



Livre Blanc de la création de site web optimisé

Version 1.0 - 12 février 2008



Conçu et réalisé par : **François Houste** (SEO Director - fhouste@lsf-corp.com)
Thomas Poulard (SEO Manager - tpoulard@lsf-corp.com)
Bertrand Hamelin (SEO Executive - bhamelin@lsf-corp.com)

Merci pour leur précieuse aide : Baptiste Bobet, Simon Caporossi, Agnès Gautier, Isabelle Rocher,
Maxime Sarrade, Laure Sommervogel



Introduction	2
Optimisation sémantique	3
Définition des mots et expressions clés	3
Rédaction des titres HTML	3
Utilisation des balises META	4
Optimisation du texte visible	4
Nomenclature des liens	5
Optimisation technique	6
Arborescence et structure des URL	6
Mise en place de redirections 301	7
Navigation	7
Conclusion : Les points SEO à ne pas oublier	8
A propos de LSF Interactive	8

Introduction

Si le référencement naturel a beaucoup changé depuis cinq ans, il repose toujours sur deux bases très solides : la pertinence accrue des textes qui composent les pages d'un site et l'accessibilité technique des moteurs de recherche. Ces deux éléments sont encore aujourd'hui la base sur laquelle doit se construire toute stratégie de référencement naturel appliquée sur internet. Deux prérequis indispensables. Et pour bien faire, ces éléments doivent être pris en compte le plus tôt possible dans le processus de création d'un site web !

Car malheureusement le référencement naturel n'est pas encore considéré, du moins en France, comme une étape indispensable de la conception d'un site. Même si les habitudes des entrepreneurs changent doucement à ce sujet. Aujourd'hui les réflexions liées au référencement naturel proviennent plus souvent d'un mauvais positionnement face à sa concurrence que d'une vraie prise de conscience du potentiel de trafic et de chiffre d'affaires des moteurs de recherche.

C'est l'une des raisons de l'existence de ce livre blanc de la création optimisée : fournir aux créateurs de sites et de commerces en ligne les premières clés leur permettant de concevoir et de mettre en place un site optimisé dès sa création. Parce que l'utilisation de bonnes méthodes de référencement au plus tôt permet de profiter, le plus vite possible, d'une audience gratuite et fortement qualifiée. Et parce que toute optimisation effectuée avant la sortie effective d'un site est un moyen d'économiser sur l'optimisation technique future de celui-ci.

Ce livre n'a bien entendu pas la prétention d'être une "bible" du référencement naturel ni de présenter l'ensemble des méthodes permettant d'acquérir un trafic optimisé via les moteurs de recherche. Un simple livre blanc n'y suffirait pas, et les recommandations précises de référencement doivent de toute façon s'adapter au cas par cas... Il se propose plutôt de fournir de bonnes bases, de bonnes méthodes, qui constituent les premiers pas d'une politique de référencement efficace.

Ce livre blanc se décompose en deux parties principales. Dans la première, vous découvrirez les astuces liées à l'optimisation sémantique d'un site, depuis le choix de mots clés stratégiques jusqu'à leur positionnement dans les pages. Dans la seconde partie, il sera question des choix techniques liés à l'accessibilité du site et à sa bonne compréhension par les moteurs de recherche. C'est en combinant ces deux domaines d'optimisations que vous obtiendrez les meilleurs résultats sur internet.

Optimisation sémantique

Parce que les recherches d'informations sur internet se font avant tout à l'aide de mots, le contenu d'un site web constitue sans doute l'aspect le plus important du référencement naturel. De la façon dont sont optimisés et présentés les textes d'un site va dépendre une bonne partie de son positionnement, et par conséquent de son audience.

Afin qu'une page soit correctement référencée pour une expression précise, on considère généralement que cette dernière doit se retrouver à la fois :

- dans le titre HTML de la page ;
- dans ses balises META liées au référencement (Description et Keywords) ;
- dans le texte réel de la page, c'est à dire à la fois le contenu textuel, mais aussi les en-têtes, légendes d'image, intitulés des liens, ...

Les pages suivantes passent en revue quelques-uns de ces postes d'optimisation sémantique, et présentent quelques façons d'en tirer le meilleur parti possible.

Définition des mots et expressions clés

Mais avant de se pencher sur l'optimisation sémantique de son site, il faut identifier les mots et expressions clés susceptibles de générer un trafic pertinent.

Il est important, lors de la mise en place d'une stratégie de référencement, de bien identifier les mots clés sur lesquels un nouveau site se doit avant tout d'être présent. C'est en effet par l'intermédiaire de ces mots clés que la majorité du public du site accèdera à celui-ci, après avoir saisi la requête correspondante dans un moteur de recherche. La sélection des mots clés est donc cruciale pour l'audience générale du site.

