

# Leçon 3 : Notions et enjeux et techniques de base du référencement web

*Cours UVCI  
GCW2104-1 : Leçon 3*

**DIABATE Nagbégna,**  
Enseignant - Chercheur

# Table des matières



<b>Objectifs</b>	3
<b>I - Section 1 : Généralités et enjeux du référencement Web</b>	4
1. 1. Définition et enjeux du référencement web .....	4
2. 2. Le référencement Naturel ou SEO .....	5
3. 3. Le référencement sponsorisé ou SEA .....	6
4. 4. La notion de visibilité et comparaison SEO/SEA .....	7
5. Exercice : Quiz de compréhension .....	9
<b>II - Section 2 : Moteurs de recherches et vocabulaire</b>	10
1. 1. Les moteurs de recherche .....	10
2. 2. Éléments de vocabulaire et lexicaux du référencement .....	11
3. Exercice : Quiz d'entraînement .....	13
<b>III - Section 3 : Principaux outils de référencement</b>	15
1. 1. Les mot-clés et l'indexation .....	15
2. 2. Les fichiers robots.txt .....	19
3. 3. Les sitemap .....	21
4. 4. Les webmaster tools .....	24
5. Exercice : Question pour vous ! .....	25
<b>Solutions des exercices</b>	27

# Objectifs

À la fin de cette leçon, l'apprenant sera capable de :

- **Expliquer** le vocabulaire et les enjeux du référencement ;
- **Décrire** le fonctionnement des moteurs de recherche.
- **Définir** la notion de mots-clés ; et,
- **Présenter** les principaux outils de référencement SEO.

# Section 1 : Généralités et enjeux du référencement Web



## 1. 1. Définition et enjeux du référencement web

### Définition

Pris dans un sens large, le référencement est l'action de référencer une chose, c'est à dire de la mentionner dans un système afin de pouvoir la retrouver grâce aux informations utilisées pour la référencer. Comme par exemple des produits dans un magasin, des informations, des fichiers dans un ordinateurs, etc...

Sur internet, le terme référencement a été repris pour définir l'action d'inscrire un site dans (sur) un moteur de recherche ou un "annuaire web". Il s'agit de faire en sorte d'apparaître, si possible, sur la première page affichée par Google, par exemple, lorsqu'un internaute effectue une recherche en lien avec le site web concerné par le référencement.

L'objectif principal visé dans un référencement, c'est bien de se faire connaître en étant rapidement trouvé dans les recherches des internautes. D'où le principal enjeux d'apparaître au moins sur la première page des résultats des recherches en liens avec votre page sinon, apparaître en première position des résultats. Aussi, faut-il que les mots ou phrase a saisir pour trouver votre page soient plus faciles et nombreux de sorte a vous trouver facilement !

Sur internet, on distingue deux types de référencement qui sont pratiquement distincts, même si cela n'a pas d'importance pour le visiteur ou l'utilisateur.

Le référencement dit payant ou sponsorisé (**SEA**) et le référencement dit naturel ou organique (**SEO**). Ces deux pratiques ne répondent pas aux mêmes règles, stratégies, techniques, ni aux mêmes besoins.

Une stratégie de référencement pour un site internet peut se faire uniquement sur la base du référencement naturel ou peut combiner les deux pratiques si besoin.

Par contre une stratégie de référencement n'intervenant que sur la technique du référencement payant n'est pas viable à long terme.

Le référencement naturel est indispensable pour la réussite et l'efficacité d'un site internet.

## 2. 2. Le référencement Naturel ou SEO

### Définition

« Le référencement naturel, ou SEO (Search Engine Optimization), englobe l'ensemble des méthodes et techniques qui visent à positionner les pages web de votre site internet dans les premiers résultats naturels des moteurs de recherche (notamment Google, mais aussi Bing, Yahoo et consorts). Cela dans le but d'améliorer la visibilité des pages en question, en appliquant une série de « règles » édictées par les moteurs de recherche. »

<http://www.eskimo2.fr/referencement-naturel/>

Le référencement fait donc appel à différentes techniques comme l'optimisation du contenu avec l'utilisation de mots-clés, la mise en place de liens hypertextes pointant vers le site web, etc. Les moteurs de recherches indexent en effet les fichiers suivant un algorithme complexe qui détermine leur pertinence pour ensuite ressortir ces fichiers dans un ordre logique lorsqu'un internaute effectue une recherche. Pour ressortir efficacement sur la page des résultats de recherches, il faut donc comprendre comment fonctionne cet algorithme (en général les moteurs de recherche) afin d'optimiser le site en fonction.

L'optimisation se fait à deux niveaux : le référencement On Site et le référencement Off Site.

