

HTML et CSS

Référencement naturel et SEO

Les principes de bases	2
Qu'est-ce que le référencement naturel ?	2
Qu'est ce que le SEO ?	2
Pourquoi suis-je concerné en tant que développeur ?	2
Les bonnes pratiques SEO	3
Utilisez les balises HTML sémantiques	3
Choisir les titres avec précaution	3
Un lien, simple, basique	4
Les images	4
Les balises meta	5
Le nom de domaine et le nom des fichiers	6
Gérer l'exploration et l'indexation des robots	6
Fichier robots.txt	6
Sitemap	6
Balises meta no-index	7
Les liens nofollow	7
Expérience utilisateur	8
Pensez au mobile	8
Performance d'un site	8
Les outils SEO	9
Google Search Console	9
Lighthouse	9

Les principes de bases

Qu'est-ce que le référencement naturel ?

Lorsque vous créez un site internet et que vous le publiez sous un nom de domaine (<https://monsite.com>), il n'est pas directement présent sur les moteurs de recherche, principaux points d'entrées du web. Chaque moteur de recherche possède des armées de robots qui explorent (crawl) le web en visitant toutes les pages, en passant de liens en liens afin de référencer tout le contenu accessible des sites.

C'est ce que l'on appelle le référencement naturel, car à priori il ne nécessite pas d'actions de votre part : les robots de Google et consort viendront d'eux-mêmes sur votre site afin de le référencer. Tout ceci dans le simple but de proposer les pages de votre site dans les résultats des recherches de leurs utilisateurs.

Qu'est ce que le SEO ?

Le SEO, pour Search Engine Optimisation, est l'ensemble des pratiques qui permettent de favoriser le référencement de vos pages par les robots des moteurs de recherche. L'objectif est de favoriser le crawl et d'être BIEN référencé :


- C'est-à-dire d'apparaître sur les résultats des recherches correspondants à la thématique de votre site, donc de cibler les bons mots clés.
- D'apparaître en bonne position dans les résultats, si possible en première position : Le Saint Graal.

Pourquoi suis-je concerné en tant que développeur ?

A priori on serait tenté de se dire :

C'est une problématique de l'équipe communication et marketing. En tant que développeur je ne me sens pas concerné.

Seulement le référencement n'est pas qu'une question de contenu et de marketing : La façon dont vos pages HTML sont structurées, où les textes pertinents sont placés, la façon dont vos pages s'affichent sur mobile, la rapidité de chargement de vos pages... Tout ceci est bien à la main des développeurs. C'est ce que l'on a tendance à appeler le SEO technique.



Comme nous allons le voir, le SEO technique est avant tout un ensemble de bonnes pratiques de développement qui vous incite à rendre vos sites plus lisibles, accessibles et performants.

Les bonnes pratiques SEO

Utilisez les balises HTML sémantiques

Le HTML 5 permet d'avoir un panel très large de balises qui permettent de donner du sens à leur contenu : ce sont des balises sémantiques. Il n'est plus question de n'utiliser que des balises `div` et `span` dans vos pages. Voici quelques balises sémantiques de base :

- **header** pour l'entête de la page
- **footer** pour le footer de la page
- **main** pour le contenu principal de la page
- **section** pour une partie du contenu de votre page
- **article** pour un ensemble de contenu dans une section
- **p** pour un paragraphe de texte
- **a** pour un lien
- **input** pour un champ de saisie de formulaire
- **label** pour une étiquette de champ de formulaire
- **img** pour une image
- **video** pour une vidéo
- ...

Vous trouverez une documentation précise de toutes les balises HTML et leur sens ici :

- <https://developer.mozilla.org/fr/docs/Web/HTML/Element>.

Il faut bien comprendre que les robots qui explorent et indexent vos pages n'ont pas vraiment d'autres informations que ces balises pour comprendre la hiérarchisation de ce qu'elles contiennent et donc l'importance de chaque texte ou image.

Choisir les titres avec précaution

Les titres représentent les textes les plus importants d'une page. Il ne doit y avoir qu'un seul titre de premier niveau par une balise **h1** et il représente le titre de la page. Nous allons donc nous assurer qu'ils contiennent les mots clés que nous souhaitons cibler : ceux pour lesquels nous souhaitons que notre page apparaisse dans les moteurs de recherche.

