



# Cours d'intégration

## 3W Academy

### Notions de référencement

---

#### [Présentation](#)

[Le SEA](#)

[Le SEO](#)

#### [Fonctionnement des moteurs de recherches](#)

[Le cas du PageRank de Google](#)

[Importance du title](#)

#### [Améliorations techniques](#)

##### [Les éléments du head du document HTML](#)

[Le title](#)

[Le charset](#)

[Les balises META](#)

[La balise META robots](#)

[La balise META description](#)

[La balise META google-site-verification](#)

[Les balises LINK](#)

[La balise LINK canonical](#)

[La balise LINK favicon](#)

##### [Les fichiers supplémentaires du site](#)

[Le fichier robots.txt](#)

[Le fichier sitemap.xml](#)

[Les fichiers RSS](#)

[Le fichier humans.txt](#)

[Le fichier .htaccess pour la configuration du serveur Apache.](#)

# Présentation

Le référencement peut être divisé en branches:

- Le référencement payant: Search Engine Advertising (SEA).
- Le référencement naturel: Search Engine Optimization (SEO).

## Le SEA

Le SEA utilise des techniques payantes pour améliorer le positionnement d'un site dans les résultats de recherche et pour sponsoriser le site sur canaux publicitaires. Cette partie n'est pas abordée ici.

## Le SEO

Le référencement naturel, c'est à dire, la visibilité et le positionnement du site sur les résultats naturels des moteurs de recherche. Le référencement naturel peut être amélioré de différentes manières.

# Fonctionnement des moteurs de recherches

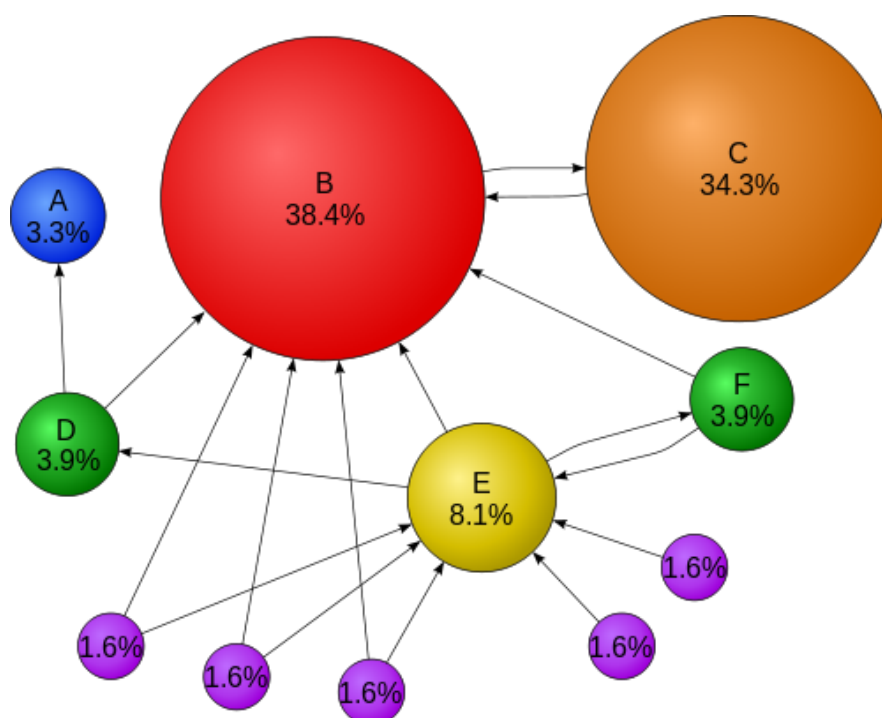
## Le cas du PageRank de Google

Pourquoi un site arrive en tête des résultats de recherche sur Google ? Un site arrive en tête des résultats de recherche sur Google si son PageRank est grand.

Le PageRank ou PR est l'algorithme d'analyse des liens concourant au système de classement des pages Web utilisé par le moteur de recherche Google. Il mesure quantitativement la popularité d'une page web. Le PageRank n'est qu'un indicateur parmi d'autres dans l'algorithme qui permet de classer les pages du Web dans les résultats de recherche de Google. Ce système a été inventé par Larry Page, cofondateur de Google. Ce mot est une marque déposée.

Le PageRank est plus grand si il existe un nombre plus important de liens pointant vers ce site que pour les sites concurrents dans le même domaine.

Le PageRank est également plus grand si les liens pointant vers le site proviennent de site ayant un PageRank plus grand.



*Exemple de poids relatif des sites en fonction du nombre de liens entrants*

Améliorer le positionnement d'un site dans un résultat de recherche revient améliorer son PageRank, c'est faire en sorte qu'il y ai plus de liens qui pointent vers ce site.

La meilleur méthode pour améliorer son positionnement est donc de de rendre son site attractif pour un sujet donné. Plus le site est apprécié par ses visiteurs, plus il y aura de chance que certains de ces visiteurs mettent des liens sur d'autres sites. Autrement dit, il faut privilégier la qualité du contenu mais aussi le design, l'ergonomie et l'accessibilité.



**Attention** Les référenceurs et les webmestres créent quelquefois massivement des liens retour, par échanges de liens ou en inscrivant un site sur une quantité d'annuaires. Google avait autrefois une vision quantitative de la popularité d'une page, cette technique permettait donc de gonfler artificiellement son indice de popularité.