#### *a) Le référencement On Site*

Ce sont les pratiques intervenant directement sur le site à référencer. On y retrouve l'optimisation technique du site, l'optimisation de la structure, l'optimisation du contenu..

Le référencement naturel on-site regroupe tous les éléments qui ont trait aux pages web elles-mêmes :

- Les balises html (title, meta description, niveaux de titres Hn, attribut alt pour les images...)
- Les extrait enrichis (rich snippets)
- La structure des URLs
- Les temps de chargement
- Le maillage interne du site
- Les fichiers robots.txt et sitemap
- La pertinence des mots-clés
- Le placement des mots-clés dans les contenus et leur densité
- Le champ sémantique autour des mots-clés
- La qualité et l'originalité du contenu
- L'aspect mobile friendly – Responsive

Ces éléments impactent plus ou moins le référencement naturel d'une page, mais ils contribuent au minimum à sa bonne indexation et/ou à l'amélioration de sa visibilité.

#### *b) Le référencement Off Site*

Ce sont les pratiques intervenant en dehors du site à référencer, mais avec pour objectif de favoriser indirectement son référencement. On parle essentiellement de stratégie de liens entrants, de **netlinking** ou d'acquisition de **backlinks**. Si des sites B, C, D font un lien vers un site A, l'algorithme du moteur de recherches va estimer que le site A est pertinent car d'autres sites le mentionnent.

Le site A va ainsi améliorer son référencement. Plus les sites faisant des **backlinks** vers le votre sont des sites biens réputés, avec une forte autorité, plus leurs liens auront un impact positif sur le site A.

Le Netlinking concerne donc les éléments extérieurs aux pages web :

- La popularité du site
- Le netlinking et les backlinks (les liens qui pointent vers vos pages en provenance d'autres sites)
- Le benchmarck concurrentiel
- Les liens sociaux

**Le moteur du référencement naturel Google off-page est la qualité** : une page qui plaît aux internautes obtiendra plus facilement des liens entrants et verra sa popularité s'accroître ; 10 backlinks en provenance de quelques sites influents sont plus importants que 100 backlinks venant de sites ayant mauvaise réputation, etc.

### 3. 3. Le référencement sponsorisé ou SEA

#### Définition

Le référencement payant, appelé aussi SEA (Search Engine Advertising) consiste, à acheter la visibilité de son site web sur un moteur de recherche donné à travers un système d'enchère qui définit sa position en marge de la page de résultat. Il s'agit, en pratique, de liens publicitaires placés sur les pages de résultats des moteurs de recherches, ou même sur des sites partenaires de régies publicitaires.

Lorsque l'on parle de référencement payant, on pense donc aux liens sponsorisés ou aux campagnes Google Ads. Pour apparaître dans les annonces Google, il faut acheter des mots clés aux enchères. Le coût payé dans ce système se fait au clic. C'est à dire que le site publiant son annonce va payer l'enchère prévue à chaque fois qu'un internaute clique sur l'annonce.

Lorsque je navigue sur Google après avoir fait une recherche, je vois apparaître des résultats comportant un cartouche jaune sur lequel il est écrit "**annonce**". Il s'agit donc de liens sponsorisés.



On met généralement en place une stratégie de référencement payant pour :

- Les sites e-commerce
- Se positionner très rapidement
- Le lancement d'un nouveau site encore mal référencé naturellement
- Une période spéciale de l'année
- Se positionner sur une recherche très concurrentielle
- Compenser une mauvaise position dans les résultats naturels

### Remarque

Les annonces et liens sponsorisés permettent d'optimiser et de travailler votre référencement naturel. Il est par exemple impossible d'avoir un bon référencement naturel du jour au lendemain sur une nouvelle offre d'assurance, alors que grâce à la publicité en ligne on peut avoir une visibilité en quelques minutes ou heures. Mais l'intérêt pour le SEO réside dans le fait que les outils statistiques (Google Analytics et autres) vont vous permettre de savoir comment optimiser votre contenu. Et ce grâce aux expressions que les internautes ont recherché sur les moteurs de recherche notamment.

### Complément

En vous penchant du côté du marketing digital ou de la publicité en ligne, on vous parlera de deux autres techniques de référencement que nous n'allons pas étudier dans le cadre de ce cours. Il s'agit du :

**SEM** : Le « Search Engine Marketing » comprend le SEO et le SEA. Il regroupe ainsi à la fois les techniques de référencement payant et naturel pour obtenir du trafic sur un site à partir des moteurs de recherche.

**SMO** : « Social Media Optimisation » en anglais et « Optimisation des Médias Sociaux » en français, fait référence au travail d'amélioration de la visibilité d'un site au travers des réseaux sociaux.

