

We have seen many cases where Google will show tables in featured snippets in the search results. And some types of structured data have the appearance of tables when Google displays it in the search results, such as event listings. But now Google is pulling information from tables to show in the search results...

<http://www.thesempost.com/google-showing-tables-regular-search-results/>

- Understanding TF*IDF: One of Google's Earliest Ranking Factors

*In this week's episode of Marketing Nerds, I am joined by Marcus Tandler, one of the most accomplished SEOs I know and also the co-founder of OnPage.org, a leading Technical SEO Tool. We talk about how Marcus got into SEO, share some funny stories, and then dive into what TF*IDF is, how to use it and why it is so important in SEO today...*

<https://www.searchenginejournal.com/understanding-tfidf-one-googles-earliest-ranking-factors/166856/>

- What's new with markup & structured data

Contributor Eric Enge recaps a session from SMX Advanced on structured data markup in its many forms...

<http://searchengineland.com/whats-new-markup-structured-data-253371>