

Recherche et Référencement

Numéro 144 - Janvier 2013

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



**LE SITE ABONDANCE
VOUS SOUHAITE
UNE BONNE ANNEE 2013 !**



Au sommaire ce mois-ci :

- > [Peut-on s'affranchir des "not provided" ?](#) (13 pages)
- > [Gozil et Askalll, une nouvelle façon communautaire de recherche l'information](#) (9 pages)
- > [WordPress et le Duplicate Content](#) (8 pages)
- > [Le Negative SEO au regard de la loi](#) (4 pages)
- > [EdgeRank et GraphRank : des algorithmes au cœur du fonctionnement de Facebook](#) (7 pages)

- > [Revue d'outils moteurs et SEO](#) (4 pages)
- > [Revue de moteurs](#) (6 pages)
- > [Revue d'URL](#) (6 pages)

Prochaine parution : Vendredi 15 février 2013

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

La lettre "Recherche & Référencement" paraît aux alentours du 15 de chaque mois (un seul numéro pour les mois de juillet-août)

Pour tout renseignement : © Olivier Andrieu, olivier@abondance.com

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

* Il est **interdit d'envoyer, par mail** ou autres systèmes de transmission, un exemplaire de la lettre à quelque format que ce soit, à une personne qui ne serait pas abonné.

* Il est **interdit de diffuser à l'intérieur d'une entreprise, à plusieurs adresses e-mails ou sur un intranet**, le contenu de la lettre "Recherche & Référencement" si vous n'avez pas souscrit un **abonnement "Entreprise"** (<http://lettres.abondance.com/rechref.html>) qui est fait pour cela.

* Il est **interdit de transmettre vos login / mot de passe** à des personnes qui ne seraient pas abonnées. Pour information, les adresses IP des machines se connectant sont relevées et tout abus pourrait être sanctionné si ce phénomène ne pouvait être justifié par vos services techniques.

* Il est **interdit d'utiliser à des fins privées ou publiques tout contenu** de cette lettre sans notre autorisation expresse et écrite : envoi à des clients, des prospects, voire reprise intégrale du contenu en changeant le nom de la lettre et de la source (eh, oui, ça arrive...), etc.

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

Merci pour votre compréhension et bonne lecture.

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

Sommaire du numéro 144 - Janvier 2013

Peut-on s'affranchir des "not provided" ?	3
Gozil et Askalll, une nouvelle façon communautaire de recherche l'information	16
WordPress et le Duplicate Content	25
Le Negative SEO au regard de la loi	33
EdgeRank et GraphRank : des algorithmes au cœur du fonctionnement de Facebook	37
Revue d'outils moteurs et SEO	44
Revue de moteurs	48
Revue d'URL	54

Peut-on s'affranchir des "not provided" ?

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Depuis plus d'un an, Google chiffre une partie des pages de résultats. La conséquence directe est que les mots-clés "referers" ne sont plus transmis aux sites distants, posant d'énormes problèmes marketing pour la mesure de la qualité du trafic fourni par le moteur de recherche. Ces requêtes désormais en (not provided) atteignent, parfois, plus de 50% des mots-clés analysés. Il est donc très important de comprendre quel est le type de trafic impacté par ce phénomène (Courte Traîne ? Longue Traîne ? Nouveaux visiteurs ou anciens ?) et de tenter de trouver des palliatifs pour mesurer de façon efficace votre trafic. Voici comment faire...

Quelle est la situation actuelle ?

La mention (Not provided) a fait son apparition en date du 20 octobre 2011. Elle signale les visites d'internautes qui utilisent une connexion sécurisée (généralement, ceux qui se sont connectés à leur compte Google). Le cryptage des données ne permet donc plus aux outils d'analyse de récupérer les requêtes saisies par les internautes sur le moteur de recherche (mots clés dits "referers").

L'apparition de ces (Not provided) a jeté un certain froid dans le monde du SEO. Une étude récente (<http://www.optify.net/forms/optify-not-provided-study>) montre qu'en septembre 2012, les (Not provided) représentaient 51% des mots-clés détectés.

Voici quelques chiffres-clés à ce sujet :

- Les (Not provided) comptent pour près de 40% du trafic issu de la recherche naturelle ;
- 64% des entreprises constatent qu'entre 30% et 50% de leur trafic est classé en « Not provided » ;
- Le nombre de mots-clés reconnu a chuté de 49%.

Quelles sont les conséquences ?

Les conséquences de ce phénomène sont très importantes. Nous pouvons notamment lister les points suivants :

- Une plus grande difficulté à relier une requête avec les statistiques Google Analytics (taux de conversion, taux de rebond, etc.) ;
- Moins de possibilités de faire des offres adaptées à la requête de l'internaute ;
- Moins d'occasions de faire du « Lead scoring » en fonction de ces mêmes termes de recherche.

Par contre, un certain nombre de leviers restent exploitables :

- Les mesures des performances globales et détaillées sont toujours possibles ;
 - La pertinence du signal statistique obtenu permet de continuer à optimiser son SEO.
- Après tout, un échantillonnage de 50% offre un signal suffisant !

Qu'en-est-il sur votre site ?

Pour le savoir, développez ces rubriques : **Sources de trafic – Sources – Recherche – Résultats naturels**. Voici un exemple de rapport obtenu :

- 275 950 visites ;
- 55 073 requêtes identifiées ;
- 48 346 visites regroupées dans la rubrique **(Not provided)**.

<input type="checkbox"/>	Mot clé	Visites	↓	Pages/visite	Durée moy. de la visite	Nouvelles visites (en %)	Taux de rebond
<input type="checkbox"/>	1. (not provided)	48 346		6,73	00:04:33	67,33 %	42,82 %
<input type="checkbox"/>	2. [blurred]	21 636		7,38	00:03:42	75,69 %	38,50 %
<input type="checkbox"/>	3. [blurred]	16 105		11,53	00:06:55	40,09 %	27,65 %
<input type="checkbox"/>	4. [blurred]	9 699		3,13	00:01:23	88,23 %	62,67 %
<input type="checkbox"/>	5. [blurred]	3 602		12,32	00:07:53	34,09 %	15,55 %
<input type="checkbox"/>	6. [blurred]	2 954		2,89	00:01:17	88,05 %	64,22 %
<input type="checkbox"/>	7. [blurred]	2 851		10,00	00:04:54	64,61 %	36,51 %
<input type="checkbox"/>	8. [blurred]	1 932		5,86	00:03:08	72,52 %	39,75 %

Nous pouvons donc poser cette formule : $(48346 / 275950) * 100 = 17,52\%$. Nous obtiendrons ce même pourcentage en cliquant sur la petite icône représentant un camembert.

Afin d'avoir une idée, sur un an, de la progression des requêtes (Not provided), suivez cette procédure :

1. Accédez au mini-calendrier afin de sélectionner la plage de dates souhaitée.
2. Dans la barre de recherche, saisissez **(not provided)** puis cliquez sur l'icône représentant une loupe.
3. Cliquez sur l'icône **Mois**.



Attention au fait que les données dans Google Analytics sont échantillonnées et que vous devrez peut-être, dans certains cas, définir une plus grande précision des rapports.



Quelles sont les premières questions à se poser ?

Tout le problème est de savoir si ces visites (et ces visiteurs) sont différentes des visites « renseignées ». Une manière simple de procéder consiste à créer deux segments : un premier qui ne va afficher que les visites en « Not provided » et un autre regroupant les requêtes « en clair »...

1. Cliquez sur le bouton **Segments avancés** puis sur **Nouveau segment personnalisé**.
2. Créez ce premier segment (« Not provided ») :
Inclure – Mot clé – Correspondance exacte : (not provided)

The screenshot shows the 'Segments avancés' interface. At the top, there's a dropdown menu for 'MOT CLÉ: (not provided)'. Below it, there are several buttons: 'Segments avancés', 'Personnaliser', 'E-mail', 'Exporter', 'Ajouter au tableau de bord', and 'Raccourci BÉTA'. A text input field for 'Nom' contains 'Not provided'. Below this, there are three main configuration sections:

- 'Inclure' (dropdown), 'Mot clé' (dropdown), and 'Correspondance exacte' (dropdown) with a text input field containing '(not provided)'. A mouse cursor is pointing at the 'Correspondance exacte' dropdown.
- 'ou' (text) with an empty text input field and a button 'Ajouter une instruction "OU"'. A mouse cursor is pointing at the 'ou' text.
- 'et' (text) with an empty text input field and a button 'Ajouter une instruction "ET"'. A mouse cursor is pointing at the 'et' text.

3. Créez ce second segment (« Provided ») :
Exclure – Mot clé – Correspondance exacte : (not provided)
ET
Inclure – Support – Correspondance exacte : organic
ET
Inclure – Mot clé – Correspondance exacte : google

4. Activez les deux segments en ayant soin de sélectionner, dans le mini-calendrier, une plage de dates offrant un signal statistique suffisant.

Une fois cette étape franchie, vous pourrez y voir plus clair sur, au moins, trois points essentiels...

Afin de croiser cette statistique avec les visiteurs nouveaux ou anciens, développez ces rubriques : **Audience – Comportement – Nouveaux vs connus**. Voici des exemples de statistiques :

- Nouveaux visiteurs : Provided : 55301 - Not provided : 19148 (soit un pourcentage de 35%) ;
- Visiteurs connus : Provided : 26819 – Not provided : 9363 (soit un pourcentage de 35%).

En bref, il n'y a aucune différence !

Voyons ce qu'il en est du taux de rebond... Développez les rubriques **Sources de trafic – Sources – Recherche – Résultats naturels**. Voici un exemple des données relevées :
 - Provided : 50,06 % ;
 - Not provided : 44,47 %.

	Visites	Pages/visite	Durée moy. de la visite	Nouvelles visites (en %)	Taux de rebond
Not provided	31 940 % du total: 9,10 % (351 122)	6,40 Moyenne du site: 5,64 (13,41 %)	00:04:16 Moyenne du site: 00:03:28 (22,86 %)	70,69 % Moyenne du site: 67,92 % (4,09 %)	44,47 % Moyenne du site: 46,87 % (-5,12 %)
Provided	165 291 % du total: 47,08 % (351 122)	5,41 Moyenne du site: 5,64 (-4,06 %)	00:03:21 Moyenne du site: 00:03:28 (-3,39 %)	68,36 % Moyenne du site: 67,92 % (0,66 %)	50,06 % Moyenne du site: 46,87 % (6,79 %)

Nous savons maintenant que la différence ne se situe pas au niveau du type de visiteur mais dans le fait que les visites classées en (Not provided) semblent plus adaptées aux contenus des pages du site.

Affinons ce constat les conversions et/ou les transactions enregistrées...

Développez les rubriques **Sources de trafic – Sources – Recherche – Résultats naturels** puis cliquez, par exemple, sur le lien **Commerce électronique**. Voici, pour notre démonstration, les valeurs que nous obtenons en ce qui concerne le taux de conversion du commerce électronique :

- Provided : 0,60 % ;
- Not provided : 1,03 %.

Si nous posons cette formule : $((1,03 - 0,60) * 100) / 0,60$, nous obtenons un lift de 71,67%. Ce n'est pas rien ! Nous pouvons supposer que :

- Les requêtes classées en « Not provided », et saisies par les internautes, « collent » plus aux contenus des pages qui sont proposées ;
- Ces mêmes requêtes dénotent, chez les internautes, un niveau de maturité plus important.

Développez ces rapports : **Contenu – Contenu du site – Pages de destination** puis cliquez, par exemple, sur le lien **Commerce électronique**. Regardons, maintenant, le nombre de pages d'entrée indexées par Google :

- Provided : 2123 ;
- Not provided : 1489.

En bref, le phénomène « Not provided » touche plus particulièrement les pages qui présentent la plus forte valeur par visite et donc affectent les recherches transactionnelles. Mais, pour l'instant, il nous semble impossible de définir si les requêtes (Not provided) affectent les recherches de type courte traîne, longue traîne ou de notoriété.

Utilisez les rapports personnalisés

Il est, dans ce cas, plus simple de créer un rapport personnalisé :

1. Cliquez sur les boutons **Rapports personnalisés** et **Nouveau rapport personnalisé**.
2. Ajoutez ce premier groupe de statistiques : **Visites – Taux de rebond – Pages/visite – Taux de conversion par objectif** ou **Taux de conversion du commerce électronique – Objectifs réalisés** ou **Chiffre d'affaires**.

Contenu du rapport

Engagement x Problèmes + Ajouter un onglet

Nom Engagement

Type Explorateur Tableau statique Synthèse géographique

Groupes de statistiques

Niveau d'interactions

Visites Pages/visite

Taux de rebond Chiffre d'affaires

Taux de conversion du c... + Ajouter une statistique

+ Ajouter un groupe de statistiques

Détails des variables

Page de destination Mot clé

+ Ajouter une variable

Filtres - facultatif

Inclure Type de trafic Regex organic

3. Ajoutez ce second groupe de statistiques : **Visites, Temps passé sur la page, Sorties (en %), Taux de conversion par objectif ou Taux de conversion du commerce électronique, Taux de rebond, Objectifs réalisés ou Chiffre d'affaires.**

Le principe est de créer deux groupes de statistiques avec, d'un côté, les éléments positifs et, de l'autre, les aspects négatifs.

4. Affichez les variables sous forme d'une liste alphabétique.
5. Ajoutez ces variables : **Page de destination – Mot-clé**
6. Appliquez ce filtre : **Inclure – Type de trafic – regex – organic**
7. Enregistrez le segment puis ouvrez ce rapport.
8. Accédez à une des pages de destination qui sont listées.

Ce type de rapport est intéressant pour afficher la diversité des requêtes qui peut mener à une page en particulier. Le nombre de visites indiqué en face de chaque requête permet de lister celles pour lesquelles votre page possède une réelle légitimité.

Mot clé	Visites	Pages/visite	Taux de rebond	Chiffre d'affaires ↓
1. queue de billard	1 722	10,23	33,39 %	3 648,70 €
2. (not provided)	904	13,52	22,90 %	897,65 €
3. billard queue	48	7,44	22,92 %	696,82 €
4. supreme billard	183	10,57	53,01 %	420,24 €
5. queue de billard	206	8,15	44,17 %	362,49 €
6. le meilleur queue de billard	1	18,00	0,00 %	315,85 €
7. queue billard	266	10,05	30,08 %	258,89 €
8. supreme	31	17,23	22,58 %	151,46 €
9. carreaux de billard	192	9,54	34,90 %	142,47 €
10. regles billard	1	184,00	0,00 %	137,41 €

Mais, là encore, nous ne savons pas à quoi s'apparentent les requêtes (Not provided)...

Utilisez les segments avancés

Le principe est d'obtenir :

- Le nombre total des visites issues de la recherche naturelle (courte et longue traîne) ;
- Le nombre total des visites provenant des requêtes de notoriété.

Commençons par créer un segment personnalisé qui va filtrer toutes les visites provenant des requêtes de type « Courte traîne » (soit moins de trois termes).

1. Cliquez sur les boutons **Segments avancés** puis **Nouveau segment personnalisé**.
2. Créez ce segment : **Exclure – Mot-clé – Correspondance avec expression régulière – $^\wedge s^* [^\wedge s] + (\backslash s + [^\wedge s] +) \{2, \} \backslash s^* \$$**
3. Combinez-le avec un segment qui exclue toutes les requêtes de notoriété.
4. Combinez-le avec un segment qui exclue toutes les requêtes en « Not provided ».

À chaque fois, la condition utilisée doit être de type ET.

Pour créer le segment « Longue traîne », utilisez une condition de type « Inclure ».

Nous obtenons quatre segments :

- Requêtes de type « courte traîne » (avec trois sous-segments) ;
- Requêtes de type « longue traîne » (avec trois sous-segments) ;
- Requêtes de notoriété ;
- Requêtes « Not provided ».

5. Activez ces quatre segments.
6. Affichez le rapport que nous venons de créer.
7. Ouvrez une des pages de destination.

Une synthèse présentée sous forme de tableau va être visible.

	Visites	Pages/visite	Taux de rebond	Chiffre d'affaires	Taux de conversion du commerce électronique
Longue traîne	96 314 % du total: 27,43 % (351 127)	4,08 Moyenne du site: 5,64 (-27,62 %)	54,92 % Moyenne du site: 46,85 % (17,23 %)	57 552,85 € % du total: 10,90 % (527 979,20 €)	0,39 % Moyenne du site: 0,67 % (-42,73 %)
Courte traîne	56 431 % du total: 16,07 % (351 127)	5,37 Moyenne du site: 5,64 (-4,63 %)	50,67 % Moyenne du site: 46,85 % (8,15 %)	33 168,09 € % du total: 6,28 % (527 979,20 €)	0,32 % Moyenne du site: 0,67 % (-52,18 %)
Not provided	32 405 % du total: 9,23 % (351 127)	6,45 Moyenne du site: 5,64 (14,45 %)	44,60 % Moyenne du site: 46,85 % (-4,79 %)	52 757,61 € % du total: 9,99 % (527 979,20 €)	0,76 % Moyenne du site: 0,67 % (12,57 %)
Requêtes de notoriété	23 511 % du total: 6,70 % (351 127)	11,09 Moyenne du site: 5,64 (96,70 %)	26,23 % Moyenne du site: 46,85 % (-44,00 %)	139 105,06 € % du total: 26,35 % (527 979,20 €)	1,98 % Moyenne du site: 0,67 % (193,27 %)

Dans ce cas, la corrélation est beaucoup plus simple à faire : les visites en « Not provided » se situent entre celles de type « Longue traîne » et les requêtes de notoriété. Poussons un peu plus loin notre analyse en ajoutant des filtres spécialisés...

Utilisez les filtres avancés

Tout d'abord, expliquons la démarche :

- Nous allons créer un premier filtre qui va extraire le rang du lien cliqué (en incluant les résultats de type « Recherche universelle ») ;
- Un second filtre va nous permettre de juxtaposer le « Ranking » du lien avec le terme de recherche saisi ;
- La création de différents segments va nous permettre de regrouper les requêtes en fonction du « Ranking » des « Snippets ».

Il vaut mieux créer un nouveau profil de site web :

1. Cliquez sur le bouton **Admin**.
2. Cliquez sur le nom du compte puis sur celui du site web.
3. Cliquez sur le bouton **Nouveau profil**.
4. Indiquez un nom de profil de création de rapports.
5. Sélectionnez le fuseau horaire voulu puis cliquez sur le bouton **Créer un profil**.

N'oubliez pas d'accéder aux paramètres du profil afin d'activer, par exemple, les sources de données électroniques et de changer la monnaie de transaction.

6. Cliquez sur le nom du profil que vous venez de créer puis sur l'onglet **Filtre**.
7. Créez deux filtres comme suit :

- **Filtre personnalisé**
- **Avancé**
- **Champ A -> Extrait A – Site référent – (\?|&)(cd)=[^&]***
- **Sortie vers -> Constructeur – Champ personnalisé 1 – \$A3**
- **Champ A requis – Oui**
- **Champ B requis – Non**
- **Remplacer le champ de sortie – Oui**
- **Sensible à la casse – Non**

Nom du filtre	Ranking	
Type de filtre	<input type="radio"/> Filtre prédéfini <input checked="" type="radio"/> Filtre personnalisé	
	<input type="radio"/> Exclure <input type="radio"/> Inclure <input type="radio"/> Minuscules <input type="radio"/> Majuscules <input type="radio"/> Rechercher et remplacer <input checked="" type="radio"/> Avancé	
Champ A -> Extrait A	Site référent	(\? &)(cd)=[^&]*
Champ B -> Extrait B	-	
Sortie vers -> Constructeur	Champ personnalisé 1	\$A3
Champ A requis	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
Champ B requis	<input type="radio"/> Oui <input checked="" type="radio"/> Non	
Remplacer le champ de sortie	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
Sensible à la casse	<input type="radio"/> Oui <input checked="" type="radio"/> Non	

- **Filtre personnalisé**
- **Avancé**
- **Champ A -> Extrait A – Champ personnalisé 1 – (.*)**
- **Champ B -> Extrait B – Terme de la campagne – (.*)**
- **Sortie vers -> Constructeur – Champ personnalisé 1 – \$B1 (\$A1)**
- **Champ A requis – Oui**
- **Champ B requis – Oui**
- **Remplacer le champ de sortie – Oui**
- **Sensible à la casse – Non**

Nom du filtre	Not provided	
Type de filtre	<input type="radio"/> Filtre prédéfini <input checked="" type="radio"/> Filtre personnalisé	
	<input type="radio"/> Exclure	
	<input type="radio"/> Inclure	
	<input type="radio"/> Minuscules	
	<input type="radio"/> Majuscules	
	<input type="radio"/> Rechercher et remplacer	
	<input checked="" type="radio"/> Avancé	
Champ A -> Extrait A	Champ personnalisé 1	(.*)
Champ B -> Extrait B	Terme de la campagne	(.*)
Sortie vers -> Constructeur	Terme de la campagne	\$B1 (\$A1)
Champ A requis	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
Champ B requis	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
Remplacer le champ de sortie	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
Sensible à la casse	<input type="radio"/> Oui <input checked="" type="radio"/> Non	

Patientez quelques heures avant que ce profil soit réellement opérationnel...