Plusieurs méthodes existent pour identifier les mots clés importants. Tout d'abord : se baser sur votre propre expertise et votre propre expérience. Les propriétaires d'un site web savent mieux que quiconque ce qu'ils proposent dans leurs pages et la façon dont ils souhaitent le mettre en avant. Une première liste de mots clés proviendra très certainement de cette "expertise métier". Au passage, utiliser les informations d'utilisation du moteur de recherche interne de votre site est également une très bonne piste pour connaître la façon dont les internautes nomment vos produits.

On peut également définir une liste de mots clés pertinents en utilisant les outils d'analyse d'audience mis à disposition par les régies marketing en ligne, telles **Google AdWords** ou **Yahoo! Panama**. A l'aide de **Google Trends**, il est possible de voir l'évolution de mots clés sur une période donnée. Ces outils permettent de hiérarchiser l'importance des mots clés et ainsi de travailler sur les mots véritablement générateurs de trafic.



Le choix des mots clés à cibler est à la fois fonction du trafic espéré et du degré d'adéquation de ces expressions avec le sujet du site final et son public réel. Une fois des mots clés à fort potentiel définis pour un site, il convient ensuite d'utiliser ceux-ci à bon escient sur l'ensemble des pages du site. Ne pas prendre assez de temps pour réfléchir aux mots clés peut amener à reprendre l'optimisation sémantique du site au tout début. C'est pour cela qu'il est indispensable de ne pas laisser cette partie de côté.

Rédaction des titres HTML

Une bonne rédaction des titres HTML des pages (Balise <TITLE> présente dans l'en-tête de chaque page) est un élément crucial de l'optimisation d'un site. D'une part, le texte présent dans ce titre possède un poids important dans la reconnaissance des mots clés d'une page par les moteurs de recherche puisqu'il apparaît en tant que titre dans les résultats de requêtes de ces moteurs de recherche. D'autre part, le ou les mots clés utilisés sont normalement ceux pour lesquels la page apparaîtra dans les résultats des moteurs.

Techniquement, le titre HTML d'une page doit être le premier « texte utile » rencontré par les moteurs de recherche lorsque ceux-ci parcourent un site. La balise **<TITLE>** doit donc être la première balise écrite à l'intérieur de la balise **<HEAD>** dans le code source de la page.

Une liste courte de mots clés génériques autour de l'article et de la rubrique en question peut être utilisée pour compléter le titre HTML sur des sujets pertinents. Prenez garde toutefois à la longueur de votre balise **<TITLE>**. Si le W3C (World Wide Web Consortium, organisme régissant les standards du web) recommande de limiter celle-ci à 64 caractères, gardez en tête que les moteurs n'afficheront qu'entre 80 et 120 caractères du titre que vous avez choisi dans leurs pages de résultats. Vous pouvez bien entendu écrire un texte plus long, mais vous diminuerez alors l'importance de chaque mot clés dans le titre.

Utilisation des balises META

Les balises META permettent d'indiquer aux moteurs de recherche certaines informations sur le contenu d'une page web. Ces balises, dont seules les variantes **Description** et **Keywords** sont pertinentes pour le

Méta :	
Nom	Contenu
description	Amazon.fr : achat et vente de DVD, jeux vidéo, CD, livres, lecteurs MP3, ordinateurs, apparei...
keywords	Amazon, Amazon.fr, achats en ligne, ventes en ligne, neufs, occasions, livres, librairie, DVD, f...

référencement naturel, ont beaucoup moins d'importance aujourd'hui pour le référencement qu'il y a quelques années. Les dernières versions des moteurs de recherche n'en tiennent quasiment plus compte dans le calcul de la pertinence. Il reste toutefois intéressant de travailler ces balises pour des moteurs mineurs, ou pour la visibilité qu'elles peuvent avoir dans les pages de résultat.

META Description

La balise **META Description** représente la phrase utilisée en général par les moteurs de recherche pour décrire le contenu de la page indexée. Cette balise permet ainsi aux internautes d'avoir un résumé de la page et de voir si celle-ci est pertinente par rapport à leur recherche. Assurez-vous donc que le texte de cette balise est pertinent vis-à-vis du contenu de votre page et possède bien les mots clés sur lesquels vous avez choisi celle-ci.