## 4. 4. La notion de visibilité et comparaison SEO/SEA

### *La notion de visibilité*

« La notion de référencement ne peut pas être considérée de façon autonome. Elle est indissociable d'une notion capitale : la notion de visibilité.

Si la notion de référencement définit l'action d'être référencé et donc présent dans une base de donnée, la notion de visibilité elle définit la capacité de la chose référencée à être trouvée. C'est une notion essentielle à toujours avoir en tête lorsque l'on parle de référencement. Imaginez un grand magasin sur 3 étages, votre produit y est référencé, mais au fond d'une allée du troisième étage au niveau du sol et en un unique exemplaire. Vous conviendrez que l'utilité de la chose est très réduite. Alors que si votre produit est référencé en tête de gondole prêt de l'entrée du grand magasin, c'est une toute autre histoire.

Le référencement sur internet fonctionne de la même façon. Les 3/4 des internautes ne visiteront même pas la seconde page de résultats de recherches. Lorsque vous pensez référencement, vous devez donc penser visibilité avec pour objectif d'être présent en première page des résultats de recherches pour des mots clés pertinents, en rapport avec votre activité et votre clientèle. » <https://www.seo.fr/definition/comprendre-referencement>

On revient très souvent à se poser la question de savoir lequel choisir entre le SEO et SEA pour être plus visible sur Internet.

En effet, le SEO et le SEA ont chacun des caractéristiques particulièrement pertinentes pour donner de la visibilité à un site internet. Pour simplifier la question, on peut se baser sur un tableau :

	SEO – Référencement naturel	SEA – Référencement payant
<b>Temps passé</b>	Long. Vous devez mettre en place plusieurs optimisations sur des domaines variés (netlinking, technique, contenu...)	Relativement court. Le paramétrage et le suivi de campagnes varient en fonction de la concurrence.

<b>Rapidité du positionnement</b>	Long. Google valorise les optimisations sur le moyen terme (plusieurs mois parfois)	Très court. Le système d'enchères est quasi-immédiat.
<b>Retour sur Investissement</b>	Long et moyen terme. Les positions acquises sont à suivre et à nécessite un entretien (liens notamment)	Court terme. Dès la fin de la campagne, vous disparaissiez.
<b>Budget</b>	Il s'agit ici de passer du temps, qui a un coût forcément !	Coûteux. Plus la concurrence est forte, plus vous devrez dépenser de l'argent.
<b>Contenus</b>	Plus les contenus sont riches et pertinents mieux ils se positionneront	Plus les contenus de la landing pages sont pertinents, plus le quality score sera élevé ce qui fera baisser le prix de vos campagnes !

**Source :** <https://www.lere-position.fr/blog/seo-vs-sea-referencement-naturel-ou-referencement-payant/>

Ainsi, que cela a été remarqué plus haut, les deux méthodes peuvent être utilisée de façon complémentaire selon la stratégie mise en place.



## 5. Exercice : Quiz de compréhension

*[solution n°1 p.27]*

### Exercice

---

1) On ne parle de référencement que dans le cadre des sites web

☐ Vrai

☐ Faux

### Exercice

---

2) Qu'il s'agisse de SEO ou de SEA, les deux pratiques répondent :

☐ aux mêmes règles

☐ aux mêmes stratégies

☐ aux mêmes techniques

☐ aux mêmes besoins

### Exercice

---

3) Le SEO peut aussi être payant

☐ Vrai

☐ Faux

### Exercice

---

4) Quel type de référencement est indiqué pour un site web de e-commerce ?

☐ Référencement Naturel

☐ Référencement payant

### Exercice

---

5) L'ensemble des méthodes et techniques qui visent à positionner les pages web de votre site internet dans les premiers résultats naturels des moteurs de recherche dans le but d'améliorer la visibilité des pages en question s'appelle :

☐ le SEO

☐ le SEA

# Section 2 : Moteurs de recherches et vocabulaire

## II

### 1. 1. Les moteurs de recherche

Il existe des milliards de pages enregistrées sur des millions de sites Web répartis sur toute la planète. Chaque page contient des informations appelées (mots clés) qui décrivent son contenu.

#### Définition

Un moteur de recherche est un service en ligne permettant de trouver facilement une page ou une ressource sur le Web grâce à un ou plusieurs mots-clés renseignés décrivant son contenu que l'utilisateur saisi dans un formulaire de recherche.

En pratique, un moteur de recherche c'est comme votre GPS pour vous déplacer sur le réseau Internet, qui est comparable à un réseau routier. Il vous guidera dans vos déplacements et vos recherche sur Internet.