Créez ensuite trois segments comme suit :

- **Inclure – Mot-clé – Correspondance exacte – (1)**
- **Inclure – Mot-clé – Correspondance avec expression régulière – \([2-5]\)**
- **Inclure – Mot-clé – Correspondance avec expression régulière – \([6-9]\)**

1. Développez ce rapport : **Sources de trafic – Sources – Recherche – Résultats naturels.**
2. Sélectionnez l'option **Pages de destination** comme variable principale.
3. Cliquez sur une de vos pages de destination.
4. Sélectionnez, comme variable secondaire, la variable **Mot-clé.**

Notez que, là encore, vous pouvez vous servir du rapport personnalisé que nous venons de créer...

5. Activez un ou plusieurs segments à la fois.

<input type="checkbox"/>	Page de destination	Mot clé	Visites	↓ Pages/visite
<input type="checkbox"/>	1. /	(not provided) (1)	32	1,19
<input type="checkbox"/>	2. /	supreme billard (1)	10	1,20
<input type="checkbox"/>	3. /	supreme france (1)	7	1,29
<input type="checkbox"/>	4. /	supreme pool and snooker (1)	4	1,25
<input type="checkbox"/>	5. /	accessories billard (1)	3	1,33
<input type="checkbox"/>	6. /	billard supreme (1)	3	1,33
<input type="checkbox"/>	7. /	supreme 9 (1)	3	1,33
<input type="checkbox"/>	8. /	achat billard (1)	2	1,00
<input type="checkbox"/>	9. /	billard vert main (1)	2	1,00
<input type="checkbox"/>	10. /	vente billard (1)	2	1,00

Nous pouvons nous dire que les requêtes (Not provided), qui possèdent le même rang que les requêtes (en clair) et les mêmes variables comme, par exemple, le taux de rebond, sont identiques.

Notez que vous pouvez aussi utiliser la barre de recherche pour extraire les requêtes de rang 1, 2, etc.

Mais le réel intérêt de ce rapport est autre :

- Il montre les requêtes sur lesquelles votre page présente une réelle légitimité (les mots-clés prioritaires) ;
- Il affiche les nouvelles pages que vous devez créer et, plus tard, le travail de maillage interne que vous devrez mettre en place.

Au final, nous constatons que :

- Si les requêtes ne suscitent pas de conversion et/ou de conversion, il ne sert à rien de se lamenter ;
- Une tentative d'analyse des (Not provided) ne suscite pas de leviers d'action ;
- Un travail fouillé sur les rapports Google Analytics permet de s'en affranchir.

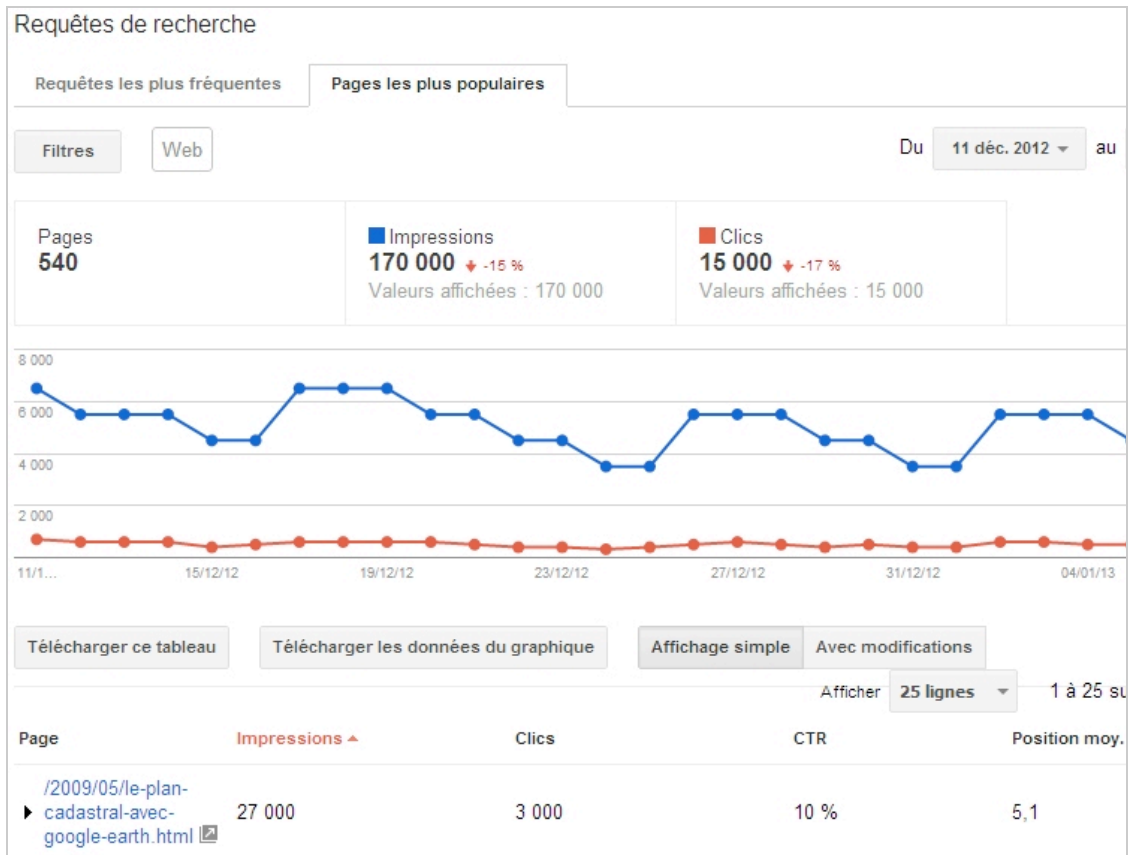
Contre le phénomène (Not provided)

Le fait que la liste des requêtes affichées risque de continuer à se réduire comme peau de chagrin nous oblige à trouver des méthodes alternatives (et à diversifier nos sources de trafic). Le principe est donc de se servir d'autres outils pour compléter les listes de requêtes saisies par les internautes. Si celles fournies par les outils Google pour webmasters est moins étoffée, elle contient des informations qui sont différentes. Rappelons que ce rapport affiche les 1 000 recherches les plus populaires et les 1 000 pages de destination les plus sollicitées au cours des 30 derniers jours.

Notez qu'il est préférable que le compte Google Analytics soit relié aux outils Google pour webmasters...

Le principe est de mettre, côte à côte, les deux rapports pour une même page de destination :

1. Accédez aux outils pour webmasters.
2. Ouvrez un des sites web listés.
3. Cliquez sur **Trafic – Requêtes de recherche**.
4. Cliquez sur l'onglet **Pages les plus populaires**.



5. Développez une des pages visibles.
6. Copiez les données.
7. Collez-le, en texte Unicode, dans une feuille de calcul (Excel par exemple).
8. Supprimez les colonnes inutiles.
9. Copiez dans une colonne le contenu du rapport personnalisé que vous avez créé juste avant et qui correspond à la même page.

Il est plus simple de la copier en (feuille2)...

10. Affichez la feuille1 puis recopiez cette formule :
=RECHERCHEV(A1;Feuil2!A1:A600;1;FAUX)

Les lignes dans lesquelles la mention #N/A apparaissent sont aussi celles qui ne figurent pas dans le rapport Google Analytics.

11. Ajoutez la liste des mots-clés correspondants à celle fournie par Google Analytics.

Il est aussi possible d'utiliser :

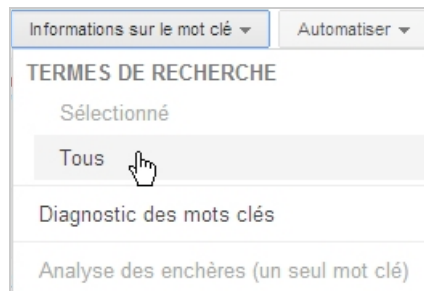
- Les autres outils de veille concurrentielle ou de sourcing (SemRush, par exemple) ;
- Le générateur de mots-clés Google AdWords ;
- Le moteur de recherche interne (sans doute, le plus intéressant !) ;
- Yahoo! Questions/Réponses ou d'autres services spécialisés ;
- Les autres moteurs de recherche (Bing, Yahoo!, etc.) ;
- Les (bons vieux) fichiers de journal de votre hébergeur.

Et sur Google AdWords ?

Le problème est strictement identique sur Google Adwords mais avec des conséquences plus « stratégiques »... Vous pouvez afficher les véritables requêtes saisies par les internautes en développant ces rubriques : **Publicité – AdWords – Requêtes de recherche avec correspondance**. Dans cette liste, la mention (No set) va être visible. Mais, dans la réalité, il existe deux niveaux d'analyse :

- Au niveau des visites et c'est le but du rapport Google Analytics.
- Au niveau des impressions et des clics et c'est le propos du rapport Google AdWords.

1. Affichez les rapports Google AdWords.
2. Cliquez sur l'onglet **Mots-clés**.
3. Cliquez sur le bouton fléché **Informations sur le mot-clé** puis sur l'option **Tous**.



En bas de ce rapport, vous allez voir cette indication : **Autres termes de recherche**.

Autres termes de recherche ?					279	12 041	2,32%	0,15 €	42,38 €	3,6
Total					539	20 189	2,67 %	0,15 €	81,32 €	3,4

4. Cliquez sur la petite icône représentant un point d'interrogation pour afficher les explications que fournit la firme de Mountain View au sujet de ce rapport tronqué :
 - « Vous n'avez pas enregistré de clics pour ce terme de recherche au cours des 30 derniers jours (d'un utilisateur qui n'a pas bloqué son URL de provenance) » ;
 - « Le terme de recherche n'a pas été saisi par un nombre suffisant d'internautes ».

Notez qu'en règle générale, le nombre de clics issus de ces requêtes « mystère » représentent un peu moins de la moitié du nombre total des clics (et des impressions).

- Les problèmes qui se posent sont de trois ordres :
- Il est plus difficile de se rendre compte de la qualité de son travail ;
 - Il est impossible d'évaluer, précisément, la pertinence de l'algorithme de Google AdWords ;
 - Votre taux de clics est affecté sans que vous connaissiez toutes les requêtes concernées.

Il faut préciser que la méthode « Google AdWords » comporte de nombreuses limites :

- Le point de départ reste vos mots-clés de ciblage ;
- *A priori*, plus les groupes d'annonces seront spécialisés et plus les méthodes de ciblage employées seront restrictives, moins vous obtiendrez de renseignements intéressants.

Quelques conclusions temporaires

Dressons un bref rappel des faits :

- Liste des requêtes (variables) qui aboutissent à une page de destination : ce que Google comprend de votre page ;
- Liste des statistiques : la perception de votre contenu par les internautes et, de manière directe, leur niveau de satisfaction (ou d'insatisfaction).

En conséquence, un travail sur le SEO passe par une réflexion sur les objectifs du site... Un objectif peut se traduire par une volonté d'obtenir une parfaite correspondance entre :

- La requête saisie par l'internaute (sous-entendu : son intention) ;
- Le contenu délivré par la page ;

- L'objectif de l'entreprise (en termes de macro-objectifs et de micro-objectifs).

En bref, un mot-clé n'a de valeur que par l'intention qu'il présuppose chez l'internaute... Dresser ou analyser une liste de mots-clés de ciblage n'est qu'une manière de déclarer à Google la thématique de la page. Mais faut-il encore que le contenu coïncide avec les intentions de recherche qui sont sous-entendues par les requêtes saisies par les internautes. C'est une autre manière de dire qu'on ne cible pas des requêtes mais les intentions de recherche qui se cachent derrière ces mêmes requêtes.

C'est dire l'importance toute relative que possède un travail de sourcing puisqu'il revient, plutôt, à effectuer un travail sur les intentions des internautes. C'est donc un préalable indispensable mais qui prendra tout son sens qu'une fois ces étapes complémentaires auront été menées à bien :

- Définition des objectifs de l'entreprise ;
- Liste des intentions de l'internaute ;
- Définition des clusters sémantiques ;
- Création d'un plan de site ;
- Mise en forme des différents contenus.

Il ne sert à rien de vendre quelque chose à partir d'une page qui est bien placée sur une requête « informative ». La légitimité accordée par Google ne coïncide pas avec les intentions de recherche des internautes (sous-entendu, avec leurs objectifs).

A quoi servent les rapports sur les mots-clés ?

La première chose qui frappe est que les mots-clés détectés sont à regrouper par clusters. Et qu'une des premières utilités du rapport « Mots-clés » est de vous permettre de créer des contenus plus spécialisés. Par exemple, si nous constatons que les requêtes "règles Snooker" et "règlement Snooker" aboutissent à la même page de destination, nous devons :

- Modifiez éventuellement le contenu existant ;
- Créez une nouvelle page, plus spécialisée, qui parlera du règlement en vigueur pour cette forme de billard (et non des règles du jeu).

Notez également que plus un mot-clé sera de type « courte traîne », moins il se « gagnera » au niveau de la page mais bien plutôt au niveau du site. En bref, vous devrez acquérir une véritable légitimité pour prétendre être bien placé sur des requêtes extrêmement concurrentielles...

Et maintenant ?

Alors, peut-on s'affranchir des (Not provided) ? Non ! Mais, *a priori*, oui car cela fait longtemps que l'on doit diversifier ses sources d'informations (et, en conséquence, de trafic). Par ailleurs, il ne sert à rien de s'intéresser à des mots-clés qui ne vous apportent pas de trafic qualifié ; c'est-à-dire qui ne suscitent pas des conversions et/ou des transactions. De fait, on peut parfaitement faire du SEO en se passant des (Not provided). Ce n'était donc pas la bonne question mais elle nous a permis de répondre à d'autres interrogations plus essentielles.

Les sources d'inspiration pour l'élaboration de cet article :

Occam's Razzor : <http://www.kaushik.net/avinash/google-secure-search-keyword-data-analysis/> ;

Seomoz : <http://www.seomoz.org/blog/show-keyword-position-using-filters-and-advanced-segments>.

Jean-Noël Anderruthy, *webmaster spécialisé dans les technologies Google.*

Gozil et Askalll, une nouvelle façon communautaire de recherche l'information

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Il n'y a pas que Google dans la vie pour recherche l'information sur le Web. Les systèmes de questions/réponses, notamment, peuvent être intéressants à utiliser dans de nombreux cas. Des sites comme Quora, déjà décrit dans cette lettre, ont amené de nombreuses nouveautés déjà dans ce domaine. Deux nouveaux outils développés par des sociétés françaises, Gozil et Askalll, s'aventurent à leur tour dans cette voie avec des systèmes intéressants et bien conçus, tenant du système de questions/réponses, du sondage en ligne et du réseau social. Exploration...

Nous évoquons très régulièrement dans ces pages le « social search » qui, rappelons-le, désigne selon les chercheurs américains Ed Chi et Brynn Evans, un mode de recherche sur le web qui :

- utilise les réseaux sociaux et réseaux d'experts et
- est mené dans des espaces de travail partagés ou
- met en jeu des techniques de data mining social ou
- met en jeu des processus d'intelligence collective pour améliorer la recherche.

Les mêmes définissent trois stratégies-types de recherche sociale :

- **Collective social search** : résultats de recherche dans les moteurs améliorés par des éléments (tendances, mots-clés, hashtags) tirés de services de type réseaux sociaux.
- **Friend-filtered social search** : recherche dans des données et résultats déjà trouvés/partagés/« likés » par les membres de son réseau (contacts/amis).
- **Collaborative social search** : quand deux personnes ou plus s'associent pour trouver la réponse à une question. C'est le cas de tous les services de type questions-réponses comme Yahoo ! Answers et Quora, qui a largement renouvelé le genre (cf. Lettre R&R n° 117 – Juillet-Août 2010).

Les services que nous évoquerons dans cet article font partie de cette dernière famille et ont pour point commun d'avoir été mis en ligne récemment par des sociétés françaises.

Gozil (<http://www.gozil.com/>)

Gozil est un service qui a été lancé par deux développeurs en novembre dernier. D'après ses fondateurs, ce mot est la contraction du verbe anglais to Go et de « zil » qui fait référence au mot « zillion » et désigne donc les grands nombres. Gozil signifie alors « le départ vers un nombre infini de connaissance ».

L'inscription au service se fait par un simple formulaire et on ne peut pas utiliser à cet effet ses identifiants Google, Facebook et autres (est-ce vraiment un mal ?).

L'interface de la page d'accueil est très épurée et l'on comprend assez vite que la barre de recherche est l'outil principal qui permet de rentrer dans le jeu.

Gozil

Intelligen|

G crid ▾

Pas

La théorie de Darwin est-t-elle valable? Question

Quel est l'objectif de la Business Intelligence ? Question

Pour enrichir votre flux d'activités, suivez des sujets, des questions ou des membres.

Sans réponse

Poser une question

En effet, dès que l'on commence à taper un mot-clé apparaissent des suggestions de questions de membres l'ayant déjà utilisé soit en « clair » dans leur question, soit dans les éléments qui la précisent :



La théorie de Darwin est-t-elle valable?

Quels sont les arguments en la défaveur de la théorie de Darwin?

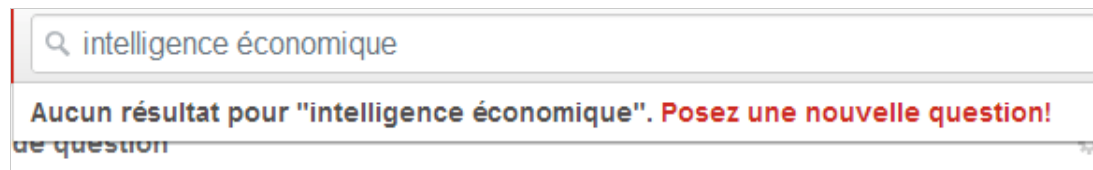
Quels sont les arguments en faveur de la théorie de Darwin?

La vie peut-t-elle être le fruit d'un créateur **intelligent?**

théorie darwin créateur

marlene a posté la question Oct 30, 2012 dans la catégorie Théologie

Lorsqu'aucune question n'existe autour de votre mot-clé, le système vous propose d'en créer une.



intelligence économique

Aucun résultat pour "intelligence économique". Posez une nouvelle question!

Un formulaire apparaît alors qui vous permet de rédiger votre question mais aussi :

- de l'intégrer dans une catégorie ;
- de la préciser ;
- de lui ajouter des tags ;
- de choisir si vous voulez recevoir une notification par email lorsqu'une réponse est apportée par un utilisateur du service.

Poser une question:

Titre (une seule phrase synthétique et distinctive):







Connaissez-vous de bons ouvrages traitant de cartographie automatique de l'information ?

Avant de poursuivre, merci de vérifier que votre question n'a pas déjà été posée:

Connaissez vous 3 habitudes indispensables pour être zen au quotidien?
Quels sujets de conférence sur le web design aimeriez-vous voir pour des évènements comme SudWeb et ParisWeb?
Quelle est la position de l'Islam par rapport à l'adoption ?
Quels remèdes contre la gueule de bois ?
Le chocolat provoque-il des crises de migraine ?

