

## Ecrire pour les moteurs de recherche (3ème partie) : le contenu

[Retour au sommaire de la lettre](#)

*Si l'optimisation des pages web pour les moteurs de recherche passe par une meilleure définition des urls, du code HTML, des liens, etc., elle doit avant tout être basée sur l'optimisation du contenu lui-même, depuis le titre (balise TITLE) jusqu'au texte qui le compose. Dans cette série de trois articles, David Cohen (de la société de référencement CVFM) nous explique comment optimiser un contenu rédactionnel afin de le rendre plus visible par les moteurs de recherche... Après le titre des pages, puis le choix des mots clés et le phénomène de "longue traîne" le mois dernier, nous terminons avec le contenu proprement dit, grâce à une stratégie en 5 "mots clés" : nommage, quantité, diversité, fraîcheur et hiérarchie de l'information...*

Après un premier article qui mettait l'accent sur l'importance du couple titre/title, après un deuxième qui couvrait le choix des mots clés et la découverte de la "longue traîne", nous voici pour ce dernier article dans la phase "projet et rédaction de site", devant le contenu proprement dit, avec l'ambition d'y insérer au mieux des mots pour les moteurs.

Cinq principes directeurs vont parsemer notre stratégie :

- Nommage.
- Quantité.
- Diversité.
- Fraîcheur.
- Hiérarchie de l'information.

Listons-les maintenant les uns après les autres...

### **1er principe : le nommage**

La conception d'un site commence par l'action de lister les mots les plus pertinents, pour en nommer tous les éléments. Nommer, c'est le début de l'écriture. Et ce n'est pas parce que le site est dynamique qu'il peut se priver de cet avantage...

L'ensemble des éléments du site doivent être nommés selon le secteur concerné avec les techniques de recherche de mots décrites dans l'article précédent.

Les moteurs "lisent" tous les éléments constitutifs du site. En fait, pensez qu'ils "voient" d'abord votre serveur FTP et listent les dossiers. S'ils y retrouvent partout des mots qui concernent la même activité (principe de thésaurus), cela leur permet de décider que le site est pertinent pour tout un secteur d'activité. Chaque mot du domaine renforce la pertinence des autres, surtout pour Google.

Alors le nommage commence par le premier élément : le nom de domaine, sachant qu'un mot-clé placé dans le domaine ou le sous-domaine lève parfois bien des ambiguïtés. Prenez la marque "Lancaster". Seule dans le nom de domaine, elle peut désigner une des nombreuses villes éponymes, une marque d'articles de cuir réputée ou de produits cosmétiques qui ne l'est pas moins. Si vous ne précisez pas le nom de domaine avec un mot du métier (sacs), du sujet (beauté) ou du lieu (britain), vos concurrents sur les résultats de recherche seront trois fois plus nombreux.

Puis il faut nommer les dossiers et chaque fichier, qu'il s'agisse d'une page, d'une image, d'une feuille CSS, d'un script JavaScript... Tout est bon pour écrire un message aux moteurs avant même qu'ils n'indexent le premier mot dans un texte. Ce nommage vous donnera en plus une avance sur l'obligation de rendre les sites "accessibles" aux malvoyants (une B.A. d'avance donc).

Et ce message va bien sûr passer dans les URL (adresses) des pages. Ainsi après un nom de domaine signifiant, vos dossiers et fichiers introduisent des mots clés dans l'URL grâce à leurs noms.

### **2e principe : la quantité**

Plus un site contient de pages, plus il peut en placer dans les moteurs, et une page mal placée dans les résultats sur une expression est souvent bien placée sur 10 autres expressions dont vous n'aviez pas idée auparavant.

Vous n'avez pas toujours les 20 000 produits d'un site de e-commerce pour multiplier vos entrées dans les moteurs, mais vous pouvez :

- Spécialiser des pages. Exemple : au lieu de lister des communiqués de presse les uns après les autres, vous écrivez un communiqué par page;
- Multiplier les rubriques. Exemple : les nombreux blogues dont les auteurs créent une rubrique à chaque fois qu'ils changent de sujet;
- Donner un lexique de votre sujet ou votre métier avec une page par mot, bien sûr.