Un **moteur de recherche** fonctionne grâce à un grand nombre de serveurs appelés **robots** qui parcourent les sites Web existants à intervalle régulier pour découvrir les liens des milliards de pages Web et leurs description. Chaque page identifiée est alors **indexée** dans une **base de données**, accessible ensuite par les internautes à partir de **mots-clés**.



Source : <http://www.coursinfo.fr/wp-content/uploads/2016/11/moteur-de-recherche-1.png>

Lorsqu'un internaute utilise un moteur de recherche :

1. Il saisit les mots clés décrivant les pages qu'il recherche, par exemple : « *cours informatique gratuit et sans pub* », soit directement dans **la barre d'adresse et de recherche** soit sur le **site web du moteur de recherche**;
2. Le navigateur prépare ensuite un **formulaire de recherche** contenant ces mots-clés et l'**envoi au moteur de recherche**;
3. Le moteur de recherche **interroge sa base de données** pour chacun des mots-clés puis affine la recherche en enlevant les pages ne convenant pas ;
4. La base de données répond avec une **liste de résultats** contenant les liens vers les pages adéquates
5. Le moteur de recherche retourne au navigateur cette liste de résultats contenant liens vers des pages, avec soit le début du texte de la page, soit le texte spécifié par le créateur de la page grâce aux balises spécifiques, appelées méta-tags, ou encore l'extrait de la page qui contient les mots recherchés.
6. Le navigateur affiche la liste de résultats à l'Internaute.

### *Rappel : Le cours d'initiation à l'informatique*

Cette vidéo introductive vous explique ce que c'est qu'un moteur de recherche et son comment il fonctionne.

Cf. ""

## 2. 2. Éléments de vocabulaire et lexicaux du référencement

Le référencement web ne se limite plus aux seuls aspect techniques dévolus aux webmestres. Il est devenu aujourd'hui un outil efficace utilisé dans marketing autour duquel se développent plusieurs stratégies et métiers.

Les professionnels du marketing digital utilisent un vocabulaire technique ponctué d'abréviations, de sigles et d'anglicisme qui vous semble parfois déroutant.

En tant que développeur, il convient de comprendre ce jargon du SEO afin de pouvoir converser et adapter vos développement aux besoins des marketeurs.

Toute une multitude de thème et d'abréviations sont utilisés, se basant souvent sur des anglicismes. Nous allons, dans le cadre de ce cours, vous présenter les principaux termes liés au référencement naturel :

- **Positionnement** : Désigne la position d'une page Web dans les pages de résultats des moteurs de recherche.
- **Mots-clés** : Les mots-clés sont les expressions rentrées sur un moteur de recherche par un internaute pour trouver un produit/service/autre.
- **Trafic (en terme de référencement)** : En référencement, le trafic d'un site Internet se compte au nombre de ses visites.
- **Duplication interne** : Contenu identique présent sur plusieurs pages d'un même site.
- **Duplication externe** : Contenu identique présent sur plusieurs pages de différents sites.
- **Indexation** : L'on dit d'une page Web qu'elle est indexée lorsque celle-ci est enregistrée dans l'index d'un moteur de recherche ou d'un annuaire en ligne.
- **Algorithme** : L'algorithme d'un moteur de recherche désigne les méthodes de calculs utilisées par ce dernier pour déterminer le classement d'une page Web selon sa pertinence.

- **Annuaire** : Un annuaire est un outil de recherche qui recense et stocke dans une base de données les sites Web classés par catégories et thématiques.
- **URL** : Une URL, « Uniform Resource Locator », fait simplement référence à l'adresse unique d'un document sur le Web, à savoir celle qui est saisie dans la barre d'adresse du navigateur Web.
- **SERP** : « Search Engine Results Page », il s'agit là des pages de résultats affichées par les moteurs de recherche.
- **Ancre (en terme de référencement)** : En référencement, une ancre représente le texte visible d'un lien sur lequel on peut cliquer.
- **Backlinks** : Les backlinks sont des liens externes, provenant d'un autre site Web, et pointant vers les pages de votre propre site.
- **Netlinking** : Désigne l'ensemble du travail réalisé visant à obtenir et optimiser des liens pointant vers les pages de votre site Web, afin d'en améliorer le référencement.
- **PageRank** : Il s'agit en quelque sorte de l'appréciation donnée par Google à une page Web : indice de Google mesurant, de 0 à 10, la popularité des pages en fonction du nombre et de la qualité des liens qui pointent vers elles.
- **CPC** : Le « Coût Par Clic » représente le montant qu'un annonceur paye aux moteurs de recherche à chaque fois qu'un internaute clique sur son lien, dans le cadre d'une campagne SEA.
- **CTR** : « Clic Through Rate » est le ratio entre le nombre d'affichages et le nombre de clics d'une annonce payante. Il mesure l'attractivité de votre annonce pour les internautes.
- **Taux de transformation** : C'est le rapport entre le nombre de visiteurs ayant effectué une action donnée (achat d'article, téléchargement d'un logiciel, demande de devis, etc.) sur le nombre total de visiteurs.
- **Taux de rebond** : Pourcentage d'internautes ayant quitté un site après y avoir consulté qu'une seule page Web.
- **CPA** : Le « Coût Par Action » permet de mesurer la rentabilité d'une campagne SEA : rapport entre le budget dépensé et le nombre d'actions.
- **Tracking** : Ensemble d'outils permettant d'étudier les comportements des internautes sur un site Web, définir des profils type de visiteurs ou encore connaître les sources de trafic.
- **Programme d'affiliation** : Le but de ce programme est principalement d'augmenter sa notoriété ou de produire des contacts commerciaux qualifiés. Le principe : l'annonceur rémunère un site affilié pour la mise en place de campagnes à la performance (CPA).
- **Liens Nofollow** : A contrario d'un lien normal qui est « dofollow », un lien nofollow comporte un attribut qui donne comme indication aux moteurs de recherche de ne pas prendre en compte ce lien.

