

L'évolution du Web Analytics et ses conséquences sur le SEO

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

Plus le temps passe et plus on s'aperçoit que le SEO ne peut pas - et ne pourra pas - se passer des outils de mesure d'audience, domaine dans lequel Google Analytics, un outil simple et puissant, se taille la part du lion. Mais cet outil présente également quelques limites et certaines annonces récentes ont également mis en lumière quelques évolutions qui nous font nous poser quelques questions quant à son utilisation dans une stratégie de référencement naturel. Petite revue d'effectif...

Le Web analytics fait partie des outils les plus utilisés par les webmasters ou chargés de communication soucieux de connaître les performances des pages web. Or, certaines évolutions risquent de rendre cet outil de plus en plus difficile à utiliser.

Cryptage des URL de recherche Google en HTTPS

Après avoir crypté en SSL les données transitant dans Gmail en janvier 2010 (<http://gmailblog.blogspot.com/2010/01/default-https-access-for-gmail.html>), puis créé au mois de mai 2010 une version sécurisée de son moteur de recherche, accessible à l'adresse <https://encrypted.google.com/>, Google a frappé fort cet automne (<http://googleblog.blogspot.com/2011/10/making-search-more-secure.html>), en mettant en place le cryptage des données pour tous les utilisateurs connectés à leur compte Google.

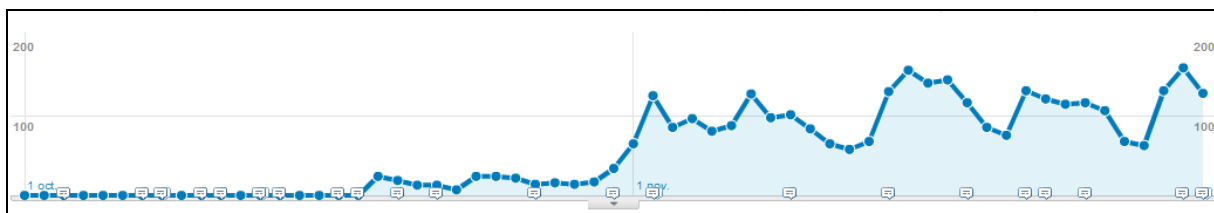
Lorsque l'on est logué à son compte, les URL de recherche Google commencent donc maintenant par un HTTPS et les outils de tracking sont dans l'incapacité de connaître les informations échangées (et notamment les requêtes des internautes). Les outils seront capables de voir les visites mais sans connaître le *referrer*.

Google avait affirmé (<http://searchengineland.com/google-to-begin-encrypting-searches-outbound-clicks-by-default-97435>) que ces données cryptées représenteraient un faible pourcentage des données Analytics. Or, après plusieurs études, il s'est avéré après coup que le pourcentage de requêtes masquées s'élevait à 12%, voire plus (<http://actu.abondance.com/2011/11/google-plus-de-requetes-masquees-par-la.html>). Une étude de SEOMOZ faisait ainsi état d'un site où plus de 50% des requêtes étaient masquées !

Pour connaître la portion de visites masquées sur un site web, il faut se reporter aux visites en référencement naturel, dont la requête est « *not provided* ».

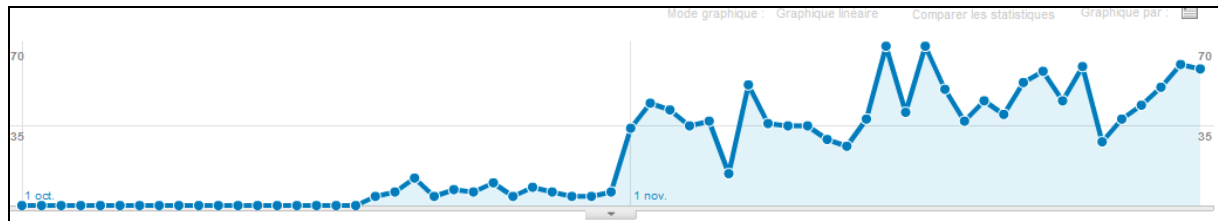
Le pourcentage de requêtes inconnues ou masqué a augmenté sans ambiguïté au cours des derniers mois, et particulièrement depuis la fin octobre.

Voici un premier exemple sur le site allopneus.com. Le nombre de visites « not provided » a nettement augmenté : il est passé de 240 visites en octobre à plus de 3000 visites en novembre !