Catégorie:

Question:

B *I* U **Abc**       **</>**

Je parle ici de représentation cartographique n'incluant pas de travail "manuel" de type mindmapping ou concept mapping

Tags - utilisez la virgule (,) comme séparateur:

information, cartographie de l'information

Exemples de tags: [voyages zen](#) [voyage](#) [habitudes zen à avoir](#) [santé](#) [cancer](#)

Me prévenir par email (christophe.de@gmail.com) si ma question a une réponse ou est commentée

Le moteur permet bien sûr d'identifier des items potentiellement intéressants pour vous qui peuvent être de deux types : il peut s'agir de questions, comme nous l'avons vu, mais également de sujets, c'est-à-dire de thèmes issus des catégories dans lesquelles les questions sont classées.

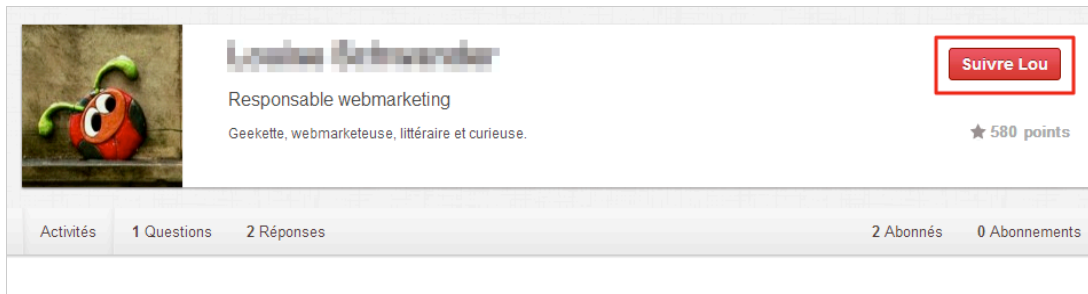
The screenshot shows the Gozil search engine interface. At the top left is the 'Gozil' logo. A search bar contains the text 'référé'. To the right of the search bar is a 'G' logo and a dropdown menu labeled 'crid'. Below the search bar is a list of search results. The first result is 'A quoi sert le référencement' with a 'Question' label. The second is 'Qu'en pensez vous du logiciel de référencement " SEO TOOLKIT" de la société TREILLAN ?' with a 'Question' label. The third is 'Pourquoi Google menace de supprimer de son index les sites de presse?' with a 'Question' label. The fourth is 'Quel est l'intérêt d'avoir des titres (attribue title) dans les liens internes HTML?' with a 'Question' label. The fifth is 'Le Trustrank est-il toujours d'actualité en 2012 ?' with a 'Question' label. The sixth is 'Référencement' with a 'Sujet' label. At the bottom right of the results list, the name 'Fabrice Fourdrinier' is visible.

Chaque page consacrée à une question fonctionne comme un forum de discussion où il est bien sûr possible d'apporter sa propre réponse, mais aussi de commenter une réponse précédente ou d'associer une nouvelle question à celle-ci.

The screenshot shows a question page on Gozil. The question is titled 'Le Trustrank est-il toujours d'actualité en 2012 ?'. The question text is: 'Le TrustRank, signifiant notion de confiance, est une étude fondée par des chercheurs de l'université Stanford et de Yahoo à la base. Google l'utilise pour combattre les spams et les copier/coller. Que représente cet algorithme dans les classements des résultats?'. Below the question text are tags: 'trustrank', 'seo', 'référencement', and 'duplicate'. Below the tags is the text: 'david a posté la question Nov 9, 2012 dans la catégorie Internet'. To the right of the question text is a 'Signaler' button and a 'Répondre' button. Below the question text is a section titled '1 Réponse'. The response is by 'valbou' and is dated 'Nov 15, 2012'. The response text is: 'Bonjour, Le TrustRank est attribué à un domaine et donc à toutes les pages de ce domaine. Le TrustRank a un impact direct sur le positionnement dans les résultats de recherche, il servira a trancher lorsque deux sources fournissent le même contenu. La source ayant le meilleur TrustRank sera dans l'index principal et la copie sera mis en sandbox dans un index secondaire. Si plusieurs sites partagent un TrustRank important alors le contenu dupliqué peut-être toléré par Google. Dans certains cas un bon TrustRank permettra d'obtenir plusieurs positions différentes pour une même requête (surtout pour des requête peu concurrentielles ou précises), allant parfois jusqu'à plusieurs pages du même nom de domaine sur la première page de résultats. Enfin le TrustRank ne permettra pas de gagner des positions sur une requête large. D'autres facteurs sont à prendre en compte avant de parler de TrustRank, comme la bonne indexation du contenu du site avec une sémantique propre. (Un petit message pour Gozil qui ne met pas de titre sur ses liens internes notamment ;)'. To the right of the response text is a 'Signaler' button, an 'Associer une question' button, and a 'Commenter' button. On the right side of the page, there is a 'Suivre la question' button, a 'Flux d'activités' section with 'Questions', 'Sans réponse', and 'Poser une question' options, a 'Questions associées:' section with three related questions, a 'Partagez la question' section with 'Twitter' and 'Facebook' options, and a 'Statistiques' section with the text: 'Cette question a été vue 10 fois' and '0 personnes suivent cette question'.

Par ailleurs, chaque question peut être suivie (principe de l'abonnement) et viendra alors alimenter le flux de votre page d'accueil. Enfin, comme on le voit sur la partie gauche de la page, il est également possible de voter pour une réponse. Le système se charge alors de faire émerger les meilleures réponses pour chaque question.

En cliquant sur le pseudo d'un utilisateur, on arrive sur une page profil qui synthétise ses contributions au dispositif : questions posées et réponses données. Si les thématiques qu'il aborde et les réponses qu'il donne vous intéressent il est possible de s'abonner à son profil afin de ne rien rater de ses interventions.



The image shows a profile card for a user named Lou. On the left is a profile picture of a red, round character with green appendages. To the right of the picture, the name 'Lou' is displayed in a stylized font, followed by the title 'Responsable webmarketing' and the bio 'Geekette, webmarketeuse, littéraire et curieuse.' A red button with the text 'Suivre Lou' is positioned in the top right corner. Below the bio, it shows '★ 580 points'. At the bottom of the card, there are statistics: 'Activités', '1 Questions', '2 Réponses', '2 Abonnés', and '0 Abonnements'.

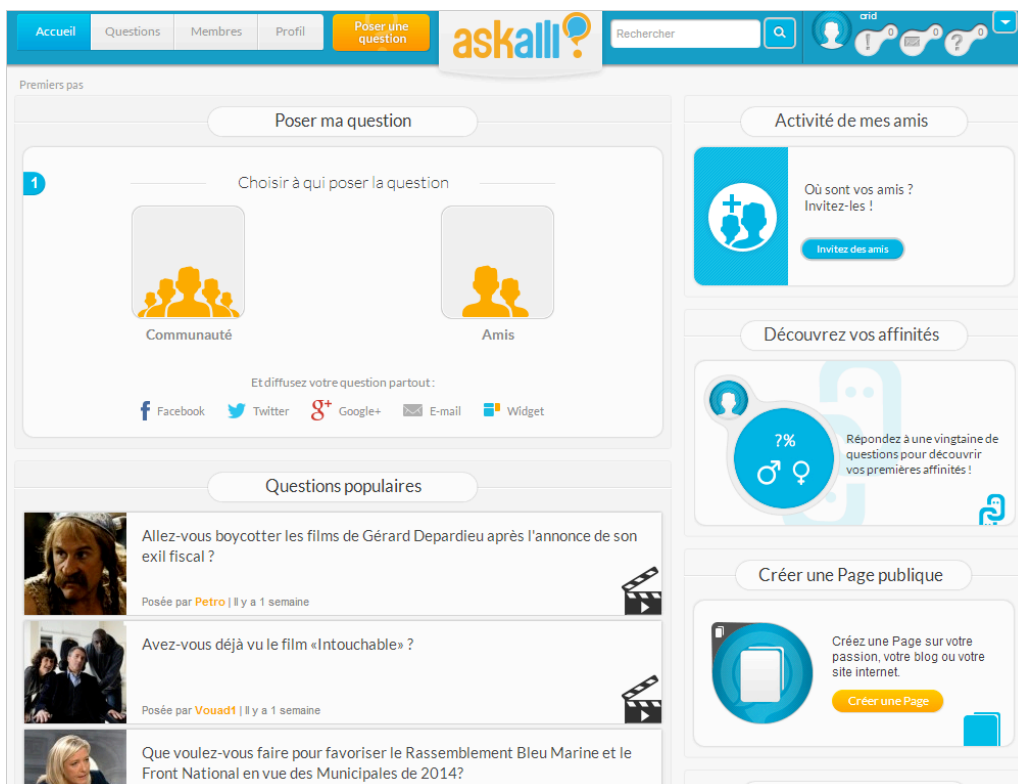
Notez qu'il est aussi possible de lui envoyer des messages directement si l'a permis dans l'interface de configuration de son compte.

Au final un service simple à comprendre et utiliser et (donc) bien conçu. Nous avons relevé quelques points négatifs cependant. On regrettera par exemple que le système ne nous propose pas des personnes à suivre en se fondant sur les thématiques pour lesquelles nous avons déclaré un intérêt (questions et personnes suivies). Et bien sûr on ne peut que constater qu'il s'agit d'une copie assez fidèle des mécanismes inventés par Quora. Gozil dispose cependant selon nous de deux chances de se démarquer de ce service. D'une part il s'adresse à un public francophone là où Quora continue à n'accepter que les contributions en anglais. D'autre part l'interface et les fonctionnalités de ce dernier se sont complexifiées avec le temps et le rendent parfois un peu abscons, alors bon vent à Gozil !

Askall (<http://www.askall.com/>)

Askall est un service qui a été lancé en décembre dernier par une équipe d'une vingtaine de personnes (et oui, il y a bien 3 « l » à la fin du nom). Il est possible de s'y inscrire en créant un compte ou en utilisant ses identifiants Facebook ou Twitter.

La page d'accueil du service est riche et le plus simple est de commencer par le commencement, à savoir poser une question.



The screenshot shows the homepage of the Askall website. The top navigation bar includes 'Accueil', 'Questions', 'Membres', 'Profil', and a prominent orange button 'Poser une question'. A search bar and user profile icons are also present. The main content area is divided into several sections: 'Premiers pas' with a 'Poser ma question' section where users can choose to ask a question to the 'Communauté' or 'Amis'; 'Activité de mes amis' with an invitation prompt; 'Découvrez vos affinités' showing a 7% statistic; and 'Créer une Page publique'. A 'Questions populaires' section features three questions with associated images and user information, such as 'Allez-vous boycotter les films de Gérard Depardieu après l'annonce de son exil fiscal?'.

Le premier choix à effectuer est de savoir si vous l'adressez à la communauté ou uniquement à vos amis. Débutant sur Askalll et donc sans « amis » pour l'instant, nous ferons le premier choix. Un second choix arrive très vite puisque vous devez alors choisir le type de question que vous souhaitez poser et ceci en fonction du type de résultats que vous souhaitez obtenir, à savoir :

Sous forme de graphique :

- Choix entre deux options (VS) ;
- Réponse unique (choix 1, choix 2 ou choix 3) ;
- Réponse chiffrée (ex : combien de fois allez-vous au cinéma par semaine) ;
- Notation (sur une échelle de 1 à 10).

Sous forme de classements :

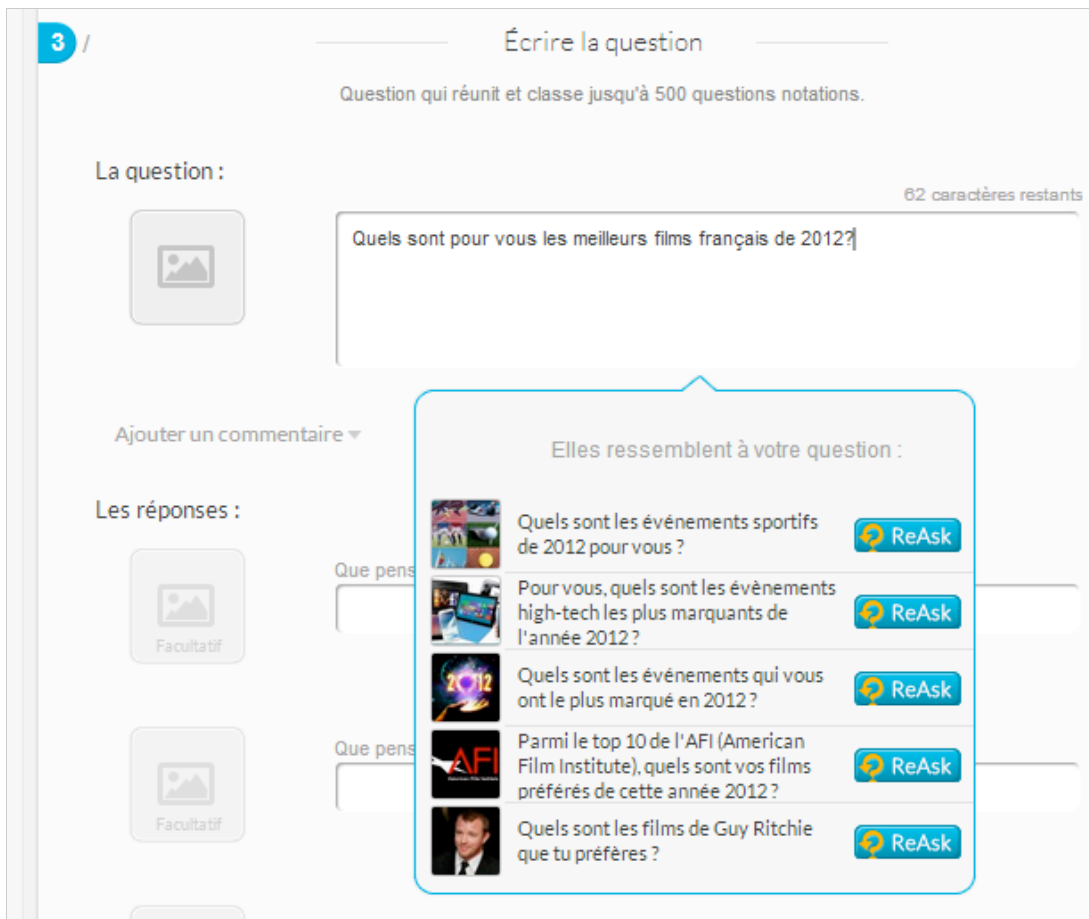
- Notation comparée (ex : quelle est la meilleure actrice française ? choix 1, choix 2 ou choix 3) ;
- Réponses multiples : (ex : quels sont vos sports favoris ? Football et/ou rugby et/ou Tennis).



Une fois ce choix effectué, ce qui n'est pas forcément simple, la troisième étape consiste à rédiger sa question ainsi que les éventuelles réponses possibles.

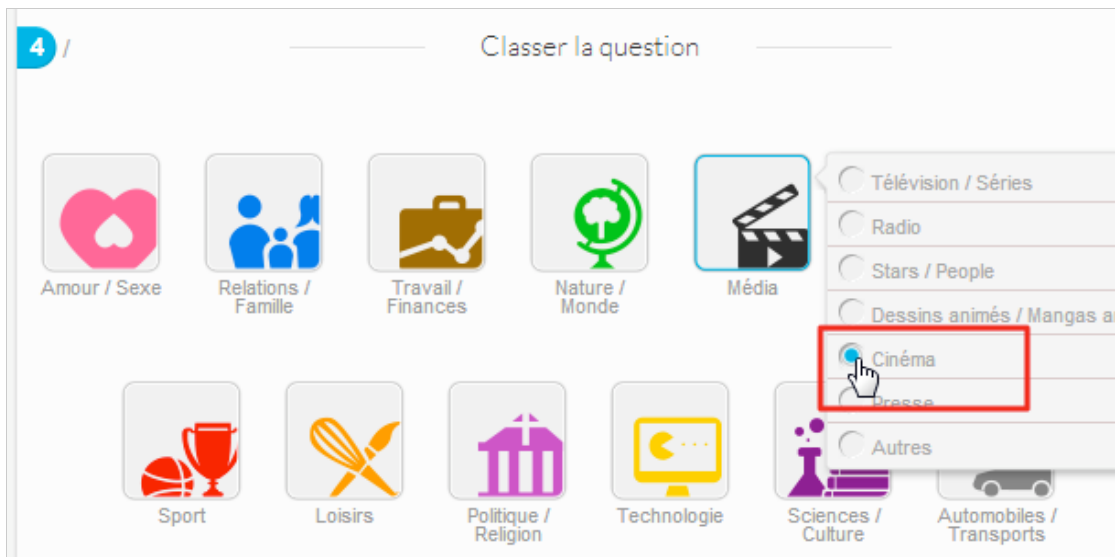
De fait on comprend ici qu'Askalll est beaucoup plus un service de création de sondages qu'un service de questions-réponses traditionnel. Après avoir bien cherché nous avons en effet constaté qu'il n'était pas possible de poser des questions ouvertes du type : quels sont selon vous les meilleurs films français de 2012 ?

Continuons donc avec cette première question.



On remarque que le système émet des suggestions automatiques en fonction de la question posée afin d'éviter les doublons.

La quatrième étape consiste à classer la question dans une catégorie existante.



Il est également possible d'ajouter une date de clôture à cette question et de choisir de la poser anonymement. Par ailleurs le service vous propose d'ajouter une image à celle-ci, à choisir parmi un ensemble proposé suite à une requête automatique.

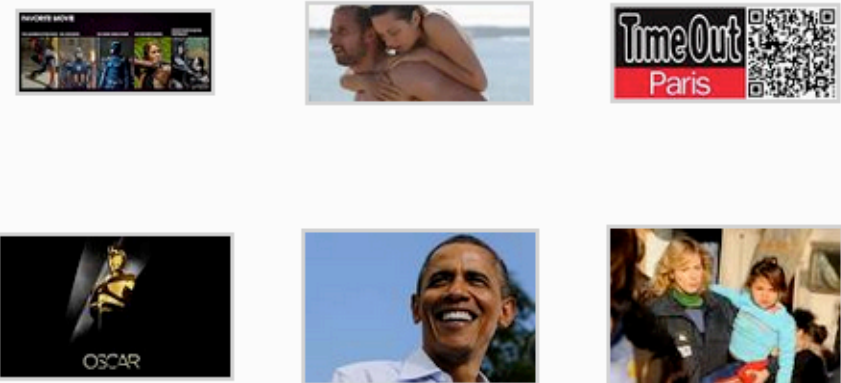
Joindre une image

Choisir une image :

De mon ordinateur

Depuis une adresse web (URL)

Rechercher une image



La question peut alors être lancée.



Posée par **crld** | Il y a moins d'une minute | 

Quels sont pour vous les meilleurs films français de 2012?

★



Répondre

A coeur ouvert

?/10

0 1 2 3 4 5 6 7 8 9 10

Être anonyme

Le prénom

?/10

0 1 2 3 4 5 6 7 8 9 10

Être anonyme

Une fois de retour en page d'accueil, les possibilités d'interaction sont très nombreuses. On peut bien sûr répondre à des questions posées à la communauté et visualiser immédiatement les résultats :



Comme on le voit, Askalll va plus loin que les services de sondage classiques en proposant également de visualiser des données sur les répondants.

Toujours à partir de la page d'accueil, il est possible d'inviter des amis en connectant son compte Gmail et de suivre ensuite leur activité sur le service. On peut également répondre à une vingtaine de questions rapides afin qu'Askalll effectue un calcul d'affinités et vous propose des profils d'utilisateurs à suivre. En effet, à l'instar de Quora ou de Gozil, il est possible de s'abonner aux activités d'une personne. La page profil de chacun, plutôt bien pensée, se présente ainsi :

Petro

Petrovski | Membre depuis le 23/07/2012

Je pense we are...

Follow me and I will follow you !

Voir les photos

Thèmes préférés

Ses questions posées

Toutes Toutes Récentes

Quelle note donneriez-vous à ce redesign de Facebook?