Bref, vous avez sûrement une idée de spécialiste de votre domaine sur la question maintenant.

### **3e principe : la diversité**

Plus les pages sont différentes par leur contenu, plus le moteur s'oblige à passer partout pour ne rien manquer.

Pour un moteur, un site de référence n'est pas axé sur un seul mot, une seule expression (voir la place laissée aux nombreux annuaires). Un site est une référence dans un secteur d'activité, de service ou de réflexion. Le site va donc parler de tous les aspects de ce secteur.

Cette diversité doit se retrouver au niveau du TITLE pour que les moteurs la voient immédiatement. Le SEO des années 90 pouvait se permettre de trouver un bon TITLE et le répéter sur toutes les pages du site (5 ou 6 mots-clés lui suffisaient et il avait tort). Le SEO de 2007 structurera son TITLE pour que chaque page apparaisse différemment.

Lorsqu'on pratique la Longue Traîne (voir article du mois précédent), la diversité des expressions devient un réflexe conditionné de rédacteur pour répondre à la diversité des requêtes. Plus votre vocabulaire sera riche, plus vous répondrez à différents niveaux socioculturels. Et "vocabulaire riche" ne se traduit pas par savant ou compliqué : il s'agit plutôt d'utiliser tous les synonymes d'un même mot, de le marier avec différents adjectifs ou verbes. Bref, de faire un travail de rédacteur professionnel dont le vocabulaire ne se limite pas aux 200 mots-clés qu'il gardera tout de même sous les yeux.

La diversité est d'ailleurs indispensable dans le cas où un mot serait ambigu. Ainsi, l'expression "studio paris" fait référence à l'immobilier, la photographie, la création ou même à "studio Pilates" (plus connu dans le monde anglophone)... C'est le contexte des autres mots qui permet à Google de cibler ses résultats sur l'immobilier, par exemple. La diversité consistera alors à utiliser aussi le mot "appartement" et l'expression "une seule pièce" pour accrocher d'autres requêtes que celles centrées sur le mot "studio", mais aussi et surtout pour bien cibler votre "studio" sur l'immobilier (si c'est votre cas).

Au passage, testez Live.com et Yahoo, qui dispersent leurs résultats sur plusieurs secteurs d'activité. C'est la preuve que Google utilise des thésaurus (pour se rapprocher du sens) et aussi la fréquence des clics. Dans le cas de "studio paris", Google privilégie l'immobilier et la photographie parce qu'il observe le nombre de clics sur les résultats, alors que MSN et Yahoo gardent toujours une pondération certaine sur la recherche d'une chaîne de caractère. Et c'est moins pertinent.

### **4e principe : la fraîcheur**

Plus vos pages sont fréquemment mises à jour, plus les moteurs s'obligent à passer souvent les "voir". Plus ils passent souvent, plus ils considèrent le site comme important.

Les sites de journaux, de news, sont indexés plusieurs fois par jour et, s'ils ne "râlent" pas, ils ont souvent la chance d'entrer sur Google News.

Faire la une de Google News est une expérience surprenante pour qui regarde les courbes d'audience. Faire la une de Google News signifie que votre information a été sélectionnée comme

répondant au mieux à une requête d'actualité. Des milliers de journaux, magazines et blogues l'ont par exemple tenté pour la sortie de l'Iphone, le mois dernier.

Mais entrer dans la sélection des sites de news n'est pas donné à tout le monde et certains refusent même d'y figurer... Tant mieux, ça laisse plus de la place pour les autres !