Cette liste, inspirée du blog : <https://leblog.powerboutique.com/lexique-lessentiel-du-vocabulaire-du-referencement-explique/>, n'est pas exhaustive, vous pouvez avoir plus de définitions et d'explications sur les liens suivants :

- [<https://leblog.powerboutique.com/lexique-lessentiel-du-vocabulaire-du-referencement-explique/>]<https://www.neoptimal.com/blog/glossaire-seo>
- <http://www.superbibi.net/glossaire-webmarketing>
- <https://leblog.powerboutique.com/lexique-lessentiel-du-vocabulaire-du-referencement-explique/>

### 3. Exercice : Quiz d'entraînement

*[solution n°2 p.28]*

#### Exercice

---

1) Trouvez l'intrus

- ☐ Bing
- ☐ Gmail
- ☐ Google
- ☐ Quant

#### Exercice

---

2) Un mot clé, c'est :

- ☐ Un dessin montrant un mot avec une clé
- ☐ Mot qu'on utilise pour faire une recherche
- ☐ Un mot avec une clé dessus

#### Exercice

---

3) Si je connais l'adresse URL d'un site, je ne suis pas obligé d'utiliser Google pour accéder à ce site

- ☐ Vrai
- ☐ Faux

#### Exercice

---

4) Que fais le moteur de recherche Google lorsqu'il ne trouve pas exactement ce qu'il recherche ?

- ☐ Il ferme le navigateur
- ☐ Il se plante
- ☐ Il vous propose des résultats de recherche similaires ou proches de ce que vous cherchez
- ☐ Il vous déconnecte de l'Internet et vous envoie une notification

#### Exercice

---

5) Avec lesquelles des connections Internet on ne peut pas faire de recherches ?

- ☐ 3G+
- ☐ ADSL
- ☐ Wimax
- ☐ 4G

- ☐ 3G
- ☐ Aucune, on peut faire des recherches avec toutes ces connexions

# Section 3 : Principaux outils de référencement

III



[https://www.aube.lu/wp-content/uploads/2018/01/referencement\\_site.jpg](https://www.aube.lu/wp-content/uploads/2018/01/referencement_site.jpg)

Le référencement web est une aventure, oui une aventure qui va allier stratégie, technique et technologie. Il n'est donc pas possible de vous donner une solution miracle ou la démarche à zéro échec pour faire un référencement ; c'est ensemble d'outils et de techniques que chacun aura à combiner pour booster la visibilité de ses ressources sur le web. Il faut savoir qu'un référencement web ne s'improvise pas, il se base sur une méthode et une stratégie. dans le cadre de ce cours, nous allons vous donner quelques outils et techniques de référencement que vous pourrez utiliser pour bâtir votre stratégie et votre méthode !

## 1. 1. Les mot-clés et l'indexation

### Définition : Indexation

L'**indexation** d'un site internet est son enregistrement dans la base de données d'un moteur de recherche (Google, Yahoo, Bing...) ou d'un annuaire.

Elle se fait par l'intermédiaire d'un formulaire de soumission comprenant un certain nombre d'informations (souvent: url du site, description, mot clé...).

L'indexation permet entre autre de ressortir dans les pages de résultats de requêtes faites par les utilisateurs.