Posée par Petro | Il y a 1 heure

Utilises-tu Foursquare?

Posée par Petro | Il y a 20 heures

Allez-vous suivre le Dakar 2013?

Ses Pages

Globetrotter

Le Globe Trotteur

Ses informations

Nom: Petrovski

50 4532 40

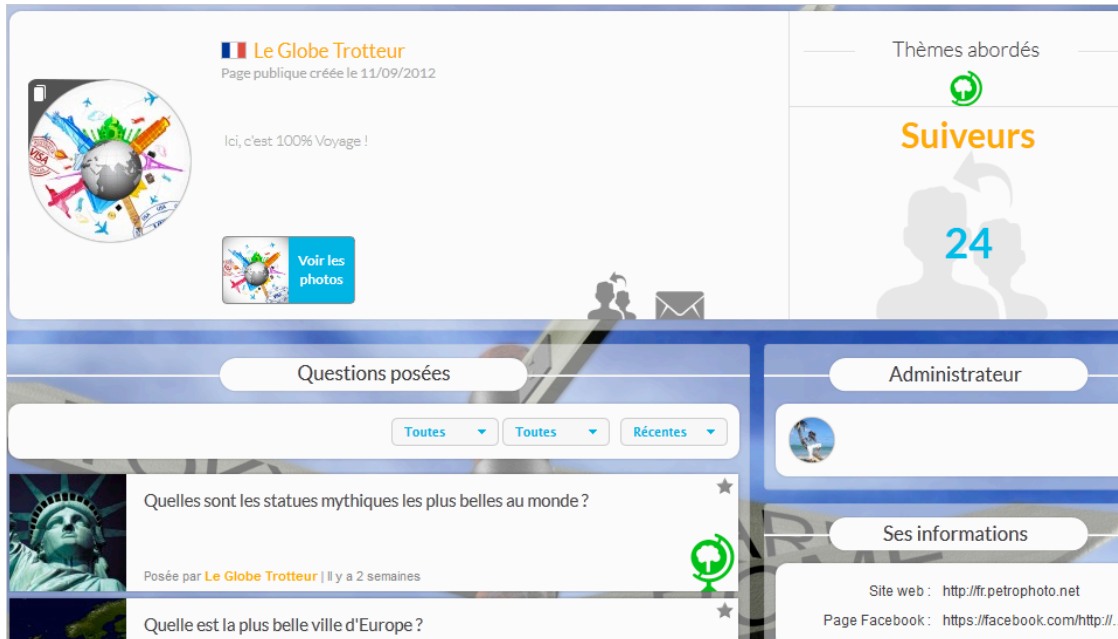
Questions Réponses Suiveurs

On y trouve :

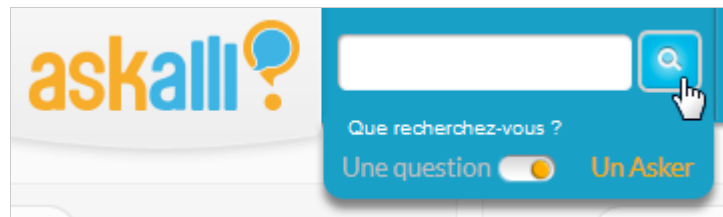
- Des informations de profil classiques (pseudo, devise, photos) ;
- La possibilité de le suivre, l'ajouter comme ami ou lui envoyer un message direct ;

- Un schéma qui synthétise ses centres d'intérêt ;
- Les questions qu'il a posées ;
- Des indicateurs de son activité ;
- Ses pages thématiques (cf. ci-dessous) ;
- Le degré d'affinité que l'on partage avec lui calculé par le système ;
- Les questions auxquelles il a déjà répondu ;
- La liste des utilisateurs qui le suivent (suiveurs) ;
- La liste des utilisateurs qu'il suit (*askers*).

A l'instar des services de curation, Askall permet de créer des pages publiques ou privées dans lesquelles on regroupera des questions thématiques. Ces pages peuvent être liées à des blogs personnels ou des profils de médias sociaux (Twitter et Facebook) et s'intégrer ainsi à une démarche de gestion de son identité numérique.



Last but not least, le moteur de recherche visible en haut de chaque page permet de rechercher des questions ou des utilisateurs « askers ».



On l'aura compris, Askall est un réseau social de mini-sondages déjà très abouti et qui gagne à être connu tant les fonctionnalités qu'il offre sont nombreuses, innovantes et bien réalisées. On regrettera cependant que les questions ouvertes ne soient pas permises (pour l'instant).

Quoiqu'à des degrés de maturité et sur des positionnements différents, Gozil et Askall sont des services que nous trouvons tous deux bien conçus et potentiellement utiles. Comme d'habitude avec les services fonctionnant en mode réseau social, le succès viendra d'abord du nombre d'utilisateurs qu'ils réussiront à capter et fidéliser, seul gage d'un crowdsourcing pertinent, et à ce jeu là Askall nous semble le mieux parti.

Christophe Deschamps

Consultant et formateur en gestion de l'information.

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

WordPress et le Duplicate Content

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

WordPress est un CMS prisé par les référenceurs et les utilisateurs pour sa simplicité d'utilisation et la facilité à l'optimiser pour le référencement naturel. Malheureusement, celui-ci a une lacune importante : le contenu dupliqué. Voici comment le limiter ou même totalement l'éradiquer sur votre site en configurant au mieux votre plate-forme...

Dans cet article, nous allons voir comment réduire au minimum le phénomène de duplicate content, à la fois pour l'internaute et le moteur de recherche.

PS : nous partirons du principe que vous utilisez le plugin WordPress SEO, l'un des meilleurs et des plus complets plugins de référencement naturel pour WordPress : <http://wordpress.org/extend/plugins/wordpress-seo/>

C'est quoi le duplicate content ?

Avant tout chose, il est bon de rappeler quelques bases. Le contenu dupliqué est le fait de trouver un contenu identique sur plusieurs adresses web différentes. Celui-ci peut-être :

- interne :
 - `pageA.html`
 - `pageB.html`
- externe :
 - `mon-site.com/pageA.html`
 - `site-concurrent.com/pageB.html`
- partiel ou complet (le contenu a pu être dupliqué intégralement ou partiellement)

En externe

Il faut bien différencier le contenu dupliqué en interne de celui qui ne l'est pas. Situé sur un autre nom de domaine que le vôtre, Google doit déterminer quelle est la copie originale. Pour cela, il peut utiliser les dates de publication, la popularité et la "crédibilité" de chaque site. Malgré les différents signaux que peut utiliser le moteur de recherche, il est fréquent pour lui de se tromper.

Si le moteur de recherche estime que votre site est celui qui copie, il le pénalisera. En d'autres termes, le contenu dupliqué externe est le plus nuisible d'entre tous.

C'est aussi le plus dur à combattre : un simple copier/coller - plus ou moins automatisé - d'une personne peu scrupuleuse permet de reproduire facilement toutes vos publications.

En interne

La duplication interne est moins dangereuse car elle ne provoque pas de pénalité directe sur votre référencement naturel. Celle-ci est plus insidieuse. Un contenu identique sur plusieurs pages de votre site va en effet avoir plusieurs effets indésirables :

- Cela ralentit l'indexation des contenus : pourquoi continuer à indexer si vos contenus sont identiques ?
- Cela force Google à choisir quelle page positionner parmi les différentes URL disponibles (et il ne choisit pas forcément la plus pertinente).
- Cela dilue la popularité du site sur des pages similaires et inutiles.

Vous n'aurez jamais de pénalité en tant que tel pour une duplication de contenu interne, mais cela peut quand même nuire à votre référencement.

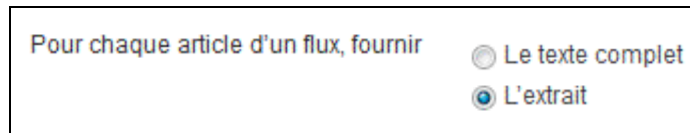
WordPress duplique le contenu en externe

Comme expliqué précédemment, il est difficile de lutter contre la duplication de contenu sur d'autres sites. Rien ne vous protégera entièrement du simple copier/coller.

Cependant, une forte partie du contenu dupliqué en externe est réalisé avec des logiciels automatisés, qui vont scanner votre site et vos articles pour les récupérer. Et le premier point d'entrée de ces robots est votre flux RSS.

Par défaut, WordPress affiche l'intégralité de chaque article dans votre flux d'actualité, à l'adresse suivante : <monsite.com/feed/>

En changeant une simple option, le CMS n'affichera plus qu'un extrait de contenu. Cela réduira considérablement les risques de contenus dupliqués, tout en incitant vos lecteurs à se rendre sur votre site plutôt qu'à utiliser un lecteur de flux RSS. Rendez-vous dans le menu "**Réglages**" => "**Lecture**", puis cocher l'extrait pour l'option "*Pour chaque article d'un flux, fournir*".



La deuxième méthode consiste à installer un plugin appelé **Read More Copy Link** : <http://wordpress.org/extend/plugins/read-more-copy-link/screenshots/>

Celui-ci va permettre :

- d'ajouter un lien "*Lire la suite*" dans tous les texte copiés avec Ctrl-C.
- de réduire la quantité de texte copié (en fonction du paramétrage que vous souhaitez).

Le contenu dupliqué en Interne

Avant d'entrer dans le détail de chaque type de page, il faut savoir une chose sur WordPress : sans paramétrage spécifique et avec un thème par défaut, un article publié sera dupliqué à plusieurs niveaux. On retrouvera ainsi l'intégralité du contenu sur :

- La page de l'article lui-même.
- La page d'accueil.
- Les pages de catégorie.
- Les pages de mots clés.
- La page d'auteur.
- La page d'archive par date.
- Les pages de Custom Taxonomies (nous en parlerons en fin d'article).
- Les sous-pages de chacune des pages citées précédemment.

Il faut donc trouver un moyen de présenter soit un contenu différent, soit un contenu tronqué sur chacune de ces pages.

Les articles et l'accueil du site

La première étape cruciale est de ne pas dupliquer l'intégralité du contenu de chaque article. Pour faire cela, il faut utiliser une simple option disponible par défaut lors de la rédaction d'un article : le bouton "`<!--more-->`". Celui-ci se situe parmi les autres boutons de mise en page :

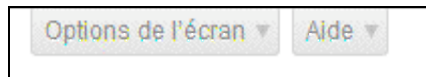


Quand vous l'insérez dans une publication, tout le contenu situé avant la balise sera visible comme avant sur toutes les pages du site, mais celui situé après ne sera visible que dans l'article lui-même. Les autres pages du site vont donc afficher un lien "*Lire la suite*" pour accéder à l'intégralité du contenu. On ne duplique donc plus que le début de chaque article et non plus son intégralité.

Les catégories

Une fois ce premier problème réglé, nous avons toujours un contenu identique sur l'accueil et les autres types de pages. Nous allons donc nous attaquer tout d'abord aux catégories.

Là aussi, on peut réussir facilement à éradiquer le contenu dupliqué. Pour cela, il faut commencer par rédiger une description courte de chacun de vos articles grâce à ce qu'on appelle l'**excerpt** (traduit en français par "*Extrait*"). Par défaut, ce bloc est caché lors de la rédaction d'un contenu. Pour l'afficher, un menu d'option se trouve en haut à droite.



Une fois cette description renseignée, il faudra modifier certains fichiers de votre thème, situé dans le répertoire `monsite.com/wp-content/themes/nom-du-theme/`

Il faut en effet remplacer toute référence à `the_content()` par `the_excerpt()`. En fonction de votre thème, il est possible de trouver les références incriminées dans les fichiers suivants :

- `category-nom.php`
- `category-id.php`
- `category.php`
- `archive.php`, si aucune présence des fichiers `category.php`
- `index.php`, si aucune présence des fichiers `category.php` et `archive.php`. Faites alors attention à placer ce code avec des conditions du type :

```
if ( is_category() ) the_excerpt();
```

Attention, `the_content()` peut parfois avoir des paramètres situés entre les parenthèses mais il est inutile de les recopier dans `the_excerpt()`.

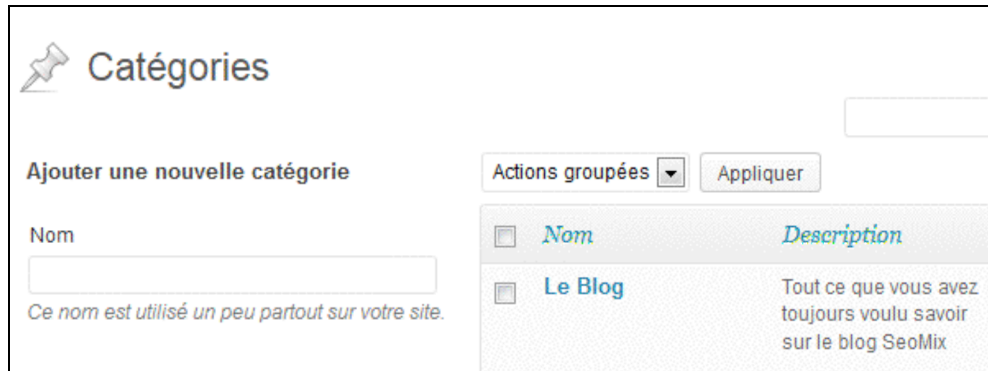
Malheureusement, c'est encore insuffisant pour avoir des catégories dignes de ce nom, car les pages de catégories possèdent encore d'autres lacunes.

La première d'entre elle est que WordPress va tester la dernière partie de l'URL de la catégorie pour savoir si elle existe ou non. Ainsi, si vous tapez `monsite.com/category/nom-de-la-categorie-36/` et `monsite.com/category/nom-de-la-categorie-36/nom-de-la-categorie-36/`, les deux URL vont renvoyer le même contenu. Pour éviter cela, il existe plusieurs possibilités :

- supprimer le `/category/` par défaut des URL de catégorie. Pour faire cela, il existe deux méthodes différentes :
 - une option dans le plugin **WordPress SEO** via le menu **SEO => Permalien**
 - un plugin qui ne fait que cela : **WP No Category Base** (<http://wordpress.org/extend/plugins/wp-no-category-base/>)
- utiliser une solution donnée par *Julio Potier* de **BoiteAWeb**, qui consiste à ajouter un code dans le fichier `functions.php` de votre thème WordPress : <http://www.geekpress.fr/wordpress/astuce/erreur-404-pagination-1422/>

La seconde lacune est d'avoir le même contenu dans toutes les catégories associées à votre article. Si une publication est associée à la catégorie A et B, le contenu sera identique sur les deux pages de catégorie. Le mieux et le plus simple pour éviter cela est de bien concevoir la structure de votre site et de vos contenus. Chaque catégorie doit être logique et unique, le but étant de n'avoir qu'une seule catégorie associée à chaque article.

La troisième lacune est qu'une catégorie devient alors une succession d'extraits d'articles, sans avoir de réel contenu unique et pertinent supplémentaire. On peut agir de plusieurs façon sur cela : on peut tout d'abord remplir la description de la catégorie dans l'administration de WordPress.



Il suffit ensuite de l'afficher avec le code suivant, toujours dans les mêmes fichiers que précédemment.

```
<?php echo category_description(); ?>
```

On peut également opter pour le fait d'avoir des affichages complètement différents par catégorie. Comme on l'a vu dans la liste des fichiers, on peut associer une mise en page et des contenus en fonction d'un ID ou d'un nom de catégorie. Par exemple, je peux très bien avoir :

- `category-36.php`, un fichier de template unique pour la catégorie 36.
- `category-38.php`, un autre template pour la catégorie 38.
- ...
- `category.php`, pour l'affichage par défaut de toutes les autres catégories.

Libre à vous de modifier chaque template en fonction de vos besoins et objectifs.

Les tags (ou mots clés)

Les étapes suivantes consistent à améliorer les pages de mots clés. Là aussi, nous allons tout d'abord chercher à ne pas avoir le même contenu. Pour gagner du temps, il est recommandé d'appliquer la même méthodologie que pour les catégories, à savoir :

- Remplissez la description de vos mots clés dans l'administration de WordPress, puis affichez-les avec le code `<?php echo tag_description(); ?>`
- Utiliser le contenu tronqué avec l'aide de `the_excerpt()` au lieu de `the_content()`

Toutes ces modifications doivent être réalisées dans les fichiers suivants de template de votre thème :

- `tag-nom.php`
- `tag-id.php`
- `tag.php`
- `archive.php`, si aucune présence des fichiers `tag.php`
- `index.php`, si aucune présence des fichiers `tag.php` et `archive.php`. De même, faites attention à placer votre code avec des conditions du type : `if (is_tag()) the_excerpt();`

On peut éventuellement aller plus loin si l'on veut vraiment n'avoir aucun texte dupliqué. Rappelez-vous que `the_excerpt()` reprend l'extrait de votre article, mais le contenu sera donc identique sur la page de catégorie et celles des mots clés. Là aussi, PHP nous donne la possibilité de prendre le contenu intégral d'un article en le tronquant après un nombre défini de caractères (260 dans l'exemples ci-dessous). Voici tout d'abord le code à placer dans le fichier `functions.php` de votre thème :

```
function seomix_description() {
    global $post;
    $contenu = strip_tags($post->post_content);
    $contenu = strip_shortcodes($contenu);
    $contenu = str_replace(array("\n", "\r", "\t"), ' ',
$contenu);
    $contenu = '<p>'.substr($contenu, 0, 260).'...</p>';
    echo $contenu;
}
```

Placer ensuite le code suivant en remplacement de `the_content()` ou de `the_excerpt()` :

```
<?php if (function_exists('seomix_description')) {
seomix_description();}?>
```

Mais attention, même en faisant les étapes précédentes les mots clés peuvent se révéler être un véritable fléau pour le contenu dupliqué, pour la simple et bonne raison que bien souvent les utilisateurs "adorent" ajouter un grand nombre de mots clés plus ou moins pertinents à leurs articles.

Un mot clé pertinent doit donc respecter plusieurs règles :

- Etre compréhensible en dehors du contexte : le mot clé "*Conseil*" ne veut rien dire alors que le mot clé "*Conseils allaitement*" a beaucoup plus de sens.
- Il ne doit pas être identique à un nom de catégorie (car cela sera un pur doublon).
- Il ne doit pas non plus être identique au titre d'un article (car dans ce cas de figure vous compliquez la tâche de Google pour savoir quel page il doit afficher dans les résultats du moteur de recherche).
- Il ne doit pas être similaire à un autre mot clé : par exemple il faudrait éviter de voir coexister des mots clés trop proches, comme avec des variantes singulier/pluriel ou des fautes d'orthographe. Par exemple, évitez les mots clés "*WordPres*" et "*Word Press*".
- Ils ne doivent pas être trop nombreux, car cela implique forcément d'avoir le même contenu sur plusieurs pages de mots clés. Si votre publication possède 10 mots clés associés, vous aurez donc dix fois le même contenu dupliqué (même si cette duplication est partielle).

Rappelez-vous toujours que Google ne juge pas un site sur le nombre de contenus, mais sur la popularité et la pertinence de ceux-ci : plus il y aura des contenus dupliqués, plus il est probable que chacun d'entre eux soit moins pertinent et moins populaire.

Les dates

WordPress va également créer par défaut des archives par date. Elles fonctionnent comme les catégories et les mots clés dans le sens où elles listent vos articles, provoquant de la même manière une contenu dupliqué supplémentaire.

Pire encore, l'aspect sémantique est ignoré car les dates de publication n'impliquent pas forcément un lien logique entre chaque article. Par exemple, l'article A publié en Février peut parler par exemple d'écologie et l'article B publié le même mois peut parler du dernier concert des Spice Girls. Dans plus de 9 cas sur dix, votre site Internet sera dans ce cas de figure.

La solution la plus adéquate est donc d'éradiquer purement et simplement ce type d'archives. Pour cela, il faut effectuer plusieurs actions :

- Dans le menu "**Apparence**" => "**Widgets**", vérifiez qu'aucun contenu et widget ne fait référence à ces archives par date.
- Faites de même dans le menu "**Apparence**" => "**Menus**".
- Dans presque tout les fichiers de votre thème, vérifiez qu'aucun lien n'est fait vers ce type d'archive.
- Dans le plugin WordPress SEO, rendez-vous dans le menu "**SEO**" => "**Titres & Métas**" => "**Autre**" => "**Archives par date**", puis cochez "*désactiver les archives par date*".