Même si votre site ne dispose pas d'une actualité quotidienne, vous avez certainement des nouvelles fraîches à mettre en ligne toutes les semaines, tous les mois. Vous n'y croyez pas? Vous n'avez donc pas de communiqué de presse, pas de nouveau produit, pas de nouveau témoignage d'un client ou d'un ami, pas de participation à un salon, à un séminaire... Prenons le cas d'un site d'hôtel (ambiance : une guerre permanente du référencement). *A priori*, l'hôtel a toujours les mêmes chambres, les mêmes tarifs et il ne change pas d'adresse... mais, s'il a un restaurant, il change de menu tous les jours et, hop, on met le menu en ligne; il change de clients tous les jours et, hop, on met le livre d'or en ligne. Au hasard d'une requête aussi fine qu'imprévisible, ce site d'hôtel apparaîtra bien placé sur Google grâce aux derniers commentaires élogieux d'un client sur sa baignoire spa ou son écran TV avec webcam incorporée...

Vous cherchez d'autres nouvelles fraîches? Votre ville vous livre chaque mois son lot d'animations, de spectacles, de manifestations, votre société remporte de nouveaux marchés, vos nouveaux stagiaires ont fait des voyages enrichissants pour leur CV, vous avez déposé un nouveau brevet... Certains rédacteurs de blogues s'obligent à écrire tous les jours pour être bien "vus", ce qui les pousse parfois à parler de la petite dépression de leur cactus (dame, la vie n'est pas toujours intense !).

Nous sommes sûrs que vous ne descendrez pas aussi bas dans la banalité, que vous resterez centrés sur votre sujet, votre métier et leurs actualités, mais vous retiendrez surtout l'impératif de fraîcheur pour attirer les moteurs.

### **5e principe : la hiérarchie de l'information**

Plus le mot est proche du début (haut) de la page et du début du titre, plus il est considéré comme important pour la page. Une fois que le moteur a repéré le plus important en haut, vous allez lui faciliter le travail en hiérarchisant l'information, en soulignant ce que vous voulez souligner avec des attributs de style (c'est pour le rédacteur) ou de code (c'est pour le webmestre).

Dans une page, les éléments de hiérarchie de l'information sont nombreux. Rubrique, sous-titre, chapeau, intertitres ont chacun un style, une graisse, un corps. Essayez de maintenir une même hiérarchie tout au long du site, cela facilitera la tâche du webmestre et du référenceur, s'il intervient - malheureusement - après le lancement (ce qui arrive trop souvent encore).

Sur le Web, la lecture est plus zappeuse, moins attentive que sur un livre. Ne vous privez pas de souligner un mot par du **gras** ou de l'**italique** et si ce mot est un des mots-clés, vous aurez deux raisons de le mettre en valeur. Le moteur, qui est un lecteur comme les autres (enfin presque ;-)), va remarquer cette mise en évidence.

Les **liens hypertextes** deviennent maintenant essentiels dans une stratégie de mise en valeur de votre contenu.

Illustrons notre propos : votre cactus (encore lui!) vous inspire deux pages et vous pointez un lien [en savoir plus](#) de la première vers la deuxième page, Google tombe sur le lien et il apprend que cette deuxième page parle de "en savoir plus"... Merci, se dit-il (oui, Google se parle à lui-même), mais il y a des milliers de rédacteurs qui renvoient à plus d'information avec cette formule.

A contrario, si votre lien s'intitule [les piquants du cactus](#), Google crédite la page visée d'une meilleure pertinence sur la requête "piquants du cactus", voire "cactus", et vous obtenez un point de plus dans la course aux 10 premières places de résultats du moteur.

Vous vous dites alors qu'il suffit de faire une longue liste de liens avec des mots-clés. Nous y avons déjà pensé et nous l'avons déjà expérimenté. Depuis, Google préfère les liens intégrés au texte, au fil de la plume, c'est plus "naturel". Donc ayez l'air le plus naturel dans vos efforts de rédaction pour les moteurs. Souriez en écrivant!

Vous l'avez compris, les moteurs aiment les textes. Selon de nombreuses sources, il leur faudrait au moins 300 caractères par page. Les pages qui ne comportent qu'un gros titre ont encore parfois leur heure de gloire mais c'est fini, il faut du contenu maintenant.

