

# Les bases du référencement

## SEO : les **trois clefs** du référencement naturel

- **L'architecture globale du site** qui comprend :
  - le maillage des pages entre elles, l'arborescence globale,
  - le niveau de profondeur des pages,
  - la technologie utilisée,
  - la propreté et la lisibilité du code source par les moteurs,
  - le temps de chargement des pages,
  - la navigabilité,
  - le taux de duplication des pages,
  - ...
- **La sémantique des pages** qui contient, entre autres :
  - l'optimisation des méta title et description,
  - la présence de contenu original,
  - le balisage sémantique des pages,
  - le maillage interne des pages,
  - l'optimisation des contenus multimédias (images, vidéos, etc.),
  - ...
- **Le netlinking** qui prend différentes formes :
  - les liens en provenance des sites d'annuaires, de communiqués de presse, blogs, forums, sites partenaires, etc.,
  - les backlinks issus des réseaux sociaux,

La sélection des **mots-clés** qui est au cœur de la **stratégie**. D'ailleurs, on peut imaginer une analogie entre le référencement et le corps humain :

- la **structure** du site correspond au **squelette**, à l'ossature du site web. Il s'agit de fondations qui, une fois mises en place, demeurent intangibles.
- la **sémantique** qui représente les **poumons**. C'est elle qui alimente le site en oxygène et permet au site de continuer à vivre au lieu de passer pour obsolète auprès de Google.
- le **netlinking** symbolise les **muscles**. Il permet au site d'être enduring et de tenir sur la durée. Mais c'est comme tout, à vouloir aller trop vite, trop mal, on finit par s'essouffler. Il faut garder un rythme de soumissions pertinent.
- les **mots-clés**, comme nous l'avons déjà dit au-dessus, caractérisent le **cœur** de la stratégie de référencement. Sans eux, votre site serait une coquille vide et l'ensemble des étapes précédentes seraient bien inutiles.

D'autres critères viendront influencer le positionnement d'un site web dans les moteurs de recherche. Parmi un nombre important de facteurs de ranking, en voici les principaux :

- l'**ancienneté** du nom de domaine,
- l'**autorité** du site dans son domaine,
- le **taux de rebond**,

- le [PageRank](#),
- le [volume de pages](#) indexées,
- une [présence active](#) sur les [réseaux sociaux](#),
- ...
- 

## Actions qui permettent d'améliorer votre référencement sur le long terme.

### La check-list des bonnes actions

#### Optimisations techniques

##### Faciliter le travail d'indexation des moteurs de recherche

###### Le fichier sitemap.xml

Placé à la racine d'un site sur le serveur, le fichier sitemap.xml permet de :

- préciser aux moteurs l'ensemble des URL qui composent votre site
- leur signaler si des modifications ont eu lieu sur le site depuis le dernier passage des spiders

Illustration d'un fichier sitemap.xml :

###### Le fichier robots.txt

Placé à la racine d'un site sur le serveur ftp, le fichier **robots.txt** permet d'indiquer aux moteurs les pages à désindexer sur votre site.

#### Améliorer l'expérience utilisateur

##### La page 404 personnalisée

Lorsqu'un internaute arrive sur votre site et échoue sur une page d'erreur 404, il risque de le quitter prématurément, augmentant ainsi votre taux de rebond qui est l'un des critères majeurs de positionnement d'un site web. Pour y remédier, il est indispensable de mettre en place une page 404 personnalisée qui redirige les utilisateurs vers les principales rubriques de votre site, l'objectif étant de [garder l'internaute sur votre site](#) plutôt que de le laisser partir.

##### Le temps de chargement des pages

Le temps de chargement des pages est un facteur de positionnement déterminant. [Les moteurs de recherche le prennent de plus en plus en compte](#) dans le calcul du classement des pages de résultats. Il est donc essentiel de veiller à ce que votre site se charge rapidement.

Un site trop lent entraîne un **départ prématuré** des internautes avant la fin du chargement de la page, ce qui a pour conséquence l'augmentation du **taux de rebond**, cette hausse étant nuisible à la visibilité d'un site.

Dans certains cas extrêmes, il peut même arriver que les Googlebot (robots qui indexent les sites web), ne parcourent pas/ plus un site internet à cause d'un temps de réponse des pages trop élevé. Par exemple dans le cas d'un serveur de mauvaise qualité ou en surcharge de trafic, il a été confirmé que le Googlebot puisse prendre la décision de ne pas « crawler » le site afin d'éviter une charge supplémentaire du serveur pouvant entraîner sa mise hors ligne.