Attention, si ces archives étaient actives et indexées par Google, il vous faudra mettre en place des redirections 301 de celles-ci vers l'accueil de votre site, soit avec un fichier htaccess, soit avec l'excellent plugin Redirection :

<http://wordpress.org/extend/plugins/redirection/>

Les auteurs

Les deux derniers types de contenus vont concerner un nombre plus restreint de sites Internet mais peuvent également nuire au référencement naturel. Le premier type est la page d'auteur. Sur une installation standard, une page de ce type contient la liste des articles publiés par l'utilisateur concerné.

Si vous êtes dans le cas de figure d'un site avec un seul auteur, et que vous ne souhaitez pas spécialement vous positionner sur le nom, prénom ou pseudo de celui-ci, cette page n'a aucun intérêt, exactement comme les archives par date. Il faudra donc supprimer tout lien pointant vers ce type de pages. Procédez de la même manière que pour une archive par date :

- Dans le menu "**Apparence**" => "**Widgets**", vérifiez qu'aucun contenu ne fait référence aux auteurs
- Faites également de même dans le menu "Apparence" => "Menus".
- Dans presque tout les fichiers de votre thème, vérifiez qu'aucun lien n'est fait vers les pages auteurs.

A l'inverse, si vous avez plusieurs auteurs ou que vous voulez vous positionner sur le nom de cette personne, il va falloir modifier cette page pour lui apporter un réel contenu. En fonction de votre thématique, les possibilités sont plus ou moins nombreuses mais voici quelques suggestions que vous pouvez mettre en place :

- Ajouter les 5 derniers commentaires de l'auteur

Dans le fichier `author.php` de votre thème, ajoutez ce code en haut de la page

```
<?php $curauth = (isset($_GET['author_name'])) ? get_user_by('slug', $author_name) : get_userdata(intval($author));  
$myid = $curauth->ID;query_posts("author=$myid&posts_per_page=-1");?>
```

Copiez ensuite à l'endroit désiré le code suivant :

```
<?php  
$comments = get_comments("author_email=$curauth->user_email&number=5");  
foreach($comments as $comment) :  
    $theid = $comment->comment_post_ID;  
    echo '<blockquote><p>'. $comment->comment_content. '</p></blockquote>  
    <p>Le '.get_comment_date('d F Y', $comment->comment_ID ).'  
dans <strong><a href="'.get_permalink($theid).'"
```

```
title="'.get_the_title($theid).'">'.get_the_title($theid).'    endforeach;  
>>
```

- Ajouter une image de grande taille pour la photo

Cette étape est relativement simple. Copiez le code ci-dessous dans le fichier `author.php` de votre thème WordPress :

```
<?php echo get_avatar(get_the_author_meta('ID'), '150'); ?>
```

- Ajouter une description plus étoffée de l'auteur

Pour réellement étoffer le contenu, on peut afficher ici une description plus longue de l'auteur. Pour y parvenir, il est nécessaire de suivre trois étapes.

1. Dans le fichier `functions.php` de votre thème copiez le code suivant :

```
//***** Infos supplémentaires sur les auteurs  
add_action( 'show_user_profile', 'extra_user_profile_fields' );  
add_action( 'edit_user_profile', 'extra_user_profile_fields' );  
  
function extra_user_profile_fields($user) { ?>  
<h3>Informations complémentaires sur le profil</h3>  
<table class="form-table"><tr><th><label for="address">Description  
détaillée de l'utilisateur</label></th><td>  
<textarea name="shortdesc" id="shortdesc" cols="8" rows="3"><?php  
echo esc_textarea(get_the_author_meta('shortdesc', $user->ID) );  
>></textarea><br />  
</td></tr></table>  
<?php }  
add_action( 'personal_options_update',  
'save_extra_user_profile_fields' );  
add_action( 'edit_user_profile_update',  
'save_extra_user_profile_fields' );  
  
function save_extra_user_profile_fields( $user_id ) {  
    if ( !current_user_can( 'edit_user', $user_id ) ) {  
        return false;}  
    update_usermeta( $user_id, 'shortdesc', $_POST['shortdesc'] );  
}
```

2. Dans l'administration de WordPress, vous aurez accès à un nouveau champ de description pour chaque profil utilisateur que vous devrez remplir.

3. Copiez pour finir le code suivant dans le fichier `author.php` de votre thème :

```
<?php the_author_meta('shortdesc');?>
```

Les custom taxonomies

Ce dernier cas de figure concerne une proportion encore plus réduite de sites Internet. Avant tout, il faut expliquer de quoi il s'agit. Dans WordPress, une taxonomie est un moyen de classer les contenus. Chacun de vos articles possède donc 4 taxonomies par défaut :

- Les catégories.
- Les mots clés.
- L'auteur.
- La date de publication.

Il s'agit de 4 moyens différents utilisés pour classer et afficher vos différents contenus. Certains plugins ou thèmes peuvent créer leurs propres taxonomies. Vous les verrez théoriquement apparaître dans un nouveau sous-menu de "Article" dans l'administration de WordPress, à côté des sous-menus "Catégories" et "Mots clés".

La question à se poser est de savoir comment elles ont été conçues, et surtout dans quel but. En référencement naturel, il existe deux cas de figures différents pour les custom taxonomies :

- une taxonomie de "gestion", uniquement utilisée pour classer et ordonner des contenus sans donner accès aux visiteurs à un nouveau type d'archives.
- une taxonomie réelle, à savoir présenter un nouveau formats de liens pour des contenus spécifiques.

La taxonomie de "gestion"

Voici un exemple sur le site de SeoMix. La page ci-dessous affiche les articles du site Internet selon une taxonomie. La taxonomie en tant que telle n'est pas affichée et celle-ci sert donc surtout à ordonner et classer des contenus : <http://www.seomix.fr/guide-referencement-naturel/>

Si votre taxonomie ne sert qu'à classer des contenus ou à administrer votre site, vous devez donc vous assurer que votre thème ne fait jamais de lien vers vos archives de taxonomies. En règle générale, c'est déjà le cas mais vérifiez quand même vos fichiers de thème WordPress.

Par contre, ce qui peut réellement vous nuire à l'inverse sont les plugins installés. Nous allons donner un simple exemple : si vous avez installé le plugin **WordPress SEO**, celui-ci va par défaut ajouter toutes les taxonomies dans le fichier sitemap. Assurez-vous donc de désactiver les archives de vos taxonomies (ainsi que celles par auteur si vous ne les utilisez pas).

La taxonomie réelle

Si votre taxonomie doit servir aux internautes, par exemple une taxonomie "films" pour afficher vos contenus *via* une page d'archive dédiée du même nom, vous prenez le même risque qu'avant : avoir un contenu dupliqué entre l'article, la page de taxonomies, les mots clés et les catégories. Il faudra donc utiliser pour chaque page de taxonomie la même méthodologie :

- Tronquer le contenu.
- Ajouter du contenu unique, comme une description de taxonomie avec le code suivant:

```
<?php $termDescription = term_description( '', get_query_var( 'taxonomy' ) );if($termDescription != '') : ?><?php echo $termDescription; ?><?php endif; ?>
```
- Evitez d'associer le même article à trop de taxonomies (que ce soit vos catégories, vos mots clés ou vos custom taxonomies).

Conclusion

Vous l'aurez sans doute compris, WordPress possède encore quelques lacunes pour le référencement naturel. Certes, le contenu dupliqué ne va pas vous bannir de Google mais cela peut vous empêcher de grappiller quelques positions de plus sur les mots clés que vous ciblez.

Pour contrer cet effet négatif, il faudra donc agir tant sur le code de votre thème et de vos plugins que sur la gestion interne de vos contenus par les différents rédacteurs de votre site. Et tout devrait bien se passer...

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

Le Negative SEO au regard de la loi

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Les pratiques de Negative SEO, visant à dégrader le référencement et la visibilité d'un site concurrent sur les moteurs de recherche, sont légion. Mais sont-elles légales ? Un site web victime de tels agissements peut-il se retourner contre Google ou contre les personnes ayant mis en place de telles pratiques ? Explications...

Le "Negative SEO" est une technique qui se répand de plus en plus. Certains pourraient voir, dans ce phénomène, une consécration de notre discipline, mais les limites légales et jurisprudentielles de cette pratique sont en réalité autant de contraintes à prendre en compte pour les acteurs se risquant dans de telles pratiques.

Le concept de Negative SEO

Même si ce qui suit peut paraître une évidence pour certains, il convient de retourner aux fondamentaux pour établir ce qui est, pour nous, le Negative SEO

L'optimisation pour les moteurs de recherche permet de :

- rendre son site visible de façon pérenne (et donc exister commercialement) ;
- capter du trafic qualifié depuis les moteurs de recherche (et donc des clients) ;
- gagner une image d'autorité auprès des internautes (et donc une image de marque).

Le but d'un moteur de recherche étant de fournir la requête la plus pertinente à l'utilisateur, les algorithmes de classement des moteurs sont constamment modifiés pour aller dans ce sens, donc dans celui de l'utilisateur.

Le référencement porte donc sur l'existence commerciale d'une entreprise, sa clientèle (et donc son chiffre d'affaires) et son image de marque. Quand on est un *pure player*, le référencement, c'est la vie (ou la mort) de l'entreprise, en quelque sorte.

Du fait de cet énorme trafic potentiel, les webmasters cherchent par tous les moyens à figurer en tête de liste, quitte à « tricher ». On a noté à une certaine époque l'utilisation de techniques de spamdexing (référencement abusif) destinées à flouer les robots des moteurs pour faire remonter le site.

Les moteurs ont réagi en modifiant leurs algorithmes de classement pour blacklister ou pénaliser ces sites "voyous". A ce titre, nous avons évoqué, dans la Lettre "Recherche et Référencement" du mois dernier, les fondements juridiques qui permettent aux moteurs (dont principalement Google) de blacklister les sites qui violaient certaines règles. Ainsi, en ce qui concerne Google, les « conseils aux webmaster » posent plusieurs principes relatifs à la qualité d'un site et précisent les pratiques considérées par Google comme répréhensibles

(<http://support.google.com/webmasters/bin/answer.py?hl=fr&answer=35769>) :

« Évitez les techniques suivantes :

- Contenu généré automatiquement ;
- Participer à des systèmes de lien ;
- Techniques de dissimulation (cloaking) ;
- Redirections trompeuses ;
- Textes ou liens cachés ;
- Pages satellite ;
- Contenu détourné ;
- Participer à des programmes d'affiliation qui n'apportent pas une valeur ajoutée suffisante ;

- Multiplication de mots clés non pertinents sur les pages ;
- Création de pages générant des opérations malveillantes, telles que des tentatives de phishing ou encore l'installation de virus, de chevaux de Troie ou d'autres programmes malveillants ;
- Utilisation abusive des extraits enrichis balisés ;
- Envoi de requêtes automatiques à Google. »

Dans un contexte fortement concurrentiel, il est donc tentant de profiter de ces règles et comportements des moteurs pour discréditer un concurrent, le rendre coupable de pratiques illicites et le dénoncer au gendarme d'Internet, Google.

Ainsi, le Negative SEO serait une action visant à nuire au référencement d'un site d'un concurrent par différentes techniques.

Dans ce cadre, il est important de noter que les Internautes (et surtout les concurrents) peuvent « reporter » à Google toute pratique considérée comme illicite en envoyant un rapport de spam à l'adresse : <https://www.google.com/webmasters/tools/spamreport>.

Le caractère illégal du Negative SEO

Le Negative SEO se traduit parfois par des comportements soit extrêmement discutables, soit complètement illégaux.

Plusieurs méthodes ont été référencées :

Actions sur l'hébergement du concurrent

Agir sur l'hébergement du concurrent induit techniquement de s'introduire dans le Système d'Information d'un tiers (le concurrent ou l'hébergeur du concurrent). Dans de tels cas, un tel comportement violerait la loi Godfrain (maintien dans les SI) et notamment l'article L.323-1 du Code pénal : « *Le fait d'accéder ou de se maintenir, frauduleusement, dans tout ou partie d'un système de traitement automatisé de données est puni de deux ans d'emprisonnement et de 30000 euros d'amende.* »

Une telle action est donc pénalement répréhensible.

Actions sur l'e-réputation du concurrent

Un tel comportement est également totalement déconseillé car il est constitutif de nombreux délits.

D'une part, dénigrer son concurrent est constitutif de diffamation condamnée par la loi du 29 juillet 1881 à une amende de 12.000 euros. Le Procureur de la République peut poursuivre de tels comportements, y compris ceux commis sur Internet.

D'autre part, dans le cas où des données personnelles seraient utilisées, la loi du 6 janvier 1978 (Informatique et Libertés) sanctionne également pénalement ce comportement, à hauteur de 5 ans d'emprisonnement et de 300.000 euros d'amende (articles 226-25 et suivants du Code pénal). En plus du procureur, la CNIL peut être saisie (notamment par un formulaire en ligne) dans de tels cas.

Actions sur le contenu (duplicate content, etc...) ou sur les backlinks

Le droit français a la particularité de disposer de principes généraux permettant de faire un procès dès que l'on estime qu'une personne a commis une faute (faire un procès ne veut pas dire le gagner, il s'agit ici d'évoquer le fondement juridique permettant uniquement d'ouvrir une procédure).

La responsabilité délictuelle est basée sur l'article 1382 du Code civil. Trois conditions sont nécessaires pour se voir reconnaître responsable d'un fait :

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

En l'espèce, nuire au référencement d'un concurrent direct est, à notre sens, une faute. Le sens commun le ressent ainsi. Mais, la jurisprudence a mis du temps à le préciser et le confirmer.

Dans la lettre R&R du mois dernier, nous évoquions l'affaire Luteciel où une société avait nu à son concurrent en faisant du duplicate content de son site. Malheureusement, le dossier avait été étrangement orienté et le tribunal n'avait apporté que peu de réponses. Toutefois, le tribunal de commerce a relevé que « *l'on doit donc considérer que la mise en ligne d'un site internet, copie servile des codes informatiques d'un concurrent, dans le but de nuire à son référencement est un acte déloyal même si son aspect visuel ne porte pas à confusion pour l'internaute moyen* » mais a rejeté les demandes de CNI car CNI n'a même pas été en mesure de démontrer que le référencement de son site ait souffert des agissements décrits.

On sait donc aujourd'hui que le Negative SEO est un acte déloyal de la part du concurrent indélicat.

Que ce soit dans le cas Luteciel ou dans n'importe quel autre cas de Negative SEO, la preuve du préjudice sera déterminante pour que le concurrent indélicat soit sanctionné.

La responsabilité des moteurs dans un cas de negative SEO

Un arrêt de la Cour d'appel de Paris du 19 mars 2009 a précisément mis en relief ces difficultés à retenir une faute de la part d'un moteur. Nous avons d'ailleurs commenté cet arrêt dans cette lettre à l'époque (<http://recherche-referencement.abondance.com/2009/05/la-responsabilite-des-moteurs-de.html>).

Pour rappel, une société a assigné Google et Yahoo! en se plaignant du référencement naturel d'un site la dénigrant et leur reprochant de ne pas avoir vérifié le contenu du site. Les moteurs se sont défendus en soulignant qu'ils procèdent à un référencement de manière automatique, sans aucun contrôle a priori sur leurs robots ne pouvant en conséquence maîtriser le référencement.

La Cour a rejeté la responsabilité de Google et Yahoo! en précisant que ces moteurs n'avaient pas à vérifier le contenu de chaque site et de s'assurer de leur conformité à la loi. La Cour a ainsi conclu que « *ce caractère automatique des résultats affichés et l'absence de toute analyse de contenu excluent une intention de nuire ou délictueuse, que ce soit de diffamer ou de dénigrer* ». En d'autres termes (et pour être schématique), pas de faute possible tant que l'algorithme est neutre et qu'aucune intervention et/ou volonté humaine n'intervient.

Cette jurisprudence est clairement transposable au Negative SEO. Tant que l'algorithme est neutre et qu'aucune intervention et/ou volonté humaine n'intervient dans le classement des résultats, aucune faute ne peut être imputée aux moteurs. Si un site est déclassé par Panda ou Penguin dans le cadre des résultats naturels, le site concerné ne pourra pas prouver de faute et donc, ne pourra pas gagner un éventuel procès intenté contre Google.

La pratique du negative SEO

Dans la pratique, il est en réalité assez fréquent de voir des comportements de Negative SEO. Dans les secteurs fortement concurrentiels, c'est même une pratique courante.

L'inertie des moteurs

Or, nous avons constaté, depuis plusieurs années maintenant, une certaine inertie des moteurs lorsque les concurrents malheureux « reportaient » leur situation. Il est même banal de voir ces « reports » laissés sans aucune réponse.

Nous estimons que cette situation est particulièrement pénible pour toute l'économie : si les moteurs imposent des règles mais ne les font pas respecter ou, pire, les font respecter à certains mais non à d'autres, cela induit de fortes distorsions de concurrence et, surtout, des situations économiquement non viables. De plus, les moteurs étant, pour ces acteurs, des arbitres centraux et incontournables, leurs décisions n'en sont que plus importantes. Et, à l'inverse d'une décision de justice, aucun appel ou aucune contestation n'est possible.

En conséquence, nous appelons de nos vœux à ce que les moteurs fassent respecter leurs propres règles et appliquent ce qu'ils demandent aux autres d'appliquer. Seul un tel comportement responsable permettrait de garantir la très fameuse « neutralité » du Net...

Le manque de moyens de la justice pénale

De même, pour ce qui concerne la justice, force est de constater que les autorités pénales ont grande difficulté à faire face à la recrudescence des faits rapportés. Lorsque l'Office central de lutte contre la criminalité liée aux technologies de l'information et de la communication est saisi, il est fréquent que le dossier n'aboutisse jamais : les fonctionnaires sont des professionnels brillants et consciencieux, mais il leur manque de réels moyens. Par exemple, cet Office dispose d'une centaine de professionnels pour toute la France...

A nouveau, saisir la justice pénale ou la CNIL (qui elle-même saisi la justice pénale) n'aboutit que rarement...

La justice civile

Finalement et malheureusement, seule la voie de la justice civile (où l'on obtient uniquement des dommages et intérêts) semble une voie statistiquement ouverte.

Cette voie est longue, mais elle aboutit à un résultat. Ce résultat peut être décevant, comme pour l'affaire CNI, mais il existe. Dans de telles procédures, l'expérience nous apprend qu'il est important de prendre des professionnels du référencement, notamment pour établir des rapports d'expert pour rapporter les pratiques de négative SEO développés par le concurrent et le préjudice subi.

Alexandre Diehl

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

EdgeRank et GraphRank : des algorithmes au cœur du fonctionnement de Facebook

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Pour classer leurs résultats, les moteurs de recherche tels que Google utilisent aujourd'hui des algorithmes extrêmement sophistiqués. Mais les réseaux sociaux comme Facebook vont également, et jour après jour, dans cette voie pour choisir et classer les informations qu'ils vont afficher sur le mur d'un utilisateur. Une connaissance approfondie de ces mécanismes devient donc essentielle dans le succès - ou l'échec - d'une campagne de promotion sur ce réseau social. Plongeons donc dans cet article pour une exploration de l'EdgeRank et du GraphRank, deux générations de systèmes d'évaluation utilisées par Facebook à cette fin...

EdgeRank Facebook

$$\sum_{\text{edges } e} u_e w_e d_e$$

Facebook est devenu un support incontournable pour ceux qui veulent développer la visibilité de leur marque, interagir avec leurs communautés ou faire connaître leur site. Mais le Net regorge d'histoires de campagnes ratées sur le leader des réseaux sociaux, certains allant jusqu'à renoncer à ce canal, jugé décevant.

En réalité, l'explication de ces échecs réside souvent dans une méconnaissance ou une mauvaise compréhension des algorithmes utilisés par Facebook pour sélectionner les items qui seront sélectionnés dans les flux d'actualités de chaque utilisateur. L'algorithme central dans ce processus s'appelle l'**EdgeRank**.

