

SEO (Search Engine Optimisation)

Pour qu'un site remonte dans les recherches d'un moteur de recherche sans être en page 2 ou plus loin, il faut respecter plusieurs détails:

Balise Meta

Dans l'en-tête d'une page HTML nous pouvons trouver des balises "meta", elles sont lu par les navigateurs et les robots des moteurs de recherche pour comprendre le contenu de la page.

Certaines vont être plus ou moins utile mais quoi qu'il arrive toute celles placées doivent respecter le but dans lequel elles ont été créés.

- **description** doit décrire votre page sans être trop longue (160 caractères max est une bonne habitude)
- **keyword** doit contenir les mots clefs correspondant à votre page (google et certains autres ne l'utilisent plus, allant chercher les mots clef directement dans le contenu de la page)
- **viewport** permet d'adapter plus facilement votre site à différent format
- **og:***** les différentes balises meta og permettent d'aider les réseaux sociaux à afficher une vignette de votre site.
 - **og:site_name** indique le nom du site
 - **og:title** indique le titre de la page
 - **og:url** le lien de la page
 - **og:description** la description à afficher sur les réseaux sociaux
 - **og:image** le lien de l'image à afficher sur les réseaux sociaux
 - **og:locale** indique la langue de base du la page
 - **og:locale:alternate** indique les autres langues disponible
 - **og:type** indique si c'est un site web, un article, une vidéo...
 - Il y en a d'autres, même parfois certaines spécifiques à certains réseaux sociaux. Pour plus d'information voir "Open Graph".

Title et bonne pratique

Les robots des moteurs de recherche sont de plus en plus pointilleux sur les bonnes pratique HTML. Une page mal organisé ou avec des arguments manquant est moins bien référencé. Quelques rappels pour l'exemple :

- L'attribut "**alt**" sur les balises **img** n'est pas optionnel.
- Un **title** clair et concis dans le head.
- Des balises de titre **h1** à **h6** claires et concises, respectant l'ordre d'importance (Et selon certaines sources, un seul **h1**)
- Et ce ne sont que certains exemples, l'utilisation de W3C validator ou de l'extension lighthouse peut vous aider à trouver les problèmes de votre site.

Meta robots et robots.txt

Il peut arriver que vous ne souhaitiez pas que certaines pages de votre site soit indexé par les robots des moteurs de recherche, ou alors que vous ne vouliez pas que certains robots accèdent à votre site.

Il y a pour cela deux solutions:

- Le fichier **robots.txt** qui doit être à la racine de votre site (attention si vous avez un routeur qu'il n'en bloque pas l'accès)
- Et la balise **meta robots**.

On notera quand même que cela fonctionne que pour les robots qui suivent les bonnes pratiques du web. Si un robot a été conçu sans ces bonnes pratiques, ces éléments ne serviront à rien.

Le fichier **robots.txt** va empêcher l'exploration de votre site par les robots, mais n'empêche pas l'indexation dans le moteur de recherche. Il contiendra de courtes informations indiquant quel robot n'a pas accès à votre site et quels sont les pages inaccessibles.

Ici on interdit à tous les robots de visiter tout notre site :

```
User-agent: *  
Disallow: /
```

Et là on bloque l'accès au robot de google à deux chemins du site :

```
User-agent: Googlebot  
Disallow: /admin/  
Disallow: /profile/
```

La balise **meta robots** permet d'indiquer aux robots des moteurs de recherche si la page actuellement visitée doit être indexée dans le moteur de recherche ou non. Ici le robot peut visiter la page mais doit ou ne doit pas la faire apparaître dans le moteur de recherche.

Pour indiquer cela au robot on écrira :

```
<!-- Valeur par défaut, la page est indexée, et le robot a le droit de suivre les  
liens dans la page -->  
<meta name="robots" content="index, follow">  
<!-- La page n'est pas indexée et le robot n'a pas le droit de suivre les liens  
dans la page -->  
<meta name="robots" content="noindex, nofollow">
```

Ce cours ne se veut pas complet, vous trouverez d'autres informations sur la gestion des robots ici :

<https://robots-txt.com>

Plus de détail sur le fichier robot.txt, sur la meta robots ou encore sur le fichier **sitemaps.xml** qui permet d'aider les robots à explorer votre site.

Conclusion

Le SEO est important pour tout site internet qui veut être connu et reconnu. Mais c'est aussi quelque chose de complexe qui évolue d'année en année.

Prenons l'exemple de la meta description, il y a 20 ans on conseillais d'y placer le plus de mot clef populaire possible. Maintenant si on fait du hors sujet, les moteurs nous descende. La meta keywords était très importante il y a quelques années, maintenant les robots cherchent les mots clef directement dans notre contenu.

La meilleur façon d'avoir un bon SEO est de continuer à s'informer régulièrement.