## Définition : Mot-clé

Wikipédia<sup>[<https://fr.wikipedia.org/wiki/Mot-cl%C3%A9>]</sup> : « Un mot clé (orthographié aussi **mot-clé**, mot clef ou **mot-clef** est un mot ou un groupe de mots qui a une importance particulière permettant de **caractériser le contenu d'un document et permettant une recherche d'informations**. Une liste de mots-clés permet ainsi de définir les thématiques contenues dans un document. »

### a) Les mot-clés



On peut définir facilement en considérant les activités de référencement web comme étant **un mot ou suite de mots tapés ou dictés par un internaute pour effectuer une recherche**. Ainsi :

- En tant qu'internaute, quand vous effectuez des recherches, vous tapez une expression clé composée d'un ou plusieurs mots en fonction des résultats que vous souhaitez obtenir du moteur de recherche.
- En tant qu'éditeur ou propriétaire de site internet, votre objectif est d'apparaître en bonne position dans le résultats de moteurs de recherches quand un internaute tape une expression clé pour vous trouver.

Le choix des mots clés est une étape décisive de votre stratégie de référencement naturel. Pour trouver un mot clé, vous devez vous mettre à la place de votre visiteur ou éventuel prospect. Demandez-vous ce qu'il taperait dans les moteurs de recherche et le mènerait jusqu'à votre site ! Il est donc essentiel de connaître les habitudes des internautes afin de dénicher les bons mots clés.

Il revient donc qu'on ait à se poser la question de savoir **ce qu'est un bon mot-clé**.

L'époque où il suffisait de renseigner la balise meta "keywords" de pleins de mots clés est désormais révolue surtout avec l'intégration de l'intelligence artificielle (*Algorithme de RankBrain*<sup>[<https://www.anthedesign.fr/referencement/rankbrain-intelligence-artificielle-critere-seo-google/>]</sup> (<https://www.anthedesign.fr/referencement/rankbrain-intelligence-artificielle-critere-seo-google/>)).

Un bon mot clé devrait répondre aux critères suivants :

**La pertinence :** le ou les mots clés doivent décrire au mieux votre activité et être utilisés par vos clients. Il doit refléter votre activité, les produits ou services que vous proposez.



**Le potentiel :** le mot clé doit générer du trafic qualifié sur votre site, sinon son potentiel sera nul. Parmi tous les mots clés envisageables, il est évidemment préférable de choisir celui qui est le plus souvent recherché dans les moteurs de recherche.

**La concurrence :** si la concurrence est trop forte sur un mot clé, il faudra envisager de trouver une variante et d'exploiter la longue traîne. Aujourd'hui, pour certains mots clés, les moteurs de recherches ont déjà référencé des dizaines ou des centaines de millions de pages. Si beaucoup de ces pages sont déjà bien optimisées pour le référencement, il risque d'être techniquement long et difficile d'atteindre les premières places. Bien que rien ne soit impossible, il est souvent préférable, dans un tel contexte, de privilégier des mots clés plus « **accessibles** »

### *Remarque : Comment trouver des idées de mots clés pour votre site ?*

---

- Votre propre connaissance de vos produits/services, de votre site
- Le site des concurrents
- Les suggestions Google dans les SERPs
- L'*outil de planification des mots clés Google* [<https://adwords.google.fr/KeywordPlanner>] (ex. générateur de mot clé Google) qui fournit non seulement des idées de mots clés (soit à partir d'une liste de base soit à partir d'une URL) mais aussi des idées des volumes de recherches mensuelles pour chaque mot clé. Il permet ainsi de déterminer combien de fois un mot est recherché chaque mois dans Google.

Source : <https://www.olivier-corneloup.com/ressources/dico-referencement/definition-mot-cle/>

### *Complément : Mots-clés LSI*

---

L'indexation sémantique latente est un procédé de traitement des langages naturels que les moteurs de recherche, comme Google, utilisent pour déterminer le sens d'un contenu.

C'est pourquoi l'utilisation des mots connexes à votre mot-clé cible sur un contenu ou l'ensemble de votre site Web ajoute du contexte, aidant ainsi Google à comprendre plus facilement le sens de votre contenu.

Et lorsque Google comprend mieux votre contenu, il peut mieux le présenter aux internautes qui vous recherchent.

Wikipédia nous donne la définition suivante : "**Les mots-clés LSI (indexation sémantique latente) sont des mots-clés liés sémantiquement à votre mot-clé principal**".

Cependant, il ne faut pas confondre les mots-clés LSI et les synonymes de ces mots.

Les mots-clés LSI sont tout simplement des mots trouvés fréquemment ensemble puisqu'ils partagent le même contexte.

Par exemple, "apple" et "itunes" sont des mots-clés LSI. En effet, ils sont souvent associés sans être des synonymes.