Qu'est-ce que l'EdgeRank ?

Un utilisateur moyen de Facebook est relié à des dizaines, quand ce ne sont pas des centaines de profils d'amis ou de connaissances. Il suit également plusieurs pages fans. Si chaque action produite sur toutes ces pages était montrée dans le flux d'actualités de l'utilisateur, ce dernier serait noyé sous un déluge d'informations.

Facebook cherche donc à filtrer et à classer ces informations. Pour y parvenir, il utilise une fonction d'évaluation : l'EdgeRank. Notons que "fonction d'évaluation" est un terme plus exact qu'algorithme dans ce contexte, même si désormais il est devenu habituel de l'appeler l'"algorithme Edgerank".

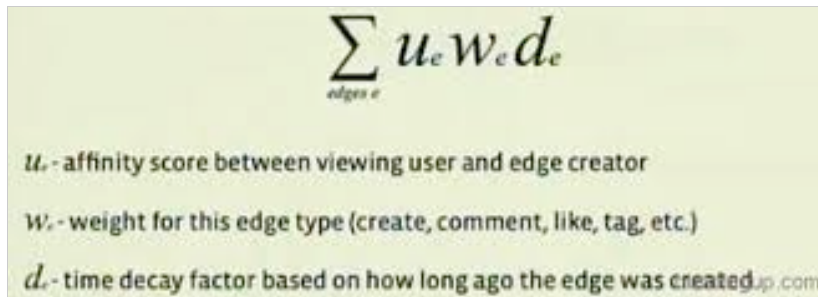
Ce qui est évalué dans l'EdgeRank, c'est le niveau d'importance de l'item à afficher dans les flux d'actualité. Notons que Facebook préfère appeler ces items des Objets.

Les Objets dont l'importance est jugée insuffisante sont éliminés et ne seront donc pas montrés à l'utilisateur.

Les critères utilisés dans l'EdgeRank

Il existe trois principaux critères utilisés dans la fonction d'évaluation :

- le niveau d'affinité (**Affinity**) ;
- le poids (**Weight**) ;
- l'érosion temporelle (**Time Decay**).



The image shows a screenshot of a presentation slide. At the top, the formula $\sum_{\text{edges } e} u_e w_e d_e$ is displayed. Below the formula, three definitions are listed: *u* - affinity score between viewing user and edge creator; *w* - weight for this edge type (create, comment, like, tag, etc.); *d* - time decay factor based on how long ago the edge was created. The source is cited as 'La Formule de l'Edgerank, tirée de la présentation au forum développeurs de Facebook F8 en 2010'.

La Formule de l'Edgerank, tirée de la présentation au forum développeurs de Facebook F8 en 2010

Le niveau d'affinité

Le niveau d'affinité mesure (comme son nom l'indique) l'étroitesse des relations entre l'utilisateur et l'Objet, ou la correspondance entre les profils. Pour un item émis par un ami dont vous n'avez pas consulté le profil depuis deux ans, et avec lequel vous n'avez rien échangé depuis longtemps, ni interagi d'aucune manière, le niveau d'affinité sera évalué comme très bas.

Présenté comme cela, on pourrait penser que le niveau d'affinité est considéré comme symétrique dans l'algorithme : si A est en forte affinité avec B, B est forcément en forte affinité avec A. Mais pas du tout. Le niveau d'affinité mesure en fait une "probabilité d'intérêt" de A pour B, ce qui ne garantit pas que B s'intéresse à A.

Le poids

Le poids d'un item combine la prise en compte de la nature de l'item et le mode et l'importance des interactions avec cet item. Facebook ne communique pas sur les poids réels attribués à certaines classes d'items, mais on considère généralement que les photos, les vidéos et les liens se voient attribués les niveaux de poids les plus importants. En ce qui concerne les modes d'interaction, un "Like" se voit attribuer un poids inférieur à celui d'un commentaire.

L'érosion temporelle

L'érosion temporelle est un facteur qui dépend de l'ancienneté de l'item. Sur Facebook, l'obsolescence des items est très rapide, les "news" disparaissent en quelques heures des flux d'actualités...

POSTS ON NEWS FEED (in order)	EDGE	AFFINITY	WEIGHT	TIME DECAY	EDGE VALUE
Brother Created recently NO ENGAGEMENT	Photo				HIGH
High School Friend Created 12 hours ago HIGH ENGAGEMENT	Status				LOW
	Mutual Friend Comment				MED
	Mutual Friend Comment				HIGH
	Mutual Friend Comment				HIGH
Brand Page Created 4 hours ago HIGH ENGAGEMENT	Photo				HIGH
	Fan Like				LOW
	Fan Share				HIGH
	Fan Comment				HIGH
Acquaintance Created 2 hours ago LOW ENGAGEMENT	Status				LOW

L'edgeRank en action : pourquoi certains Objets sont affichés et d'autres pas (illustration de Edgerankchecker)

Euh ... pourquoi cela s'appelle l'EdgeRank, au fait ?

Edge, dans ce contexte, se traduit par arête"... Il s'agit d'un terme typique de la théorie des graphes, une théorie mathématique qui connaît bien sûr de nombreuses applications sur le graphe social.

Un "edge" est ce qui relie deux points du graphe social. Toute relation entre deux points du graphe, quelle que soit sa nature, est qualifiée du terme générique "edge". Et toute interaction entre deux points du graphe produit une nouvelle relation, et une nouvelle "arête".

Sauf que Facebook n'appelle plus cet algorithme "Edgerank". Même si par habitude, Edgerank est un terme encore communément employé par les experts du réseau social et même en interne chez Facebook, un ingénieur a révélé en juillet dernier que ce terme avait été officiellement abandonné.

Comment connaît-on cet algorithme ?

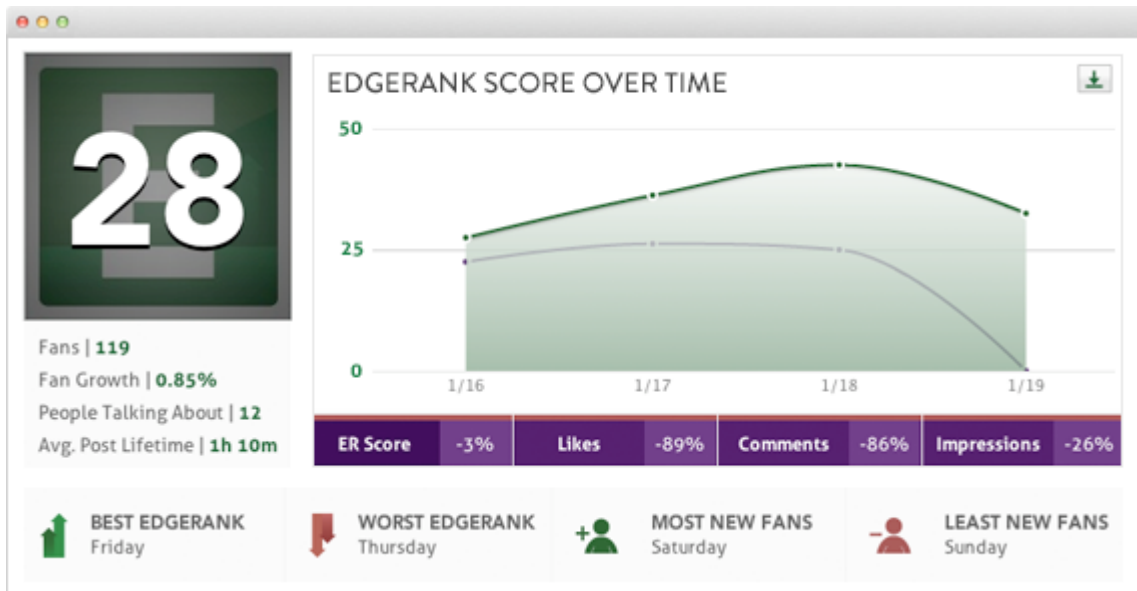
Ce sont deux ingénieurs de Facebook, Ruchi Sanghvi et Ari Steinberg qui ont présenté pour la première fois l'algorithme au cours de la conférence développeurs Facebook F8, le 21 avril 2010.

La vidéo de cette conférence n'est hélas plus accessible dans la zone développeurs de Facebook.

Peut-on estimer son Edgerank ?

La formule exacte de calcul de l'Edgerank est inconnue. Et Facebook ne communique pas les valeurs d'EdgeRank. On peut donc uniquement créer une fonction d'évaluation similaire et tenter de mesurer soi-même des critères comme le niveau d'affinité, le poids, et l'érosion temporelle des objets. Mais l'exercice a ses limites, et elles sont vite atteintes. Même si l'Edgerank n'est pas un algorithme sophistiqué, et même si la plupart des réseaux sociaux ont développé leur propre version de l'Edgerank (LinkedIn par exemple), on sait trop peu de choses sur la façon dont il est réellement calculé pour parvenir à prédire exactement ce qui va réellement apparaître dans les "newsfeeds".

Un outil (payant) disponible sur le Net fournit une évaluation des EdgeRanks aux propriétaires de pages Fan (par exemple) qui veulent suivre l'évolution de l'EdgeRank estimé. Cet outil (Edgerankchecker : <http://edgerankchecker.com/>) fournit une estimation intéressante qui permet de voir si ses efforts d'optimisation de l'Edgerank portent leurs fruits, mais il ne faut surtout pas prendre les valeurs produites par l'outil pour argent comptant.



Edgerankchecker, un outil qui sert à estimer l'Edgerank de ses objets

Le GraphRank

Au cours de la conférence développeurs F8 de 2011, Facebook a présenté un nouvel algorithme similaire au Edgerank, destiné à "sélectionner" les Objets à afficher produit par un nouveau type d'applications alimentant le graphe social en actions (comme le fait d'écouter un morceau de musique). Cet algorithme s'appelle le **GraphRank**. On sait peu de choses sur son mode de calcul réel, mais il semble reprendre les principes de l'Edgerank.

Un algorithme qui a été récemment fortement modifié

Au cours de l'automne 2012, un grand nombre d'observateurs ont constaté que le "reach" (la "couverture", le nombre d'utilisateurs qui voient s'afficher un post donné) des Objets qu'ils postaient sur Facebook avait brusquement chuté. Des rumeurs à propos d'un nouveau score de qualité ou d'autorité des pages ont commencé à se répandre.

En réalité, Facebook avait commencé à être confronté à des méthodes de spam "industrielles", et a décidé d'introduire un nouvel élément dans l'algorithme pour

minimiser la visibilité de ce spam. Ce nouveau score tient compte du "negative feedback" sur les pages. Une page qui a fait l'objet de nombreuses "plaintes" sera pénalisée, alors qu'une page qui a un nombre de plaintes plus bas que la moyenne sera favorisé.

Ce changement dans l'algorithme a été confirmé par Will Cathcart, product manager des "newsfeeds" chez Facebook, dans un entretien accordé au magazine en ligne

Techcrunch :

<http://techcrunch.com/2012/11/07/killing-rumors-with-facts-no-facebook-didnt-decrease-page-news-feed-reach-to-sell-more-promoted-posts/>

Comment la connaissance de cet algorithme permet-elle de réussir ses campagnes d'action sur Facebook ?

La formule de l'EdgeRank nous apprend pourquoi certaines marques font des « flops » sur Facebook même en dépensant des millions de dollars en achat de produits publicitaires sur Facebook.

Ce n'est pas tout d'avoir des fans ou des amis, il faut qu'ils soient « engagés »

Si vous n'avez pas réussi à créer un « engagement » de vos fans sur une page Facebook, c'est-à-dire si vos précédents posts et objets n'ont engendré aucune réaction et aucune interaction avec votre communauté, alors toute nouvelle publication risque de n'être visible que par un tout petit nombre d'utilisateurs, y compris si vous lâchez une information qui est une véritable « bombe » ou si vous publiez une vidéo virale particulièrement réussie.

Postez les bons objets, et des objets populaires

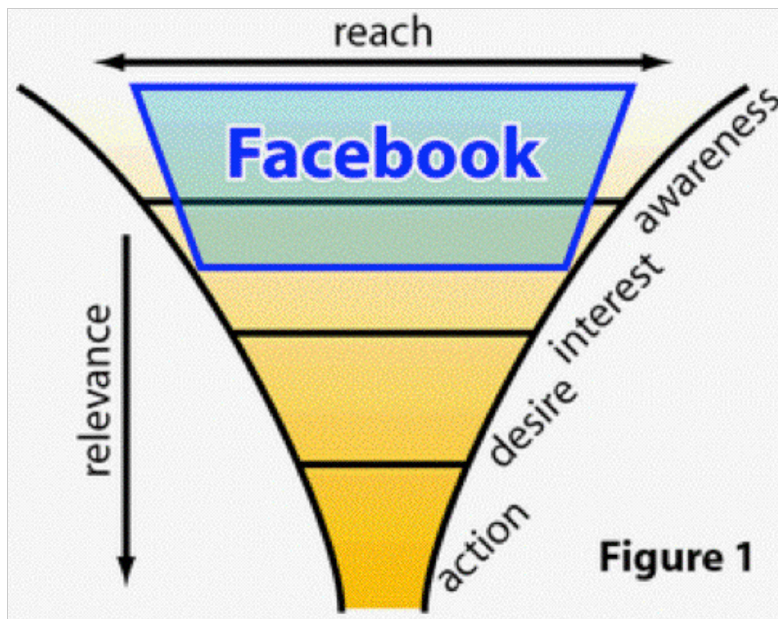
Images, vidéos et liens ont un poids important, c'est ce qu'il faut privilégier. Communiquer *via* un changement de statut vous permettra de toucher un nombre plus réduit de membres.

Avant l'heure, c'est pas l'heure, après l'heure c'est plus l'heure

Le facteur « time decay » a une implication évidente : si vous publiez vos posts ou si vous déclenchez une action quelconque un jour ou à une heure pendant lesquels un faible nombre des utilisateurs sont connectés, il y'a peu de chances que l'Objet Facebook en question soit vu.

Et ne faites pas n'importe quoi

Attention aux feedback négatifs : le temps où l'on pouvait faire n'importe quoi sur Facebook en passant sous le radar parce qu'il n'y avait pas assez d'évaluateurs humains pour éliminer les spammeurs est révolu. Même les comportements « limites » tolérés par Facebook sont maintenant pénalisés. Si vous faites l'objet de plaintes d'utilisateurs, le « reach » de vos campagnes risque de tomber à la portion congrue.



Source Michael WU : attention, le « reach » d'une campagne n'est pas tout, il faut aussi faire attention à ce que la publication soit « pertinente » pour les personnes touchées. En fait, plus le « reach » est important, plus la pertinence risque d'être dégradée. Il faut donc trouver un juste milieu...

Conclusion : le SMO s'est sophistiqué et demande aussi qu'on s'intéresse aux algorithmes du graphe social

Les applications du graphe social ont atteint un niveau de maturité et d'usage suffisant chez les utilisateurs pour obliger les principaux leaders du secteur à déployer des algorithmes destinés à trier, classer et filtrer ce qui est montré aux utilisateurs.

Le résultat est qu'il n'est plus possible de se contenter de connaître par cœur les fonctionnalités et les produits proposés par les réseaux sociaux pour savoir mener des campagnes efficaces sur ces supports. Il faut maintenant s'intéresser aux mécanismes au cœur du fonctionnement de ces outils. Pour le moment, nous ne sommes pas encore confrontés à des algorithmes d'une complexité comparable à ceux que l'on rencontre dans les moteurs de recherche, mais on s'en rapproche à grand pas...

Car les volumes d'information à traiter sont considérables, et les mathématiques décrivant des graphes aussi complexes que le graphe social très sophistiquées. Gageons qu'avec la recherche effrénée lancée sur le sujet au sein des outils sociaux et des universités, de nouvelles découvertes vont entraîner le monde du SMO vers des sphères inexplorées...

Bibliographie :

La page de facebook sur le sujet :
(plutôt laconique, voire elliptique)
<http://www.facebook.com/help/search/?q=edgerank>

L'outil Edgerankchecker :
<http://www.edgerankchecker.com>

L'article de Techcrunch sur l'entretien avec Will Cathcart :
<http://techcrunch.com/2012/11/07/killing-rumors-with-facts-no-facebook-didnt-decrease-page-news-feed-reach-to-sell-more-promoted-posts/>

Un article scientifique passionnant sur le sujet :

All Friends are Not Equal: Using Weights in Social Graphs to Improve Search

Sudheendra Hangal Diana MacLean Monica S. Lam Jeffrey Heer
Computer Science Department
Stanford University, Stanford, CA 94305
<http://xenon.stanford.edu/~hangal/weighted-social-graphs.pdf>

Le brevet sur l'Edgerank ?

Dynamically providing a news feed about a user of a social network
<http://www.google.com/patents/US7669123>
Brevet attribué à Facebook

Philippe YONNET , *Directeur de l'agence Search-Foresight / Groupe MyMedia.*
Président de l'association SEO Camp (<http://www.seo-camp.org/>)

Revue d'outils moteurs et SEO

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

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

Nom : [Facilement.com](#)

URL : <http://www.heyhop.com/facilement/>

Domaine : Recherche

Type : Site web

Descriptif : Un "moteur de résultat" dont le principe est de rediriger automatiquement l'internaute dès lors que la requête est identifiable de manière unique. Concrètement, la requête "Orange" vous redirige automatiquement vers le site Orange.fr...



Nom : [AuthorCrawler](#)

URL : <http://www.tomanthony.co.uk/tools/author-crawler/>

Domaine : Recherche - SEO

Type : Logiciel

Descriptif : Utilise la détection de l'attribut rel="author" développé par Google pour définir qui fait des liens vers vous sur le Web.

AuthorCrawler

Find who is linking to you. A free open-source SEO tool for link building.

[About](#) [Example Reports](#) [Instructions](#) [Release Notes](#) [Download](#) [Contact](#)

AuthorCrawler is a proof-of-concept tool that I built to highlight the ways in which the SEO community could use Google's authorship markup to harvest data about links.

The tool examines the backlinks to a URL and crawls them looking for authorship markup, which it will follow to download a wealth of information on the author of those links. It uses the [SEOMoz API](#) for the links, and then crawls them each.

You'll end up with a report allowing you to sort your backlinks by the social standing of the authors of those links. See the [example reports](#) to see how this might look.

[Download AuthorCrawler](#)

Author	Circles	Followers	Contributes to
 Gianluca Fiorilli Google+ Twitter	3592	3862	www.abondance.com www.seomoz.org/blog/ www.seomoz.org/blog/ www.abondance.net/
 Ben Lurie Google+ Twitter	4358	11120	www.conversionmarketing.com www.porkbut.com www.thegigland.com www.thegigland.com/
 Randy Fichtel Google+ Twitter	36115	54832	

Nom : [SimplyTestable](#)

URL : <http://simplytestable.com/>

Domaine : SEO

Type : Site Web

Descriptif : Vérifie la compatibilité W3C de toutes les pages d'un site dont on fournit l'URL de la page d'accueil. Attention : peut être très long en fonction de la taille du site testé...



Automated front-end web testing

Save time and take the pain out of routine front-end web testing. One-click testing of your entire site.

- ★ site-wide HTML markup validation, CSS validation and JavaScript linting in one click
- ★ [overview report](#) of your entire site's successes and failures
- ★ detailed [per-URL error report](#)

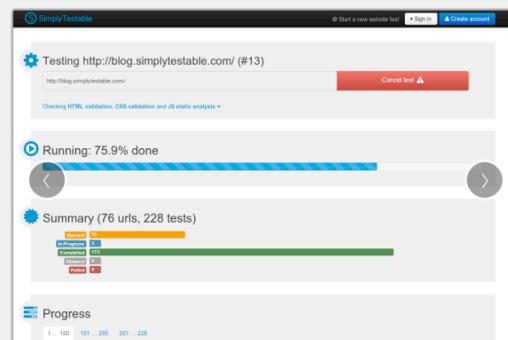


Enter website homepage URL ...