On utilisera ensuite des balises **h2** pour structurer la suite de notre contenu. Puis des balises **h3** et **h4** si un sous découpage se révèle judicieux. Cette structuration générale des contenus de la page va de paire avec les balises **section** et **article**.

Un lien, simple, basique

Pensez que le lien est la brique de base d'internet qui permet de tisser le réseau et le moyen par lequel les internautes et les robots naviguent d'un contenu à un autre. Il s'agit donc de l'élément clé du SEO !

Tout lien hypertexte qui permet d'accéder à un contenu (sur la même page, une autre page ou un autre site) doit impérativement être une balise **a** comportant un attribut **href** pointant sur ce contenu.

Attention: Bannissez toute pratique visant à transformer une autre balise qu'un **a** en lien via du Javascript.

Le texte du lien doit décrire de manière pertinente le contenu cible. En effet, un lien impacte le positionnement de la page cible pour les recherches faites pour les mots qu'il contient. Exemple :

```
<a href="petits-pois.html">Cultivez vos petits pois au jardin</a>
```

Dans cet exemple on va renforcer le positionnement de la page *petits-pois.html* dans les moteurs de recherche pour les expressions suivantes :

- cultivez
- petits pois
- jardin

Un lien aura encore plus de poids sur le positionnement d'une page s'il provient d'un site externe.

Pour résumer, plus une page aura de liens pointant vers elle, avec des mots clés variés et pertinents, plus elle aura de poids sur internet et mieux elle sera classée par les moteurs de recherche : c'est le principe du Page Rank de Google. De nos jours, les résultats des moteurs se basent sur bien d'autres critères que le Page Rank ou équivalent.

Les images

Voici quelques conseils concernant les images dans les pages web pour un bon SEO :

- Dans l'attribut **alt** des images, mettez un texte qui a du sens et qui correspond bien à l'image à afficher. C'est important pour l'accessibilité autant que pour le SEO.
- Pensez à compresser les images tant que possible pour minimiser le poids total de vos pages. (Voir Performance d'un site)
- Dimensionner vos images à leurs tailles d'affichage dans la page afin de minimiser également le poids total de vos pages.
- Pensez à donner des noms à vos fichiers image qui ont du sens : *lapin-bleu.jpg* sera plus pertinent que *photo033.jpg*.
- Lorsque vous mettez en place du lazy load, veillez à ce que les balises **img** soient toutes initialement affichées pour que les robots les trouvent.

Les balises meta

Les balises **meta** sont des balises que vous intégrez dans le **head** de vos pages afin de donner des informations sur celles-ci. Ces contenus ne sont pas affichés sur les pages mais sont utilisés par les moteurs de recherche.

Voici les 2 balises les plus importantes à renseigner pour le SEO :

- La balise **title** correspond au titre affiché sur l'onglet du navigateur. Le texte présent dans cette balise a énormément d'importance dans le référencement de la page, au même titre que le **h1** de la page. Elle doit comporter peu de mots soigneusement choisis. Chaque page de votre site doit avoir une balise titre différente.
- La balise **meta** description correspond à une description courte du contenu de la page. Le contenu de cette balise semble ne plus avoir d'impact sur le positionnement de la page dans les résultats de Google. Cependant ce texte peut être affiché dans les pages de résultats.

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Page title</title>
  <meta name="description" content="Here is a short description of the
page.">
```

```
</head>
```

Le nom de domaine et le nom des fichiers

Pour une stratégie SEO complète et anticipée, il est important de choisir un nom de domaine qui contient un ou des mots clés pour lesquelles vous souhaitez positionner le site sur les moteurs de recherche.

De la même manière, il est important de nommer les fichiers des pages et des images du site avec des noms parlant. Par exemple : */nos-services-de-developpement-de-sites-internet.html* aura un meilleur impact que : */page3.html*

Gérer l'exploration et l'indexation des robots

Fichier robots.txt

Le fichier *robots.txt* est un fichier qui se place à la racine de votre site et qui est lu par les robots afin de leur spécifier les contenus qu'ils ont le droit d'explorer ou non sur votre site à l'aide de directives **Allow** et **Disallow** associées à des chemins.