## b) Indexation



<https://optimiz.me/wp-content/uploads/2014/12/google-search.png>

De manière très simplifiée, le fonctionnement d'un moteur de recherche comme Google peut être expliqué en 4 grandes étapes :

- La phase de crawl ou phase de collecte des données
- La phase d'indexation
- La phase de traitement des données
- La phase de restitution

« L'indexation désigne l'action du **robot** (c'est-à-dire d'un programme) d'un moteur de recherche qui passe sur un site, le **parcourt** (phase dite de « crawl » du verbe « parcourir » en anglais) et **indexe** son contenu. Lorsque l'on dit que le robot (le spider) indexe un site, cela signifie qu'il visite le site, en copie le contenu et le stocke dans les serveurs du moteur de recherche. » <https://optimiz.me/indexation-comprendre-le-fonctionnement-de-lindex-google/>

Cependant, plutôt que d'attendre que les moteurs de recherche au cours des crawl automatiques découvrent votre site, vous pouvez le soumettre vous même pour indexation aux moteurs de recherches (comme une déclaration de naissance de votre site !) ; cette méthode est d'ailleurs conseillée.

Pour soumettre votre site, chaque moteur de recherche a sa méthode, vous avez par exemple :

- Soumettre le site à Google : <http://www.google.fr/addurl/> ou utiliser le *Google Search Console* [<https://search.google.com/search-console>]
- Suggérer un site à Yahoo et à Bing : <http://www.bing.com/toolbox/submit-site-url>

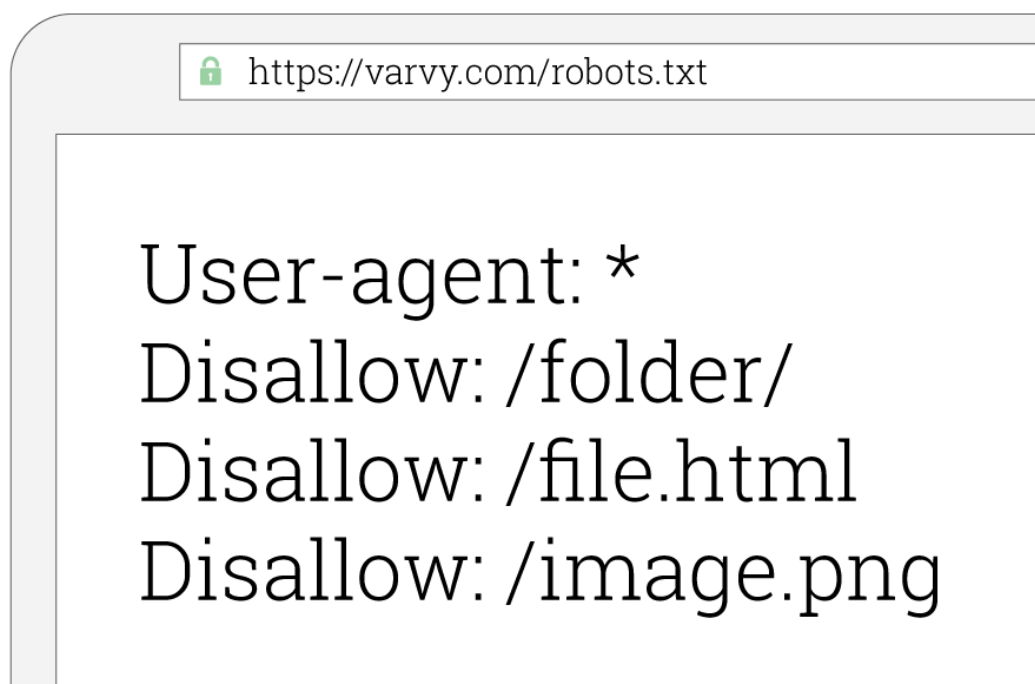
### Remarque : Supprimer un site de l'index de Google

Supprimer un site de l'index Google revient à supprimer le site de la base de données des sites connus par Google. Ainsi, ses pages ne seront plus prises en compte par le robot lors de la prochaine exploration du web :

**Suppression en urgence** : ajoutez les URL à supprimer dans un fichier sitemap et utilisez la balise <expires>.

Dans les **Google Webmaster Tools** : rendez-vous sur "Index Google" puis dans la section "URL à supprimer".

## 2. 2. Les fichiers robots.txt



*https://varvy.com/images/robots-txt-defined.png*

Pour référencer un site, les moteurs de recherche envoient de façon régulière leurs robots pour indexer les nouveaux contenus de ce site. Ces robots constituent donc le trait d'union entre les moteurs de recherche et les sites web.

Ces robots sont simplement des petits programmes qui détectent qu'un nouveau contenu a été créé ou qu'un contenu existant a été modifié.