META Keywords

La balise **META Keywords** se compose en général d'une vingtaine de mots-clés, tous optimisés pour le référencement naturel. Elle n'est aujourd'hui plus reconnue par les moteurs majeurs que sont Google, Yahoo! ou Microsoft Live. Elle ne présente donc qu'un intérêt très limité pour le référencement naturel, mais peut toutefois parfois s'avérer utile pour souligner les fautes de frappe et mettre en avant certains contenus sur des moteurs mineurs.

Optimisation du texte visible

Le **texte visible** d'une page web se compose de l'ensemble de ses éléments textuels, mais aussi de ses titres et des légendes de ses images. Ce sont l'ensemble de ces éléments qui doivent être pris en compte pour le référencement naturel.

Le texte d'une page

La longueur idéale d'une page optimisée se situe entre 250 et 300 mots. Elle est très souvent impossible à garantir, c'est pourquoi on se concentre en général sur l'optimisation des premiers 250 à 300 mots de la page. Les moteurs de recherche cherchent dans le texte d'une page un mot clé, une phrase, ou une combinaison de mots (deux ou trois) qui est répétée plusieurs fois. Idéalement cette phrase est également répétée dans le titre HTML de la page, dans ses balises META.

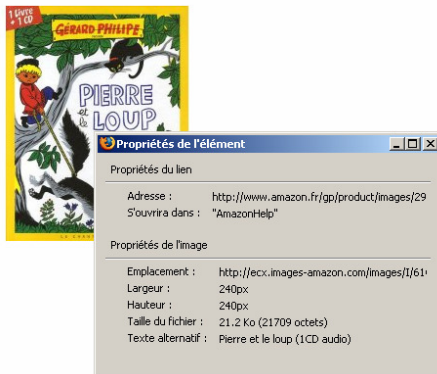
Les moteurs de recherche peuvent reconnaître les pages dont les mises à jour sont fréquentes et visiter celles-ci plus souvent afin de garder leur index le plus à jour possible. Un site de boutique en ligne peut parfaitement profiter de cette particularité afin d'optimiser son référencement pour les moteurs de recherche. Une page d'accueil dont le contenu n'est pas figé mais évolue de jour en jour peut ainsi être visitée plus régulièrement par les robots des moteurs de recherche. Et garantir une meilleure indexation dans les moteurs de recherche.



Titres <H1>/<H2>

La balise <H1>/<H2> en HTML est utilisée pour afficher les en-têtes, c'est-à-dire les titres de page et de section. Cette technique permet surtout de hiérarchiser le contenu d'une page. Les robots des moteurs de recherche prennent en compte des phrases qui apparaissent entre les balises <H1>/<H2> et leur attribuent généralement une importance plus grande. Une page bien optimisée devrait donc utiliser cette balise au maximum.

Légende des images (attribut ALT)



L'attribut **ALT** de la balise Image () permet d'associer un texte de description à chaque image d'une page web. La fonction originelle de cet attribut est d'aider les malvoyants à accéder à l'intégralité du contenu du site. Cet attribut permet également aux robots des moteurs de recherche de lire des images et de mieux comprendre le sujet d'une page. Il se met en place sous la forme :

```

```

Il est donc possible d'augmenter le contenu utile d'une page et de mettre en avant les phrases les plus importantes en donnant un attribut **ALT** à toutes les images de son site. Par ailleurs, cet attribut va donner aux images une pertinence pour les moteurs de recherche d'images et la recherche universelle.

Nomenclature des liens

Les principales technologies de moteur de recherche laissent également la part belle aux liens. La qualification d'une page en fonction de certains mots clés est ainsi déterminée par les textes utilisés sur les liens pointant vers celle-ci. Ainsi, une page citée par de nombreux sites à l'aide du terme "**Blog musical**" se verra naturellement favorisée dans les moteurs de recherche pour cette expression. Si le nombre de liens utilisant ces mêmes mots est très important, le positionnement de la page cible sur ce mot clé sera naturellement amélioré. En revanche, un site dont les articles sont systématiquement liés par un lien "**En savoir plus**" ou "**Lire la suite**" se verra, lui, optimisé pour ces deux expressions dans les résultats de recherche. C'est la raison pour laquelle ces dénominations (ainsi que le fameux "**Cliquez ici**") sont de moins en moins utilisées sur les sites professionnels.



Il est donc très important de prendre garde à la nomenclature des liens utilisés sur le site et menant aux différentes rubriques et aux différents articles de celui-ci.

Une attention toute particulière doit être portée au choix de nom des différentes rubriques du site. Ces noms de rubriques sont en effet utilisés à la fois en tant que titre des pages (aussi bien titres HTML qu'en en-tête de page) dans les URL mais également comme ancres des liens menant vers les pages d'accueil de ces rubriques.