```
User-agent: Googlebot  
User-agent: AdsBot-Google  
Disallow: /private/
```

Dans cet exemple, on indique aux robots nommés *Googlebot* et *AdsBot-Google* de ne pas explorer le contenu du répertoire *private*.

Il faut bien comprendre que le fichier *robots.txt* donne des indications concernant l'exploration (le crawl) et non pas l'indexation des pages d'un site.

Pour obtenir plus d'information sur les robots.txt voici une documentation de Google plus détaillée : <https://developers.google.com/search/docs/advanced/robots/create-robots-txt?hl=fr>

Sitemap

Le sitemap d'un site est un fichier qui répertorie l'ensemble des pages d'un site pour en favoriser l'indexation par les moteurs de recherche. Un sitemap n'est pas indispensable pour un site de petite taille. Il est recommandé pour une prise en compte plus rapide des nouveaux contenus. Surtout quand il s'agit d'un site volumineux ou avec des mises à jour fréquentes.

Un sitemap se présente le plus souvent au format XML (pour eXtensible Markup Language, un format proche du HTML mais plus généraliste). Il est déclaré à l'aide du fichier *robot.txt* et se positionne à la racine d'un site comme ceci :

- https://www.arte.tv/static/sitemap_fr.xml

Voici une documentation complète sur la création d'un sitemap :

- <https://www.sitemaps.org/protocol.html>

Balises meta no-index

Les balises meta que l'on a abordées plus haut permettent de spécifier qu'une page ne soit pas indexée par les robots des moteurs de recherche.

A cette fin, vous pouvez ajouter cette balise dans le head des pages concernées :

```
<meta name="robots" content="noindex">
```

Attention: Pour qu'une page ne soit pas indexée, il faut que le robot puisse lire cette balise meta et donc "crawler" cette page. Il ne faut donc pas la bloquer par un fichier *robots.txt*.

Les liens nofollow

Il est également possible d'indiquer directement aux robots de ne pas suivre un lien de l'une de vos pages lors de l'exploration de celle-ci. Ceci se fait avec l'attribut **rel** avec la valeur **nofollow**, comme ceci :

```
<a href="https://www.example.com/" rel="nofollow">cliquez ici</a>
```

Concrètement, cela ne veut pas dire que la page ciblée par le lien ne sera pas indexée. En effet, la page cible peut être déclarée dans un sitemap ou bien il peut exister d'autres pages qui ont un lien vers elle.

Info: Généralement on utilise nofollow pour spécifier aux algorithmes des moteurs de ne pas donner de poids à ce lien dans la logique du Page Rank.

Expérience utilisateur

La grosse tendance dans l'évolution du positionnement des sites dans les moteurs de recherche et de Google en particulier est la prise en compte de l'expérience utilisateur et de la satisfaction des internautes.

Cela consiste à considérer que si un internaute a été satisfait par une page pour une recherche donnée, alors cette page semble pertinente pour cette recherche.

Pensez au mobile

L'expérience utilisateur sur mobile est primordiale : la part du trafic internet généré avec les smartphones est désormais supérieure à celle générée par les ordinateurs. Les sites n'offrant pas une interface navigable satisfaisante en version mobile sont désormais pénalisés dans leurs positionnement par Google.

Performance d'un site

Dans le même état d'esprit, une grande importance est désormais donnée à la performance des sites internet dans les approches SEO et d'expérience utilisateur.

C'est à dire qu'un site qui s'affiche plus rapidement et qui mettra moins de temps à se charger sera favorisé dans les résultats de recherche.

Il existe beaucoup d'outils en ligne permettant de mesurer la performance d'un site et d'en connaître les pistes d'amélioration :

- <https://pagespeed.web.dev>
- <https://www.dareboost.com>

La performance des sites web est un sujet technique à part entière qui mérite que l'on s'y intéresse au-delà de son impact SEO.

Les outils SEO

Google Search Console

Google met à disposition la Google Search Console, qui permet d'avoir une vision globale de l'indexation de vos pages par Google, d'obtenir des statistiques, d'obtenir des alertes en cas de problème. Cet outil est disponible ici :

- <https://search.google.com/search-console/welcome>

Notez qu'il n'est pas nécessaire de s'inscrire à la Search Console pour que votre site apparaisse dans les résultats de recherche de Google.

Lighthouse

Google met également à disposition Lighthouse : un outil d'analyse SEO dans son navigateur Google Chrome qui permet de détecter les principales failles d'un site internet.