Dernier conseil de journaliste sur le contenu : il faut **légender** les photos et illustrations. Outre le fait qu'une photo sans légende est interprétée de façon très différente par chaque individu (cours du CFPJ, *Centre de Formation et de Perfectionnement des Journalistes*), les moteurs ne comprennent pas le contenu des images (même Google images). Il leur faut du texte pour savoir de quoi parle cette image. Sans légende, votre image est nue, ou mal habillée par les textes environnants.

### **Hiérarchiser le code de la page**

Pour cette dernière partie, disons que même les webmasters doivent apprendre à écrire pour les moteurs lorsqu'ils structurent les pages web avec leur code. C'est aussi la partie la plus connue des référenceurs.

Une association de standardisation du web, le W3C, a posé depuis une dizaine d'années les bases d'une hiérarchie du contenu. Il n'est pas encore prouvé qu'un respect de cette standardisation donne un avantage certain sur les moteurs. "Pas prouvé", cela signifie qu'il y a autant d'experts "pour" que d'experts "contre" et que personne n'a eu le temps d'expérimenter proprement ce fait. Cela dit, cette standardisation est utilisée par les règles d'accessibilité et cadre avec le code XHTML. Donc un effort dans ce sens ne sera pas perdu et pourrait bien se trouver au menu officiel des classements des moteurs. Vous l'avez compris, nous sommes plutôt "pour"...

### **Explication**

Tous les éditeurs web proposent 6 niveaux de hiérarchie nommés **h1, h2... h6**. Dans les premiers temps du Web, le titre de la page s'attribuait d'office (et dans le code) la balise h1, ce qui lui donnait immédiatement un caractère de corps 32 (pas terrible le look!).

Heureusement, grâce aux feuilles de style, nous pouvons (et il faut) reprendre les bonnes habitudes du standard pour mieux parler aux moteurs : **h1** pour le titre, **h2** pour le sous-titre et ainsi de suite. La feuille de style se réservant la tâche d'attribuer à la balise **h1** corps, graisse, couleur et type de caractère.

Bien sûr, on peut donner une taille respectable au titre et laisser le moteur deviner que c'est bien le titre, mais si le titre est encadré par les balises `<h1>Titre</h1>` et qu'il n'est appliqué qu'une seule fois dans la page, alors cela le désigne immédiatement comme tel. Et les balises en minuscules parce que la tendance passe du HTML au XHTML dont la norme est "balises en minuscules"

Étonnamment, nous assistons au retour de la balise `<strong>` (gras selon le W3C), malgré l'insistance des éditeurs web à utiliser la balise `<b>`.

De même, une liste de sujets ou de produits peut se coder par des retours à la ligne et des puces, mais il est préférable d'intégrer les balises `<ul>`, `<ol>`, balises officielle des listes.

Les moteurs ne voient pas les images, les non-voyants non plus : pour ces deux publics, la balise `<alt>` (alternative text) déclare le contenu de l'image de façon plus directe encore que la légende pour les moteurs et pour les navigateurs sonores. Ce qui ne vous dispense nullement de mettre une légende.

Et si vous le faites avec des mots-clés sans rapport, ce serait une mauvaise action pour les navigateurs des non-voyants. Si le contenu du ALT répète la légende et que un ou deux mots-clés s'y glissent, votre devoir est complet.

Enfin, votre optimisation du texte du lien peut avoir une redondance plutôt positive pour les moteurs par un "title" de lien qui décrit la cible du lien et vous respectez ainsi les règles d'accessibilité. Exemple : `<A href="/marques.htm" title="marques">`

Donc le webmaster doit réviser son W3C pour être "accessible" à tous les internautes et donc aux moteurs de recherche...

### **Conclusion**

Nous espérons que ces trois articles sur l'art de bien écrire pour Google (et les autres...) vous auront ouvert de nouveaux horizons rédactionnels. Ces quelques conseils font partie de l'ensemble "écrire pour le Web" qui est bien plus vaste. Ils ont été écrits pour vous donner un choix d'écriture et l'assurance de tout faire pour exposer votre prose à la lecture des internautes. Les moteurs sont des lecteurs comme les autres... Enfin presque. :-)

**David Cohen**  
CVFM (<http://www.cvfm.com/>)