### Exemples d'optimisations possibles :

- compression gzip des **images**,
- mise en cache des ressources statiques dans le navigateur,
- spécification de la dimension des images dans le code HTML,
- usage des sprite CSS,
- limitation du nombre de requêtes entre le serveur et le navigateur,
- différer l'analyse du JavaScript par les navigateurs web,
- éviter les hébergements gratuits,
- ...

### Partir à la chasse au duplicate content

Le **duplicate content** sur un site est l'un des principaux **fléaux** susceptible de lui causer une **pénalité**. En effet, si Google aperçoit deux pages de mêmes URL ayant le même contenu, il peut prendre des mesures qui consistent à **exclure** un site (ou une page) **des résultats des moteurs de recherche**, rendant votre **visibilité nulle**.

Il existe plusieurs types de duplication :

- la **duplication totale** : un site web identique entre 2 noms de domaine (Ex : [www.site1.com](http://www.site1.com) et [www.site1.fr](http://www.site1.fr)), l'absence de redirection en 3W (Ex : [www.site1.com](http://www.site1.com) et [site1.com](http://site1.com)), etc.,
- la **duplication partielle** : plusieurs pages sont optimisées sur le même titre, le même titre, le même contenu repris en partie.

Par ailleurs, des dizaines d'autres cas de duplication de contenu coexistent.

### Optimisations sémantiques

#### Donner de la lecture aux moteurs de recherche

Vous l'aurez deviné, la matière première d'un site web, c'est son **contenu**... Sans cela, les moteurs de recherche sont bien incapables de le positionner au sein des pages de résultats, faute d'avoir suffisamment d'**informations** pour **comprendre** ce dont votre site parle. Mais, ce n'est pas tout... Avoir du **texte**, c'est bien, encore faut-il qu'il soit **optimisé** pour le référencement. Pour ce faire, veillez à placer vos **mots-clés** dans les **zones chaudes** de chacune de vos pages (méta title, **titres** et sous-titres, header, liens, **début de paragraphe**, etc.). Ne commettez pas l'erreur suprême de ne cibler qu'une page pour une quantité infinie d'expressions. Un site bien fait est principalement un site qui

permet de **cibler un mot-clé par page**. Si ce n'est pas le cas, vous limitez vos chances d'être bien positionné dans les moteurs de recherche sur une expression ciblée car, en effet, vous **diluez le poids de vos mots-clés** relativement les uns par rapport aux autres.

Bien penser **l'architecture** de votre site en amont et la façon dont vos pages seront **organisées** et **reliées les unes aux autres** sera alors d'une **importance capitale**. En termes d'optimisation sémantique, le maillage interne est également fondamental pour favoriser certaines pages dans les moteurs de recherche. Il permet également d'éviter que le site ne reçoive que des liens entrants, indicateur éventuel de sur-optimisation et de spam pour Google notamment.

## Le netlinking : variez les backlinks

Le **netlinking** est devenu un véritable travail d'équilibriste. **Trop de liens** depuis une même ancre et, qui plus est, en provenance du même type de plateformes, autant vous dire tout de suite que vous serez, à terme, **pénalisé**. Des liens en **quantité insuffisante**, voire inexistant, et votre site souffrira d'un **manque évident de visibilité**. Quoiqu'on en dise, les **liens sont l'essence même du web**. Ils ont encore et toujours le vent en poupe, à cela près qu'aujourd'hui, la façon de les travailler a changé de cap. Il est temps désormais de naviguer avec prudence et de manœuvrer avec **parcimonie** pour éviter de sombrer dans l'œil du cyclone... le fameux ouragan Google qui pénalise les sites frauduleux, usant de techniques dites « spammy » telles que la sur-optimisation des ancres de liens et le bourrage de mots-clés sur une page ou encore la duplication et l'absence de contenus.

### Les règles à respecter :

- **varier** les ancres de liens, c'est-à-dire les **mots-clés sur lequel le lien est placé** (notoriété, mots-clés génériques, URL, mots-clés insignifiants, etc.),
- **diversifier son réseau de liens** (partenaires, commentaires de blogs, annuaires, blogs, communiqués de presse, etc.),
- **varier les zones** de mise en place du lien (contenu, footer, blogroll, etc.),
- obtenir des **liens depuis des sites d'autorité** à fort PageRank (5/10) par exemple, mais aussi des sites au PageRank faible (1/10 par exemple),
- acquérir des **liens depuis des sites de même thématique** ou thématique proche,
- posséder un ratio de liens dofollow/nofollow partagé à 50%,
- développer les **liens en provenance des réseaux sociaux** (Google +, Facebook, Twitter, etc.),
- **réaliser des liens sortants vers d'autres sites**