[Test your site now](#)

HTML validation CSS validation JS static analysis

[Advanced new test form \(fine tune your test\) ▶](#)



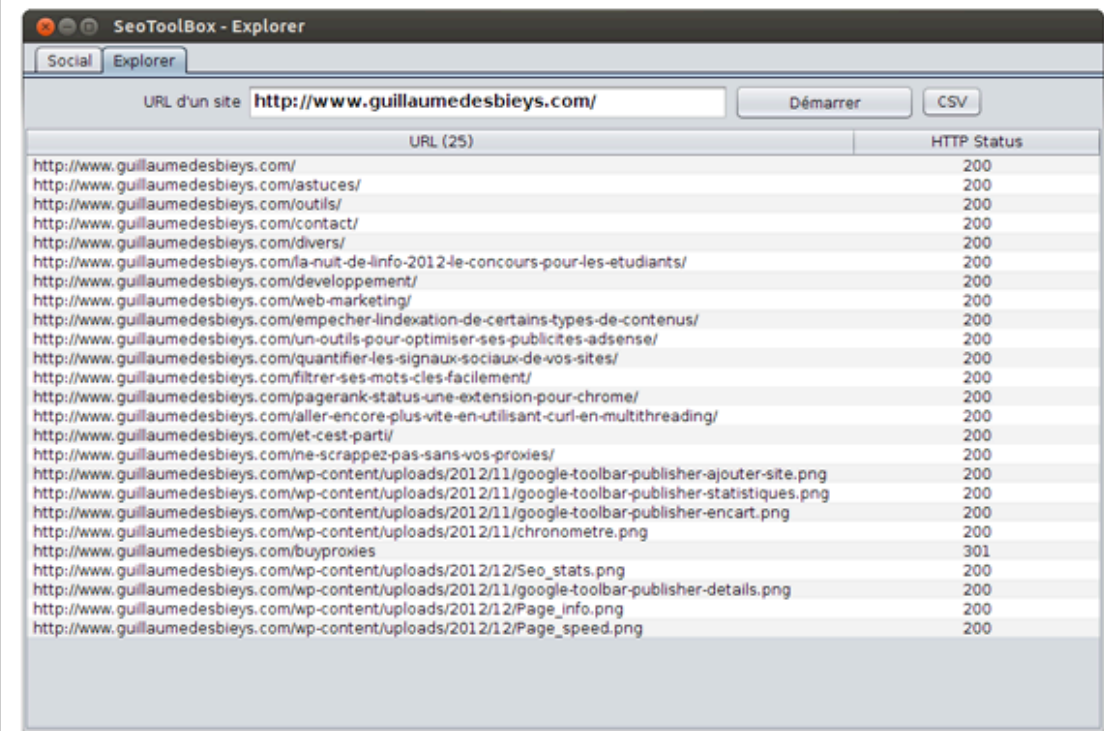
Nom : [SEO Toolbox](#)

URL : <http://tools.guillaumesdesbieys.com/seotoolbox.html>

Domaine : SEO

Type : Logiciel

Descriptif : Ensemble de petits outils SEO comme l'exploration d'un site web, nombre de signaux sociaux, etc.



The screenshot shows the 'SeoToolBox - Explorer' interface. At the top, there are tabs for 'Social' and 'Explorer'. Below the tabs, the 'URL d'un site' field contains 'http://www.guillaumesdesbieys.com/'. To the right of this field are buttons for 'Démarrer' and 'CSV'. The main area displays a table with two columns: 'URL (25)' and 'HTTP Status'. The table lists 25 URLs, all of which return a status of 200, except for one which returns 301.

URL (25)	HTTP Status
http://www.guillaumesdesbieys.com/	200
http://www.guillaumesdesbieys.com/astuces/	200
http://www.guillaumesdesbieys.com/outils/	200
http://www.guillaumesdesbieys.com/contact/	200
http://www.guillaumesdesbieys.com/divers/	200
http://www.guillaumesdesbieys.com/la-nuit-de-linfo-2012-le-concours-pour-les-etudiants/	200
http://www.guillaumesdesbieys.com/developpement/	200
http://www.guillaumesdesbieys.com/web-marketing/	200
http://www.guillaumesdesbieys.com/empêcher-lindexation-de-certains-types-de-contenus/	200
http://www.guillaumesdesbieys.com/un-outils-pour-optimiser-ses-publicites-adsense/	200
http://www.guillaumesdesbieys.com/quantifier-les-signaux-sociaux-de-vos-sites/	200
http://www.guillaumesdesbieys.com/filtrer-ses-mots-cles-facilement/	200
http://www.guillaumesdesbieys.com/pagerank-status-une-extension-pour-chrome/	200
http://www.guillaumesdesbieys.com/allier-encore-plus-vite-en-utilisant-curl-en-multithreading/	200
http://www.guillaumesdesbieys.com/et-cest-parti/	200
http://www.guillaumesdesbieys.com/ne-scrapez-pas-sans-vos-proxies/	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/11/google-toolbar-publisher-ajouter-site.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/11/google-toolbar-publisher-statistiques.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/11/google-toolbar-publisher-encart.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/11/chronometre.png	200
http://www.guillaumesdesbieys.com/buyproxies	301
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/12/Seo_stats.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/11/google-toolbar-publisher-details.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/12/Page_info.png	200
http://www.guillaumesdesbieys.com/wp-content/uploads/2012/12/Page_speed.png	200

Nom : Watussi Box

URL : <http://box.watussi.fr/>

Domaine : SEO

Type : Ana

Descriptif : Analyseur de logs orienté SEO : visite des robots, codes HTTP, volume de pages d'un site web, etc. Présentation ici : <http://www.watussi.fr/presentation-de-watussi-box-analyseur-de-log-seo-open-source>

Le 1^{er} analyseur de logs SEO gratuit

Vous cherchez à étudier le comportement de Googlebot ? Vous galérez à analyser vos logs en ligne de commande ? Les données de Google Webmaster Tools ne vous suffisent pas ?

Téléchargement GRATUIT

Suivi de crawl

Identifiez le nombre de pages crawlées quotidiennement et sur les 30 derniers jours.

Pages actives

Découvrez combien de pages vous apportent réellement du trafic.

Volume de pages

Savez-vous combien de pages de votre site sont connues de Google ?

Temps réels

Analysez en temps réel les pages visitées par Googlebot et celles qui vous génèrent du trafic SEO.

Code Retour

Vérifiez que les codes retour de votre serveur sont corrects et identifiez rapidement les problèmes.

Evolutif

La Watussi Box est gratuite... et Open Source ! Les possibilités d'évolution sont infinies !

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

Revue de moteurs

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Cette rubrique signale les informations réellement utiles émanant des moteurs de recherche eux-mêmes en termes de référencement naturel. La source étant directe, l'information est bien souvent intéressante puisque ce sont les moteurs qui dialoguent ainsi, le plus souvent au travers de blogs spécialisés et d'interviews, avec les webmasters. Voici donc une synthèse de l'information officielle publiée par les moteurs de recherche sur le référencement, ainsi que le signalement des articles particulièrement pertinents écrits par les sites spécialisés ce mois-ci...



Google, et notamment son blog destiné aux webmasters (<http://googlewebmastercentral.blogspot.com/>) ou ailleurs, a surtout publié quelques articles et vidéos plus ou moins intéressants ce mois-ci :

- *Webmaster Tools verification strategies* (<http://googlewebmastercentral.blogspot.fr/2012/12/webmaster-tools-verification-strategies.html>) qui explique les différentes façons de faire vérifier un site dans les Webmaster Tools.
- *An update on Google's webmaster videos* (<http://youtu.be/7qQojfZoR2w>) où Matt Cutts "pête un câble" en direct et nous montre la prochaine version de l'algorithme de Google : "Dinosaure"...
- *What has been the biggest surprise of 2012 for you and your spam team?* (http://youtu.be/D2H_Y7h1roQ)

Les autres actualités Google du mois (et notamment quelques unes de ces vidéos de Matt Cutts et bien d'autres informations, non reprises dans cette rubrique) sont à découvrir ici : <http://www.abondance.com/actualites/>



Bing a publié ce mois-ci sur son blog pour Webmasters (http://www.bing.com/community/site_blogs/b/webmaster/default.aspx) un post intéressant :

- *Bing Search Quality Insights: Building a State-of-the-Art Speller* : (http://www.bing.com/community/site_blogs/b/searchquality/archive/2013/01/03/bing-search-quality-insights-building-a-state-of-the-art-speller.aspx) ou comment fonctionne le correcteur d'orthographe de Bing...
- *Crawling Basics Webinar Video now posted* (http://www.bing.com/community/site_blogs/b/webmaster/archive/2012/12/27/crawling-basics-webinar-video-now-posted.aspx) : une vidéo sur le crawl et l'indexation de votre site par BingBot.



D'autre part, une fin d'année (et par conséquent le début d'une autre) est toujours l'occasion de faire des bilans, des rétrospectives sur les 12 mois qui viennent de s'écouler et des prédictions sur les 365 jours qui viennent. Des prédictions à vérifier, bien sûr, fin 2013 :-)

Voici donc quelques articles qui nous ont semblé intéressants à lire à ce sujet. En français tout d'abord :

-> Demain, le Search selon Google...

Monétiser la recherche sur internet est le cœur d'activité de Google. En cette fin 2012, Google domine outrageusement avec une part de marché mondiale de 79%. Et pour rester leader, Google n'a d'autre choix que d'innover. Vers quoi ?...

<http://www.journaldunet.com/solutions/expert/52996/demain--le-search-selon-google.shtml>

-> Les 4 hacks SEO de 2012

Voir les backlinks de ses concurrents dans Google Webmaster Tools, accéder à des informations confidentielles, remplacer son rival dans les résultats... Voici les "hacks" SEO qui auront marqué l'année 2012...

<http://www.journaldunet.com/solutions/seo-referencement/hack-et-faille-seo/>

-> Bilan : les 12 points clés SEO de l'année 2012

Fin d'année 2012 et j'ai bien évidemment largement mis mes mains dans le cambouis du référencement cette année. Dans quelques poignées de jours nous passons en 2013 (sauf si la fin du monde à réellement eu lieu)...

<http://www.keeg.fr/2012/12/20/bilan-les-12-points-cles-seo-de-lannee-2012/>

-> Tendances webmarketing pour 2013

Tendances digitales pour 2013, c'est un sujet ambitieux. Tant il est vrai que tout bouge rapidement ! Bref retour sur 2012 pour mieux comprendre 2013...

<http://www.mauricelargeron.com/emarketing-sem-et-seo-en-2013/>

-> SEO : bilan 2012 et tendances 2013, entretien avec Simon Gardette, expert SEO

En ce dernier jour de l'année, j'avais envie de dresser un état des lieux du SEO en 2012, parcourir les moments marquants de cette année dans le domaine et tenter de dégager quelques tendances pour 2013...

<http://www.zdnet.fr/actualites/seo-bilan-2012-et-tendances-2013-entretien-avec-simon-gardette-expert-seo-39785787.htm>

-> 10 nouvelles qui ont marqué l'actualité SEO en 2012

Il n'y a pas à dire, l'année 2012 fut riche en rebondissements, particulièrement si vous travaillez dans l'industrie SEO...

<http://www.davidcarlehq.com/2012/12/10-nouvelles-qui-ont-marque-lactualite-seo-en-2012/>

-> Ce que vous réserve l'e-business en 2013

Télécoms, paiement électronique, e-marketing, mobile, e-publicité, e-books, gaming... Voici les tendances qui devraient marquer le secteur de l'économie numérique cette année...

<http://www.journaldunet.com/ebusiness/le-net/tendances-numerique-2013/>

-> Rétrospective webmarketing 2012

Des faits, je les ai compté, il y en a 99 ! Simplement des faits pour rappeler les différents moments importants à garder en mémoire pour l'année qui vient de s'écouler...

<http://www.mauricelargeron.com/seo-adwords-analytics-et-reseaux-sociaux-en-2012/>

-> 2013, l'année web démarre en fanfare

Obligation de payer pour faire un lien vers un journal. Suppression de la publicité. Plus de statistiques Google analytics ! À peine remis des agapes de fin d'année, les webmasters ont été ensevelis sous une avalanche de nouvelles surprenantes qui bousculent des schémas que l'on pensait établis...

<http://blog.axe-net.fr/liens-payants-pub-interdite/>

-> Prophéties SEO - 2013

Comme bon nombre de référenceurs, j'ai essayé de deviner les tendances SEO de l'année 2013. Une belle année qui se profile là, pleine de surprises et de rebondissements...

<http://www.webmaster-referencement.fr/propheties-seo-2013.html>

-> Référencement localisé en 2013

Un semestre s'est déjà écoulé depuis le dernier article sur le référencement local. Pour ceux qui n'auraient peut-être pas suivi je ferai un petit rappel sur les derniers mouvements dans le Google land local et énumérerai les facteurs favorables pour être retrouvé localement en 2013...

<http://www.mauricelargeron.com/recherche-locale-sur-google/>

-> La structure du référencement naturel

Beaucoup de sites Internet proposent leur vision du référencement naturel, mais je suis souvent resté sur ma faim en y trouvant des erreurs ou des oublis...

<http://www.seomix.fr/structure-seo/>

-> L'article ultime sur le robots.txt !

Un article tout bête sur le robots.txt aujourd'hui ? Oui ! Enfin pas tout à fait, ce n'est pas un article tout bête puisque j'ai pu remarquer à plusieurs reprises que ce fichier était souvent mal compris et donc mal utilisé...

<http://noviseo.fr/2013/01/larticle-ultime-sur-le-robots-txt/>

-> Google 2013

En 2012, il ne s'est pas passé un mois sans que Google n'annonce une modification de son algorithme, l'ajout d'un filtre ou la nouvelle version d'un animal de malheur pour de nombreux webmasters...

<http://blog.axe-net.fr/google-2013/>

-> 5 résolutions SEO à suivre en 2013

Si le récent passé est garant de l'avenir, nous sommes en droit de nous attendre à une année stimulante et remplie de défi, qui viendra définir l'industrie à tout jamais...

<http://www.davidcarlehq.com/2013/01/5-resolutions-seo-a-suivr-en-2013/>

-> La recherche sémantique, le défi SEO de 2013 ?

Et si après Penguin et le Knowledge Graph, de nouvelles stratégies SEO basées sur la recherche sémantique étaient déjà en train d'émerger. C'est ce pense l'expert David Degrelle...

<http://www.journaldunet.com/solutions/seo-referencement/seo-knowledge-graph-et-recherche-semantique.shtml>

Et en anglais :

-> SEO Revelations for 2013

The past 2 years have brought in a tremendous amount of change in the world of search, and 2013 will be no different. More Panda releases, Penguin, The EMD update, top heavy ads, and the DMCA update all mark a shift in capabilities for Google. There is more to come...

<http://searchenginewatch.com/article/2235925/SEO-Revelations-for-2013>

-> Where Link Building is Heading in 2013

Another year has come and gone. For link builders, 2012 may well be remembered as the year that link building truly changed forever...

<http://searchenginewatch.com/article/2233829/Where-Link-Building-is-Heading-in-2013>

-> The future according to Google's Larry Page

Google CEO Larry Page envisions a future in which computers plan your vacations, drive your cars, and anticipate your whims. Audacious? Maybe. But Page's dreams have a way of coming true...

http://tech.fortune.cnn.com/2013/01/03/google-larry-page/?hpt=te_r1

-> 15 experts on SEO tips for 2013

2012 was eventful to say the least. Last year I did an interview on SEO tips for 2012. This year is no different. I asked 15 trusted experts in the SEO community about the changes we saw last year and what we can expect in 2013. This is a must read for anyone working with SEO...

<http://edgyseo.com/seo/seo-tips-2013/>

-> The Best of 2012: Top Posts and People of the Year

Now that 2012 has come to a close, it's time to take a look back at the posts and community members that made the biggest impacts on the Moz blog throughout the year!...

<http://www.seomoz.org/blog/the-best-of-2012>

-> Link Building the Right Way in 2013

Link building has always been a risk:reward proposition. This continues to be true in 2013. What has changed over the years is the degree of risk associated with operating outside of the Google Webmaster Guidelines...

<http://searchenginewatch.com/article/2233495/Link-Building-the-Right-Way-in-2013>

-> The Hottest Reads In Search & Social In 2012

We already shared the news that Jordan Kasteler took home the top prize and had a second entry in our overall list of most read columns for the year, both of which appeared here in the Search & Social column, alongside several other worthy reads...

<http://searchengineland.com/the-hottest-reads-in-search-social-for-2012-143975>

-> 10 Predictions for Inbound Marketing in 2013

As is tradition here at Moz, I'm conducting my annual analysis of my predictions from 2012, and if I score high enough, predicting what will happen in 2013. I like to use this process because it keeps me honest - if I suck at predicting what will happen in a 12-month span, should you really listen to me for the next 12 months?...

<http://www.seomoz.org/blog/10-predictions-for-inbound-marketing-in-2013>

-> Whole Lotta Link Love In 2012: Our Best Link Building Posts Of The Year

Link building certainly was a hot topic in 2012 - maybe it's because Penguins are so gosh darn cute that we just can't help ourselves, right?...

<http://searchengineland.com/whole-lotta-link-love-in-2012-our-best-link-building-posts-of-the-year-143875>

-> Top Search Marketing Columns & Contributed Features Of 2012

As an industry news publication, the majority of our visitor traffic comes to our site to read daily news updates, in-depth coverage of search engine technologies and ever-evolving search engine algorithms...

<http://searchengineland.com/top-search-marketing-columns-contributed-features-of-2012-143883>

-> Gnome Likes: Google Changes, SEO Predictions & Blog Highlights

In last week's recap, we highlighted a few of our favorite discussions in the world of SEO. This week, we're back to summarizing our favorite posts... with a slight twist. Since the end of 2012 is rapidly approaching, this week's recap focuses on posts that review 2012 and make predictions about 2013...

<http://www.webgnomes.org/blog/gnome-likes-google-changes-seo-predictions-blog-highlights/>

-> 5 Trends Shaping SEO & Content Marketing in 2013

The Mayan calendar ended this month. Contrary to the "predictions," it wasn't the end of the world, just the end of Mayan content. To avoid having your company come up short on content, let's take a quick look at the history of web content and trends for 2013...

<http://searchenginewatch.com/article/2232360/5-Trends-Shaping-SEO-Content-Marketing-in-2013>

-> Search & Analytics Year In Review

As we head into the holidays, we're looking back at our own metrics for top posts of the year, and we wanted to share a recap of the most read stories in several categories on Search Engine Land...

<http://searchengineland.com/search-analytics-year-in-review-143278>

-> International Search Predictions For 2013 – Perhaps More Accurate Than The Mayans?
As I write this piece, the Mayan calendar is about to end bringing about the demise of everything. So, if you're reading this, we survived and the predictions were wrong...

<http://searchengineland.com/international-search-predictions-for-2013-perhaps-more-accurate-than-the-mayans-143236>

-> Search Engine Land's Most Popular Stories On Google+ In 2012: Panda, Knowledge Graph & SEO

It's started, the top lists of 2012! Over the next few days, we'll be posting a series of these, including the most read Search Engine Land stories, the most tweeted ones and the ones most liked on Facebook...

<http://searchengineland.com/most-popular-stories-on-google-143594>

-> How To Build A Better B2B SEO Strategy In 2013

It's the last week of 2012. If you're reading this column, you are forsaking at least a bit of much needed rest for some serious B2B SEO planning in 2013...

<http://searchengineland.com/how-to-build-a-better-b2b-seo-strategy-in-2013-142884>

-> Five Linking Myths That Need To Go Away In 2013

Goodbye 2012, hello 2013. For my last column of the year, I'm selecting five link building myths that I hope go away completely in 2013, and giving my rationale as to why they should be gone...

<http://searchengineland.com/five-linking-myths-that-need-to-go-away-in-2013-143394>

-> 2012 Analytics Year in Review & Predictions for 2013

Around this same time last year, we were discussing the inevitable increase and appreciation of data-driven decision-making with respect to web analytics insights and back-office data insights...