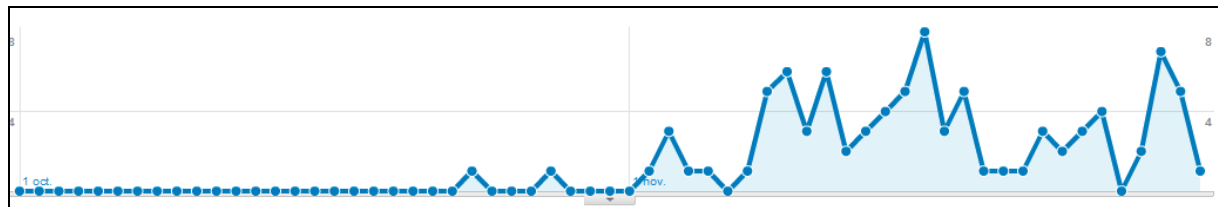
Ce pourcentage est cependant noyé dans la masse et ne représentent même pas 1 % des visites sur le site au mois de novembre.

Autre exemple, sur partenaire-europeen.fr, le nombre de visites « not provided » par mois est passé de 80 à plus de 1300 entre octobre et novembre. Le nombre de visites masquées ne représente pas tout à fait 1% des visites en novembre.



L'impact du cryptage des URL est donc très limité, du moins pour les comptes que nous avons analysés. Il est clair que cela dépend fortement de l'activité du site et du pourcentage d'internautes qui utilisent un compte Google et surtout qui utilisent Google.com (le chiffrement n'est pas encore en place sur Google.fr). Par exemple sur un site d'actualité web, on peut penser que le pourcentage de visites masquées est plus important car les internautes utilisent davantage les services Google.

Nous avons ainsi constaté sur notre site brioude-internet.fr, qui propose régulièrement des actualités sur le référencement, une évolution beaucoup plus importante du nombre de visites masquées : au mois de novembre cela représentait un peu plus de 4% des visites en référencement naturel.



Quoiqu'il en soit, le nombre d'internautes possédant un compte Google et naviguant en étant connecté, ne peut qu'augmenter au fil du temps. En effet les efforts de communication de Google et la mise à disposition de services toujours plus diversifiés tendant à attirer les internautes. Par exemple le pourcentage de Français ayant rejoint le réseau Google+ s'élève à 12%, ce qui n'est pas négligeable (source : <http://www.journaldunet.com/ebusiness/le-net/les-reseaux-sociaux-preferes-des-francais/appartenance.shtm>).

Il est capital pour un site web de connaître avec précision les requêtes qui génèrent du trafic, ainsi que les pages visités, le comportement des internautes sur le site, l'équipement des internautes... Toutes ces informations permettent de déterminer ce qui fonctionne sur un site web et de détecter les pages nécessitant de l'optimisation et du webmastering. Le fait que des données sur les visiteurs soient de plus en plus masquées par Google n'est pas de nature à rassurer les analystes !

On s'est beaucoup posé de questions sur l'offre Analytics Premium qui allait être proposée aux grandes entreprises pour un prix « modique » de 150 000 \$ par an. Est-ce que cette offre payante permettra de connaître toutes les données, y compris celles qui sont masquées en HTTPS ? Le mystère demeure sur les détails de cette offre.

Arrivée du multi-levier dans Google Analytics

Il y a bien longtemps que la navigation web ne se limite plus aux moteurs de recherche. Les portails thématiques (actualité, vidéo, images) et les réseaux communautaires ont changé la donne et le comportement des internautes a évolué : désormais nous passons beaucoup plus de temps à naviguer d'un site à l'autre et à « zapper » en allant voir par exemple les offres de produits proposées par plusieurs sites concurrents ainsi que les différents tests présents sur le web.

Google a longtemps été critiqué car il ne tenait compte que du dernier site visité pour afficher les données de conversion dans Analytics, ce qui revenait à diriger un projecteur sur Adwords. En lançant le principe des entonnoirs multicanaux (<https://www.google.com/support/analytics/help/bin/answer.py?answer=1191180>) dans Analytics au mois d'août, Google souhaitait lever le voile sur le parcours des internautes avant la page de conversion. On peut désormais connaître les différents leviers de conversion et découvrir le trafic issu des moteurs de recherche mais aussi issu des sites référents, des réseaux sociaux, des newsletters, des annonces publicitaires, des comparateurs de prix... Chacun des parcours est détaillé dans un scénario Google.

Chaque chemin d'accès peut être connu en affichant le nom de la source ou du support dans les variables secondaires du rapport Analytics. On peut ainsi déceler quel site référent ou régie publicitaire envoie le plus de visites, ce qui est d'un intérêt précieux pour le référencement, qu'il soit naturel ou basé sur des liens publicitaires.