Ces noms de rubrique sont donc d'autant plus précieux en termes de référencement qu'ils constituent une masse importante de liens disponibles sur l'ensemble des pages du site au travers de la navigation du site. On évitera ainsi l'utilisation systématique du terme "**Accueil**" pour renvoyer vers la première page d'un site, cette utilisation favorisant avant tout le référencement sur le terme "**Accueil**"...

Il est donc important de veiller à ce que les mots choisis pour ces rubriques soient des mots qu'il est possible de référencer et de placer dans les moteurs de recherche.



Optimisation technique

Si l'optimisation des textes d'un site est critique pour le référencement naturel, elle n'est pas le seul paramètre à prendre en compte pour être positionné de manière pertinente dans les moteurs de recherche. **L'optimisation technique** d'un site a également beaucoup d'importance, et sert toujours de socle aux meilleures stratégies de référencement naturel. S'il est important que les moteurs de recherche puissent comprendre le sujet qu'aborde une page, il est encore plus important que ceux-ci puissent accéder à l'intégralité des contenus de celle-ci. C'est là tout le rôle du référencement technique.

Arborescence et structure des URL

Pour profiter pleinement d'une indexation dans les moteurs de recherche, un site web se doit aujourd'hui de posséder des URL (*Uniform Resource Locator*, adresse d'une page sur le réseau internet) optimisées. A chaque page d'un site correspondra donc une adresse unique qui restera valable dans le temps. Cette adresse doit contenir différents mots clés, titres de rubriques ou titres d'articles utilisés sur la page, et susceptibles d'être recherchés par les internautes dans les moteurs de recherche. En cas de présence d'espaces dans les titres des rubriques et des articles, ceux-ci devront être remplacés par des tirets (-), ces caractères étant interprétés comme des séparateurs par les moteurs de recherche. La présence d'espaces dans l'adresse d'une page peut en effet présenter des inconvénients sur certains types de serveurs ou lors de la propagation d'une URL par mail ou flux RSS.

De la même façon, l'ensemble des caractères spéciaux et accentués qui pourraient éventuellement être utilisés dans les titres des rubriques et articles devra être remplacé par leurs équivalents non accentués ou supprimés des URL réécrites, ceci afin d'éviter tout problème technique dans la construction des adresses des différentes pages.

Les sites développés à partir de technologies dynamiques (telles ASP, PHP ou encore JSP et ColdFusion) présentent en général des adresses de page non optimisées. En effet, la plupart des pages ont besoin de différents arguments afin d'aller chercher les données nécessaires à la construction de celles-ci dans une base de données. Ces arguments sont présents à la fin de l'adresse de la page à afficher, séparés par des caractères spéciaux du type "?" ou encore "&". Par exemple, pour demander à une page d'afficher l'article n°5 d'une base de données, un site construit en PHP pourrait utiliser – suivant la façon dont le site est développé bien entendu – une syntaxe proche de <http://www.site.com/article.php?id=5>. Ce type d'adresse possède deux inconvénients dans le cadre de l'optimisation du référencement :

- Une multitude de pages porte une adresse similaire (**article.php**), simplement différenciée par une suite d'arguments ;
- Certains moteurs de recherche n'interprètent pas les données situées derrière les caractères spéciaux (**?id=5**) et sont donc incapables d'indexer les différents articles d'un site.

Pour pallier ces défauts, on utilise en général la technique dite de l'URL Rewriting (ou réécriture d'URL). Cette technique permet de traduire une URL dynamique présentant divers arguments en une suite de dossiers et de fichiers pleinement compréhensible par les moteurs de recherche.

Ainsi, une page originale nommée <http://www.site.com/article.php?id=5> pourra être accessible par l'adresse <http://www.site.com/article/5/>, qui ne possède pas de caractères spéciaux et est donc pleinement compréhensible par les principaux moteurs de recherche.

Une utilisation plus poussée de cette technique permet même de mettre en avant des données dynamiques au sein des URL d'un site. Ainsi, pour un site de e-Commerce, on peut facilement imaginer que les noms de produits ou les catégories auxquelles ils appartiennent soient cités dans les adresses pour une meilleure optimisation. On se retrouve ainsi avec des URL du type : <http://www.site.com/musique/supergrass-diamond-hoo-ha/5/>, qui offrent l'avantage de répondre à plus de requêtes dans les moteurs de recherche.

La technique de réécriture des URL utilisées dépend en grande partie du type de serveur sur lequel est hébergé le site final. Dans tous les cas il est impératif que les URL des différentes pages du site soient optimisées et fixes.