<http://searchenginewatch.com/article/2232475/2012-Analytics-Year-in-Review-Predictions-for-2013>

-> Critical Components to SEO Success in 2013

2012 has been an exciting year for search marketers. Google released algorithm updates at an unprecedented rate. New concepts like "AuthorRank," "Link Disavow Tool," and "Social Signals vs. Backlinks" have challenged SEO experts to review current strategies...

<http://searchenginewatch.com/article/2232891/Critical-Components-to-SEO-Success-in-2013>

-> The Good, The Bad & The Indifferent: Reviewing SEO results in 2012

With the end of the year upon us, I wanted to review my team's 2012 performance to inform our approach for 2013 and set ourselves some targets. As we've been preaching to clients that good content is honest content, I thought why not post-mortem our performance in public and let the data serve as a vehicle to describe how we'll be approaching SEO in 2013?...

<http://searchenginewatch.com/article/2232916/The-Good-The-Bad-The-Indifferent-Reviewing-SEO-results-in-2012>

-> 13 New Year's Resolutions Every Marketer Should Make in 2013

The new year is the perfect time to kick your professional butt into gear. Whether you are a devoted resolution-maker or just want to know how you can be a better marketer, here are 13 changes every marketing professional should make in 2013...

<http://searchenginewatch.com/article/2232888/13-New-Years-Resolutions-Every-Marketer-Should-Make-in-2013>

-> Which Top SEO Tactics Will You Focus On In 2013?

SEO tactics have morphed again over the past year, including the many Google updates we've seen...

<http://searchengineland.com/top-seo-tactics-for-2013-141993>

Bonne lecture !!

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

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Revue d'URL en français... ■ ■

-> Gagnez votre outil de scrap made in France

Bonjour à tous, aujourd'hui je vous propose de repartir avec une jolie petite licence d'un outil que certains d'entre vous connaissent déjà suite au VLC2012 et un article paru ici même il y a quelques semaines...

<http://www.visibilite-referencement.fr/blog/outil-de-scrap>

-> Les en-têtes HTTP ou comment maîtriser son indexation sur Google

Voici un petit article sur l'en-tête X-Robot-Tag et l'en-tête Link rel=canonical.

X-Robot-Tag pour simplifier, c'est l'équivalent de la meta <meta name= »robots »> (que vous pouvez placer dans le code source de vos pages HTML) mais dans les en-têtes HTTP, de même pour la Canonical...

<http://www.yapasdequoi.com/apache/2983-len-tete-x-robot-tag-ou-comment-vite-desindexer-des-pages.html>

-> SEO : 6 suites logicielles "tout en un"

Analyse de backlinks, suivi de positionnement, audit des éléments techniques, suggestion et analyse de mots clés... Certaines solutions dédiées aux SEO sont de véritables plateformes "tout en un". En voici 6...

<http://www.journaldunet.com/solutions/seo-referencement/suite-seo/>

-> Les indicateurs clés de performance pour le SEO

Puisque le référencement naturel (SEO pour les intimes) est devenu un véritable investissement pour les entreprises (en temps aussi bien qu'en argent), il devient évident de pouvoir créer des rapports de données efficaces...

<http://www.webmarketing-com.com/2012/12/17/18103-indicateurs-performance-seo>

-> Le Knowledge Graph : dommages collatéraux dans les résultats Google ?

Suite à la mise en place du Knowledge Graph (« graphe du savoir ») en France la semaine dernière et après quelques constatations sur la manière d'en tirer partie grâce à Google Images, nous avons étudié plus spécifiquement quelques groupes de mots clés afin déterminer les sites qui pouvaient potentiellement être affectés en terme de visibilité dans les résultats naturels avec l'arrivée du Knowledge Graph (KG)...

<http://blog.resoneo.com/2012/12/le-knowledge-graph-dommages-collateraux-dans-les-resultats-google/>

-> Pourquoi les citations sont importantes pour le référencement local de votre entreprise

Les Citations peuvent être définies comme des « mentions » de votre entreprise et de votre adresse sur d'autres pages Web, même s'il n'y a pas de liens vers votre site. Un exemple de citation peut se faire via le répertoire d'entreprise en ligne des pages jaunes où votre entreprise est inscrite, mais n'est pas relié directement...

<http://www.vaisonet.com/blog/referencement/pourquoi-citations-sont-importantes-referencement-local-votre-entreprise.html>

-> 8 stratégies pour tirer le maximum possible de Pinterest

Vous considérez-vous comme étant « toxicomane » par rapport à la plateforme Pinterest ? Plusieurs d'entre vous répondront non à cette question simplement parce qu'ils sont du sexe masculin...

<http://www.webmarketing-com.com/2012/12/18/17984-strategies-pinterest>

-> Ce que Facebook doit envier à Google+ Communities

Avec l'arrivée des Communautés Google+, Google nous a donné des espaces où nous pouvons discuter, partager et découvrir pleins de choses pour des sujets qui nous passionnent...

<http://www.derfranzose.ch/ce-que-facebook-doit-envier-a-google-communities/>

-> Google recrute en France pour améliorer son Knowledge Graph
Google est en quête d'un ou plusieurs contributeurs pour l'aider à améliorer son Knowledge Graph. Un service de recherche enrichie que le moteur vient de déployer pour les résultats français...

<http://www.journaldunet.com/solutions/emploi-rh/recrutement-google-en-france-pour-le-knowledge-graph-1212.shtml>

-> Domptez les Sitelinks de Google en toute simplicité

Lorsqu'un site arrive en 1ère place des résultats de recherche (par exemple sur une requête exacte du nom de domaine), il arrive que des liens additionnels, les « sitelinks » s'affichent sous le lien principal...

<http://www.webmarketing-com.com/2012/12/21/18254-domptez-les-sitelinks-de-google-en-toute-simplicité>

-> Le musée océanographique de Monaco accueille le simulateur Liquid Galaxy de Google
Le Liquid Galaxy permet d'accéder au contenu Google Earth propre aux océans, dans un environnement qui met le visiteur en position d'explorer intuitivement les océans. L'installation du Musée est la seule en Europe à proposer ce contenu...

<http://cote-d-azur.france3.fr/2012/12/22/le-musee-oceanographique-de-monaco-accueille-le-simulateur-liquid-galaxy-de-google-169527.html>

-> Les ancrés dans vos pages

Dans la gestion de nos sites, de nos articles, il y a tout ce qui est contenu, catégories, articles. Il faut cependant ne pas oublier de mettre des liens dans vos articles, des liens vers les catégories et des liens vers vos articles les plus importants ou des liens vers des articles un peu oubliés ou anciens...

<http://www.redigeons.com/?p=607>

-> L'efficacité des annuaires en référencement naturel (SEO)

Les annuaires sont-ils efficaces en référencement naturel ? D'abord, il faudrait se poser cette question : est-ce que les annuaires sont toujours utiles en référencement naturel ? Devez-vous faire confiance aux annuaires web ?...

<http://www.ya-graphic.com/2012/12/efficacite-des-annuaires-en-referencement-naturel/>

-> Faire un audit de site internet, un audit référencement

Lorsque l'on fait un audit SEO, il y a plusieurs types de rapport SEO, du plus succinct au rapport très complet. Il y a aussi des outils, des sites à utiliser, un audit SEO c'est un mélange de tout cela...

<http://www.webmarketing-seo.fr/?p=282>

-> 18 mois dans la peau d'un référenceur

Toi aussi, vis ma vie de référenceur à travers un article détaillant les plus importantes mises à jour de Google lors de ces 18 derniers mois...

<http://oseox.fr/blog/index.php/1026-vie-referencieur>

-> Tutoriel Gestionnaire de balises Google

Google Tag manager est un service gratuit offert depuis quelques semaines qui aide bien pour faciliter son suivi dans google analytics...

<http://www.mauricelargeron.com/guide-pour-demarrer-avec-le-google-tag-manager/>

-> Le SEO n'est pas mort avec Penguin, mais il faut le faire évoluer

L'année 2012 du référencement a été chargée, en particulier du fait du lancement de l'algorithme Google Penguin. Suite à cette mise à jour, beaucoup ont même annoncé la mort du SEO. Voici pourquoi ces personnes ont tort et quels changements nous réserve 2013...

<http://www.journaldunet.com/solutions/expert/53081/le-seo-n-est-pas-mort-avec-penguin--mais-il-faut-le-faire-evoluer.shtml>

-> 4 idées de Link Building pour cette fin d'année

Le Link Building est un art, essentiellement pour ceux qui connaissent l'importance des liens entrants pour leurs blogs ou sites web, sous l'objectif d'améliorer le positionnement dans les moteurs de recherche...

<http://blog.maroketing.com/idees-linkbuilding-intelligent/>

-> Humour : le bêtisier 2012 du site WebRankInfo

Comme chaque année désormais, voici le bêtisier WebRankInfo version 2012, issu des mails les plus amusants ou insolites que j'ai reçus pendant l'année 2012...

<http://www.webrankinfo.com/dossiers/webrankinfo/betisier-2012>

-> Audit référencement du site de l'Élysée

Le 18 décembre dernier, le site de l'Élysée montrait son nouveau visage : un site moderne, avec un beau design, 100% Made in France, peu coûteux, résolument tourné vers le social, etc....

<http://noviseo.fr/2013/01/audit-referencement-du-site-de-lelysee/>

-> Analyse du crawl de Googlebot avec la Watussi Box

Je tenais à vous présenter une application open-source développé par Jean-Benoît Moingt aka Mr Watussi (très bon blog SEO orienté technique/PHP/logs/Apache) qui va vous permettre de suivre le crawl de GoogleBot sur vos sites, via une interface Web et sans avoir à aller trifouiller les logs en ligne de commande...

<http://www.yapasdequoi.com/scripts/3003-analyse-du-crawl-de-googlebot-avec-la-watussi-box.html>

-> SEO : 5 conseils pour optimiser vos articles de blog sous WordPress

Il est très difficile d'obtenir de fervents lecteurs sur son blog. En effet, il existe déjà des millions de blogs sur le web. Alors pourquoi un internaute irait sur le vôtre plutôt que sur un autre parmi la masse disponible ?...

<http://www.ecommerce-webmarketing.com/2012/12/31/seo-5-conseils-pour-optimiser-vos-articles-de-blog-sous-wordpress-10285178/>

-> Comment rendre vos recherches Google plus efficaces

Voici quelques conseils qui vous permettront de gagner du temps et d'arriver plus précisément à l'objet de votre recherche...

<http://www.atlantico.fr/decryptage/comment-rendre-vos-recherches-google-plus-efficaces-582456.html>

-> Pour sauver la presse, faire payer Google : un avis de référenceur

Cela fait maintenant des semaines que la campagne de lobbying des grands acteurs de la presse française bat son plein pour faire payer à Google une taxe sur l'indexation des liens dans Google Actualités, et par ricochet dans les blended SERP's de la recherche universelle...

<http://www.cafe-referencement.com/actu-seo/pour-sauver-la-presse-faire-payer-google-un-avis-de-referenceur-515>

-> Tutoriel sur le fichier sitemap pour le référencement

Voici tout ce qu'il faut savoir sur le fameux fichier sitemap et son impact sur le référencement. Ce tuto explique les risques d'une mauvaise compréhension de l'utilité du sitemap et donne de nombreux conseils...

<http://www.webrankinfo.com/dossiers/sitemaps/tutoriel-fichier-sitemap>

-> Google fait tout pour qu'on fasse des recherches connectés à notre compte Google

Google a publié une offre d'emploi pour un responsable marketing chargé d'augmenter le nombre d'utilisateurs connectés à leur compte Google quand ils font des recherches. Cet article explore les conséquences de cette démarche pour le référencement de vos sites...

<http://www.webrankinfo.com/dossiers/google-search/connectes-compte-google>

-> Guest blogging : 5 conseils pour en tirer parti

Le guest blogging a de nombreux avantages que cela soit en tant que blogueur invité ou propriétaire de blog...

<http://www.webmarketing-com.com/2013/01/09/17689-guest-blogging-5-conseils-pour-en-tirer-parti>

-> Nettoyer, balayer, optimiser votre fichier CSS pour le SEO

Bonjour à tous, aujourd'hui un petit billet qui va faire un mix entre bonnes pratiques et outils. Vous le savez probablement, les performances d'un site n'ont jamais été aussi

importantes tant dans le processus de référencement que dans celui de l'expérience utilisateur et ce qui en découle...

<http://www.visibilite-referencement.fr/blog/fichier-css>

-> Linking interne : les stratégies performantes

Si les liens entrants sont très importants dans le référencement naturel, l'organisation du maillage interne de votre site est le prolongement de vos backlinks. Un mauvais linking redistribuera mal l'apport en « link juice » venant de l'extérieur, comme le PageRank...

<http://www.htitipi.com/blog/linking-interne.html>

-> Comment sortir d'une pénalité sans trop de bobos

Vous possédez un site e-commerce et vous (ou votre prestataire) avez abusé des techniques de link building « border line » (en connaissance de cause ou non), ou alors vous avez (trop ?) basé votre stratégie de référencement sur le link building et malheureusement pour vous, Panda dans un premier temps, puis Pingouin (et possiblement d'autres pénalités) sont venu handicaper vos site jusqu'au point de mort cérébrale ?...

<http://lemusclereferencement.com/2012/07/25/comment-sortir-dune-penalite-sans-trop-de-bobos/>

-> Podcast SEO Vol. II Ep. 1 : Francois Goube – Majestic SEO

L'année 2013 marque le début du Volume II du podcast référencement. Pour le 1er épisode de cette nouvelle édition, j'ai l'honneur d'inviter François Goube, ambassadeur en France de Majestic SEO...

<http://www.laurentbourrelly.com/blog/1325.php>

Revue d'URL en anglais...

-> Building Google's Knowledge Base and Identifying Locations in Web Pages

When we talk about indexing and crawling content on the Web, it's usually within the context of pages being ranked on the basis of a number of signals found on Web pages that might be ranked in response to queries...

<http://www.seobythesea.com/2013/01/knowledge-base-locations-in-web-pages/>

-> Semantic Web and Link Building without Links > The Future for SEO?

Rand's recent WBF about co occurrence was a real wake up call for those still transfixed with link building practices of old. While anchor text based links may still have some effect there is little arguing the fact that the factor's importance is dwindling...

<http://www.seomoz.org/blog/semantic-web-and-link-building-without-links-the-future-for-seo>

-> Secrets to Writing Engaging Titles & Content for SEO

SEW recently revealed the Most Engaging Posts of 2012. As a shameless plug, my article "10 Old SEO Methods You Need to Stop" ended up being among them and I proudly Tweeted and bothered my co-workers with the news and the article gained another round of attention...

<http://searchenginewatch.com/article/2235149/Secrets-to-Writing-Engaging-Titles-Content-for-SEO>

-> How to Craft Phenomenal SEO Landing Pages That Rank & Convert

When it comes to landing pages, most of the talk, tips, and advice revolves around creating landing pages for paid search...

<http://searchenginewatch.com/article/2234402/How-to-Craft-Phenomenal-SEO-Landing-Pages-That-Rank-Convert>

-> Visualizing Duplicate Web Pages

We've just changed the way we detect duplicate or near-duplicate web pages in our custom crawler to better serve you. Our previous code produced good results, but it could fall apart on large crawls (ones larger than about 85,000 pages), and takes an excessively long time (sometimes on the order of weeks) to finish...

<http://www.seomoz.org/blog/visualizing-duplicate-web-pages>

-> Why You Need To Complement Your In-House Team With External Expertise

I am a big fan of large enterprise organizations having an in-house SEO team. However, there are strong arguments for supplementing the efforts of your in-house team with help from external SEO expertise. Below are a two major reasons why...

<http://searchengineland.com/complement-you-in-house-team-with-external-expertise-143352>

-> Does Google Search Google? How Google May Create and Use Synthetic Queries

When someone searches at Google, their query might not express the informational or situational need that they have. It might be too broad, too ambiguous, or vague in some other manner...

<http://www.seobythesea.com/2013/01/google-synthetic-queries/>

-> Local SEO Citations – The New Link Building

Local SEO citations and mobile madness have hit the T-minus mark and are ready for launch. The previous two years have just been test flights...

<http://searchenginewatch.com/article/2234086/Local-SEO-Citations-The-New-Link-Building>

-> Learn About Robots.txt with Interactive Examples

One of the things that excites me most about the development of the web is the growth in learning resources...

<http://www.seomoz.org/blog/interactive-guide-to-robots-txt>

-> How Google Might Ignore Insignificant Terms in Queries

The most important step in doing keyword research is entering a keyword phrase into a search engine like Google, and seeing what results show up, and trying to understand why the pages that appear within results are there. If you can't do that, then it's time to dig down and start learning...

<http://www.seobythesea.com/2013/01/insignificant-terms-in-queries/>

-> 10 Reasons Why Public Relations is a 'Must-do' for SEO in 2013

Many people today recognize that search engine optimization (SEO) and public relations (PR) have much to offer each other...

<http://searchenginewatch.com/article/2232614/10-Reasons-Why-Public-Relations-is-a-Must-do-for-SEO-in-2013>

-> How Link Building Really Works These Days

I do not write a lot about link building. When I do I try to expand the concept usually or at least question it. The problem with link building is that it doesn't go away. No matter how often I and other people say that you have to get links, earn links or at least create link bait...

<http://seo2.0.onreact.com/how-link-building-really-works-these-days>

-> Do You Believe in These 10 SEO Myths?

Given the number of different sources online, determining what constitutes legitimate information on the practice of search engine optimization and which assertions are either factually incorrect or maliciously wrong can be incredibly challenging...

<http://www.searchenginejournal.com/do-you-believe-in-these-10-seo-myths/56807/>

-> An Updated Guide to Google Webmaster Tools

Google Webmaster Tools was once Google Webmaster Wasteland. But the past year has been a fruitful one as Webmaster Tools has rolled out improvements faster than Facebook does new privacy statements...

<http://www.seomoz.org/blog/an-updated-guide-to-google-webmaster-tools>

-> An Open Letter to Google

While I need to talk to you about a few issues, I believe it's only fair to begin this letter with a genuine thank you. For years you have provided a very solid set of search results and your market share reflects a true talent for understanding what people think they want, how they think they want it, and serving it to them quickly...

<http://searchenginewatch.com/article/2232358/An-Open-Letter-to-Google>

-> How to Create Websites Optimized for Mobile

Mobile Internet is on the rise all over the world. Many commentators have predicted that mobile search will overtake desktop by 2015 and in some countries, including China and India, this is already the case...

<http://searchenginewatch.com/article/2232894/How-to-Create-Websites-Optimized-for-Mobile>

-> 7 Secrets For Dealing With Quality Score

I'm back to writing my own controversial thoughts this month; although, if you didn't check out the interview I did last month with Andrew Goodman and Matt van Wagner, you should, those are smart guys...

<http://searchengineland.com/7-secrets-for-dealing-with-quality-score-142224>

-> Make News, Win Links

No longer can we get away with creating content. It has to be newsworthy content. It has to be unique (or you'll find Twitter abuzz with people fussing about how sick they are of seeing x types of posts) and it has to be crafted with your audience and reach in mind...

<http://searchenginewatch.com/article/2232319/Make-News-Win-Links>

-> A Simple Guide to Using rel="alternate" hreflang="x"

The world is shrinking and with it the need to support content for more users means supporting more languages...

<http://searchenginewatch.com/article/2232347/A-Simple-Guide-to-Using-relalternate-hreflangx>

-> Insiders Guide To Selecting The Right Local SEO Tools

Last week's announcement by Raven tools that they would remove all 'scraped data' (e.g., ranking reports) from their toolkit came as an unwelcome surprise to many of their customers. This action effectively makes a big portion of their service redundant, and the reaction to this news hasn't been all positive for Raven...

<http://searchengineland.com/insiders-guide-to-selecting-the-right-local-seo-tools-142570>

-> Forget Rankings: Here Are 6 SEO Metrics That Matter

It would be an understatement to say that the SEO landscape has changed a lot through the years. In just the last two years the changes have been dramatic...

<http://searchenginewatch.com/article/2231719/Forget-Rankings-Here-Are-6-SEO-Metrics-That-Matter>

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

Merci pour votre lecture... Pour toute suggestion : olivier@abondance.com