	Chemin de groupe de canaux de base	Chemin de la source/du support	Conversions
1.	Publicité payante → Accès directs	Leguide / cpc (direct) / (none)	8
2.	Publicité payante → Publicité payante	Leguide / cpc → Leguide / cpc	6
3.	Publicité payante → Publicité payante	Shopping / cpc → google / cpc	6
4.	Résultats de recherche naturels → Résultats de recherche naturels	google / organic → google / organic	6
5.	Résultats de recherche naturels → Résultats de recherche naturels → Résultats de recherche naturels	google / organic → google / organic → google / organic	6
6.	Accès directs → Accès directs → Accès directs → Accès directs	(direct) / (none) → (direct) / (none) → (direct) / (none) (direct) / (none)	4
7.	Accès directs → Résultats de recherche naturels	(direct) / (none) → google / organic	4

Ce type de données Analytics permet d'élargir son champ de vision et de connaître la vie d'un site dans son environnement web. Il est ainsi possible de connaître l'univers qui gravite autour d'un site marchand et qui lui apporte du trafic.

Quel impact pour le SEO ? On pourra désormais arbitrer entre les différentes solutions proposées sur le web et tirer ou non certains leviers en fonction de leurs performances. Selon les tendances relevées par le multi-leviers, on pourra faire davantage de référencement naturel sur son site, ou plutôt privilégier les partenariats et les accès depuis des sites tiers (comparateurs de prix, blogs, réseaux sociaux...).

Il est de plus évident que le référencement naturel n'est qu'un moyen parmi d'autres de parvenir une bonne visibilité et à un bon taux de conversion. Grâce au multi-leviers, on pourra définir la bonne stratégie à adapter.

Projet de loi sur les cookies opt-in

Au mois de février, le gouvernement votait un projet de loi sur les cookies publicitaires, pour mettre la France aux normes européennes en matière de communication. L'objectif est que l'internaute puisse refuser le stockage des cookies sur son ordinateur (opt-in) au lieu de les accepter par défaut (opt-out). Les cookies indispensables à la navigation (identifiants et mots de passe ou éléments d'un panier d'achat) ne seront pas concernés par ces restrictions.

Le « paquet télécom » est destiné à protéger davantage la vie privée des internautes ; selon les recommandations de la CNIL, les internautes devront être informés de façon claire sur l'installation de cookies, par exemple par le biais d'un message informatif ou d'un formulaire comprenant une case à cocher. L'utilisation des réglages du navigateur permettant d'accepter ou refuser n'est pas considérée comme une méthode suffisamment transparente.

Si la loi est votée, les sites marchands, qui se servent de cookies pour mémoriser les préférences des internautes vont être privés d'une source importante d'informations. Connaître les derniers produits achetés, les sommes dépensés, les produits consultés... permet de présenter des pages personnalisées aux internautes et des produits susceptibles de les intéresser. Les cookies comportementaux, qui sont dans la ligne de mire du Gouvernement et de la CNIL, seront soumis d'une part à l'affichage d'information claire pour les internautes, d'autre part à un consentement de l'internaute.

Difficile dans ces conditions de satisfaire aux exigences légales sans faire fuir les visiteurs ! Or il est plutôt stratégique de pouvoir proposer des produits ciblés en fonction du profil des acheteurs, de façon à améliorer la mise au panier. Ceci n'est d'ailleurs pas sans intérêt pour les internautes eux-mêmes, qui verront les dernières nouveautés et des actualités soigneusement sélectionnées s'afficher à l'écran.

L'analyse du comportement des internautes risque donc d'être de plus en plus difficile pour les sites marchands, ce qui ne sera pas sans impact sur le SEO : quels produits mettre en avant ? Comment fidéliser les internautes ? Ces problématiques devront amener la mise en œuvre de nouvelles solutions (utilisation des réseaux communautaires par exemple)

Conclusion

Les possibilités offertes par le Web analytics évoluent en bien et en mal. D'une part le cryptage des URL de recherche Google et l'obligation d'acceptation des cookies vont diminuer les sources d'information mais d'autre part le multi-leviers et la connaissance des chemins de conversion apporte des données de plus en plus pertinentes.

Ce qui semble se dessiner, c'est un environnement analytics qui protège de plus en plus l'internaute et respecte mieux la vie privée : le référencement devra donc s'adapter à cette nouvelle donne et utiliser d'autres moyens pour connaître les préférences de l'internaute (sondages, forums, réseaux sociaux...)

Guillaume Thavaud, cellule de veille de la société *Brioude-Internet Référencement*
(<http://www.brioude-internet.fr>).

Réagissez à cet article sur le blog des abonnés d'Abondance :
<http://blog-abonnes.abondance.com/2011/12/levolution-du-web-analytics-et-ses.html>