Mise en place de redirections 301

Lors de la construction d'un nouveau site, l'élément le plus important lors de la migration va être la gestion des redirections. Il existe plusieurs types de redirections. Seules les redirections 301 sont utiles au référencement naturel car elles permettent de conserver l'historique de positionnement déjà acquis sur les moteurs de recherche par la version précédente du site.

Théoriquement ces redirections doivent se faire de page à page. En cas d'impossibilité, on choisira donc la page la plus adéquate du nouveau site (en-tête de catégorie) ou à défaut sur la page d'accueil.

La mise en place technique de ces redirections peut varier suivant la technologie de serveur ou suivant le langage utilisé. Dans tous les cas il convient de bien se renseigner sur le type de serveur et ne surtout pas migrer un site avant la mise en place de ces redirections sous peine de perdre son historique de référencement sur les différents moteurs de recherche.

Navigation

Les robots des moteurs de recherche ne peuvent pas lire le contenu des scripts JavaScript, ni celui des fichiers Adobe Flash et Shockwave. Certains moteurs ne peuvent également pas lire les liens définis dans une Image Map.

Le JavaScript est un langage de programmation de type script créé en 1995 qui sert surtout à contrôler les données saisies dans des formulaires HTML., mais il est aussi souvent utilisé pour réaliser des services dynamiques. Certains sites ont de plus en plus tendance à abuser de ce type de script, notamment depuis l'apparition de la norme de développement AJAX et son application étendue aux sites dit "Web 2.0".

Les fichiers Adobe Flash sont issus d'un programme permettant la création d'animations vectorielles interactives. Ils sont souvent des fichiers insérés dans des pages web dans le but de fournir des animations à celles-ci. De plus en plus de sites internet sont entièrement construits grâce à cette technologie, ce qui pose un réel problème de référencement, les robots des moteurs de recherche étant incapables de lire les textes inclus dans ce genre d'animation. Toutefois ceci ne devrait pas empêcher les créateurs de site d'utiliser ces technologies. Il suffit de garder en tête qu'une solution de navigation alternative, adaptée aux moteurs de recherche, doit être mise en place sur le site afin de permettre son référencement.

La structure idéale de site n'existe pas. Il faut pouvoir à la fois créer un site agréable à parcourir pour l'utilisateur et facile à parcourir par les robots des moteurs de recherche.

Conclusion : Les points SEO à ne pas oublier

Plutôt qu'une conclusion rapide à l'ensemble de ces pistes d'optimisation, un petit résumé des points cruciaux du référencement naturel s'impose. Voici donc un pense-bête, rapide et pratique à utiliser, listant les principaux points dont il est bon de s'assurer de l'optimisation avec la sortie d'un site :

- Définir une liste de mots clés pertinents et générateurs de trafic ;
- Ecrire des titres HTML optimisés et différenciés pour chacune des pages du site ;
- Optimiser ses balises META Description et Keywords et les personnaliser pour chacune des pages ;
- S'assurer de la présence de mots clés ciblés dans le texte visible de ses pages ;
- Hiérarchiser ses titres et mettre en avant le sujet principal des pages dans des balises de type <h1> ;
- Accoler un légende à chacune des images du site ;
- S'assurer de la pertinence des ancrs de lien et des URL vis-a-vis de sa liste de mots clés ;
- Concevoir une navigation claire aussi bien pour les moteurs que pour le visiteur, en évitant les techniques hermétiques comme le JavaScript ou Adobe Flash.

Ces différents points ne sont bien entendu qu'une base pour le référencement naturel d'un site. D'autres points, d'autres stratégies sont à prendre en compte pour positionner un site de manière correcte sur les moteurs de recherche. Les réseaux sociaux et l'impact de plus en plus important de la recherche universelle sur les résultats des moteurs obligent à développer de nouvelles stratégies de référencement et d'acquisition de trafic sur le web. Mais quelles que soient les pistes explorées, la possession d'un site accessible et sémantiquement optimisé est une étape indispensable à ces stratégies !



A propos de LSF Interactive

LSF Interactive est une division du groupe LSF Network, créé en 1999. Installée dans la Silicon Valley, LSF Interactive élabore des campagnes en ligne intégrées pour les plus grandes marques. Le groupe travaille directement avec les annonceurs et leurs agences et fournit à ses clients des contacts – leads – qualifiés sur une base de paiement à la performance via l'ensemble des principaux canaux de distribution internet qui incluent notamment un réseau partenaire affilié de plus de 200 éditeurs sélectionnés pour la qualité de leur audience, des compétences pointues dans le search et l'e-mail, et un site propriétaire de plus de 16 millions d'utilisateurs enregistrés dans le monde.

Pour en savoir plus : <http://www.lsfinteractive.fr>