Google a réagi: d'une part en instaurant des filtres, tels que la Sandbox. Il détecte et sanctionne les campagnes massives de liens artificiels, par ailleurs il intègre des critères qualitatifs à l'analyse des liens (sémantique, confiance : indice TrustRank, comportement des utilisateurs).

## Importance du title

Le title (dans le head) de chaque page est très important. Il est visible comme titre de l'onglet dans le navigateur. Il est repris par défaut comme titre du favoris. Il est visible comme titre dans les résultats de recherche.

Le title doit être court (moins de 100 caractères). Un titre trop long sera tronqué dans les résultats de recherche.

Il doit être compréhensible et bien refléter le contenu de la page. Il peut reprendre le nom du site, par exemple:

"Contactez nous - Mon super site" au lieu de "Contact".

# Améliorations techniques

## Les éléments du head du document HTML

### Le title

Comme vu plus haut, le titre du document est de loin l'élément le plus important à renseigner.

```
<title>Le titre de ma page offre une bonne description de son contenu</title>
```

### Le charset

Ne pas oublier de déclarer le charset, sous peine d'avoir des erreurs d'interprétation de l'encodage du texte par les moteurs de recherche et des erreurs d'affichage.

```
<meta charset="utf-8" />
```

### Les balises META

Les métadonnées du document HTML offre le moyen de spécifier certaines informations complémentaires.

### La balise META robots

La balise META robots permet d'indiquer aux robots des moteurs de recherche quoi faire:

- index = merci d'indexer ma page!
- follow = merci de suivre les liens dans la page.

```
<meta name="robots" content="index, follow" />
```

## La balise META description

La balise META description permet de déclarer une petite description de la page. Cette description peut être reprise par certains moteurs de recherche ou par certains annuaires. Elle ne doit pas être trop longue (environ 100 mots).

```
<meta name="description" content="Le description de ma page" />
```

## La balise META google-site-verification

Une balise META de vérification. Afin de prouver à Google que l'on est le propriétaire du site, il faut ajouter ce code fourni par Google.

```
<meta name="google-site-verification"  
content="6kUDFYb0tiSIH73nGQ7xtSAehBs8N6cjfdrmXBjBwBf" />
```

## Les balises LINK

Les balises LINK permettent de relier le document à des ressources externes. Le type de relation entre le document et la ressource est spécifié par l'attribut **rel**.

### La balise LINK canonical

La balise LINK de l'url canonique permet de spécifier l'url sous sa forme optimale représentante de la page. Elle permet dans certains cas de regrouper plusieurs sous pages accessibles via des paramètres dans l'url et ainsi d'éviter le "duplicate content" très mal vu par les moteurs de recherche.

```
<link rel="canonical" href="http://www.mon_site.fr/" />
```

### La balise LINK favicon

La balise LINK du favicon. Le favicon est le "logo" du site.

```
<link rel="icon" type="image/png" href="favicon.png" />
```

ou

```
<link rel="shortcut icon" type="image/x-icon" href="favicon.ico" />
```

## Les fichiers supplémentaires du site

### Le fichier robots.txt

Exemple de contenu d'un fichier robots.txt :

```
User-agent: *  
Disallow:  
Sitemap: http://www.mon-domaine.fr/sitemap.xml
```

Ici, il permet que tous les robots (directive User-agent: \*) peuvent tout indexer (par la négative car rien est indiqué dans la directive Disallow).

La directive Sitemap permet de donner l'url d'accès au fichier sitemap.xml.

### Le fichier sitemap.xml

Il permet de lister l'ensemble des page d'un site pour faciliter l'indexation par les moteurs de recherche.

Voici un exemple de fichier sitemap.xml

```
<?xml version="1.0" encoding="utf-8"?>  
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"  
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"  
  xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9  
    http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd">  
  <url>  
    <loc>http://example.com/</loc>  
    <lastmod>2006-11-18</lastmod>  
    <changefreq>daily</changefreq>  
    <priority>0.8</priority>  
  </url>  
</urlset>
```

Pour chaque url, on peut spécifier la date de dernière modification (lastmod), la fréquence des modifications (changefreq) et la priorité relative de la page par rapport aux autres pages du site. Inutile de spécifier une priorité élevée de toutes les pages, ça n'a aucune influence sur le positionnement dans les résultats de recherche avec les autres sites.

## Les fichiers RSS

Les fichiers RSS sont des fichiers XML permettant de décrire les flux de syndication RSS. il s'agit de la liste des articles du site mis à jour récemment. Des applications permettent de s'abonner à ces flux et de présenter une liste des nouveaux articles dans tous les sites préférés de l'utilisateur.

## Le fichier humans.txt

Il s'agit d'une initiative non officielle (réciproque du fichier robots.txt) qui permet d'indiquer le noms des personnes qui ont travaillées à la réalisation du site.

## Le fichier .htaccess pour la configuration du serveur Apache.

La gestion des redirections

Il est très important de bien gérer les redirections d'un site. En effet si vous modifiez le nom d'une page (et que son url d'accès se retrouve modifiée) bien référencée, vous risquez de perdre le référencement de cette page car il est lié à l'url.

Il faut donc indiquer au moteur de recherche que la page a changé d'url. Il y a plusieurs moyens mais pour un site simple hébergé sur un serveur Apache, le plus simple est de faire la déclaration dans le fichier de configuration d'Apache .htaccess à la racine du site.

Se référer à la documentation en ligne pour modifier un fichier .htaccess d'Apache.