### Définition

Un fichier **robot.txt** est un fichier unique qui indique aux robots des moteurs de recherche ce qu'ils doivent et ce qu'ils ne doivent pas indexer. Par convention, les robots des différents moteurs de recherche consultent le fichier **robots.txt** avant d'indexer un site Web. Placé à la racine du site web, il permet de communiquer des instructions à aux robots des moteurs de recherches.

**Le fichier robots.txt est en fait un protocole dit d'exclusion des robots.** C'est une convention qui vise à empêcher les robots d'exploration (web crawlers) d'accéder à tout ou une partie d'un site web.

Le fichier robots.txt, contient une liste de ressources du site qui ne sont pas censées être explorées par les moteurs de recherches. Il permet également d'indiquer aux moteurs l'adresse du fichier **sitemap.xml** du site.

Les cas d'utilisation courants incluent par exemple :

- Éviter l'exploration de contenu dupliqué
- Éviter l'exploration d'un moteur de recherche interne
- Empêcher les moteurs de recherche d'indexer certaines images de votre site
- Empêcher les moteurs de recherche d'indexer certains répertoires de votre site
- Spécifier l'emplacement du plan de site

## Fondamental

Dans un soucis de maintenance et de suivi des modifications apportés à un robots.txt, il est également possible d'ajouter des commentaires au sein du fichier.

Pour finir, deux choses à savoir lorsque l'on utilise un fichier robots.txt :

- Les robots peuvent ignorer votre fichier **robots.txt**, en particulier les robots malveillants qui crawlent les sites à la recherche de vulnérabilités ou d'adresses e-mail ;
- Le fichier **robots.txt** est un fichier accessible au public, ce qui signifie que n'importe qui peut voir ce que vous souhaitez ne pas faire indexer par les moteurs.

## Méthode : Syntaxe et écriture d'un fichier robots.txt

Ce fichier vous permet de laisser des instructions aux **crawlers**. Vous avez la possibilité :

- d'interdire ou autoriser l'exploration de votre site à certains **crawlers** (également appelés "**agents**" ou "**robots**");
- d'interdire ou autoriser l'exploration de certaines pages de votre site aux **spiders** ;
- d'interdire ou autoriser l'exploration de certaines pages à certains **crawlers** ;

Le fichier **robots.txt** est constitué d'un ensemble de blocs d'instructions et optionnellement de directives **sitemaps**.

Chaque bloc contient deux parties :

- **User-agent** : cette ligne indique quel robot est concerné par la règle ou l'interdiction. A quels robots ce bloc s'adresse. Le caractère \* signifie que cela concerne tous les robots.

Si par exemples on veut appeler le seul robot de Google, on peut écrire :

```
1 # Les commentaires commencent par les # Cette ligne est donc un commentaire
2 # Pour appeler que les robots de google :
3
4 User-agent: Googlebot
5 User-agent: Googlebot-News
6
7 # Pour tous les Crawlers:
8 User-agent: *
9
```

- **Disallow** : Permet d'interdire aux robots le crawl d'une portion du site ou répertoire. Ainsi, on peut écrire

```
1 # Interdire le repertoire CSS
2 Disallow : /css
3 # Exclure l'interface d'administration
4 Disallow : /administration
5
```

- **Sitemap** : Optionnelle, cette commande permet d'indiquer l'adresse à laquelle se trouve votre fichier de sitemap.

```
1 Sitemap: http://www.votre-site.com/sitemap.xml
```

Ne vous en faites pas pour l'instant avec les sitemaps ! La prochaine section du cours lui est consacrée.

Pour l'heure, voila le contenu d'un fichier **robots.txt** avec un sitemap qui accepte tous les bots, mais qui interdit l'exploration du répertoire d'administration :

```
1 User-agent: *
2 Disallow : /administration
3 Sitemap: http://www.votre-site.com/sitemap.xml
```

### Complément

Pour plus d'information sur les fichiers **robots.txt**, consultez les liens suivants qui vous détaillent le protocole d'exclusion ou le langage utilisé par les fichiers **robots.txt** :

- [<https://blog.smartkeyword.io/robots-txt/>]<https://blog.smartkeyword.io/robots-txt/>
- [https://craym.eu/tutoriels/referencement/fichier\\_robot\\_txt.html](https://craym.eu/tutoriels/referencement/fichier_robot_txt.html)

## 3. 3. Les sitemap

### Définition

Le sitemap est un fichier XML écrit pour décrire le plan de votre site. Plus précisément, c'est un fichier dans lequel sont listées les adresses de toutes les pages d'un site de façon structurée.

Présentée de cette façon, cette page permet aux robots des moteurs de recherche d'en lire efficacement le contenu.

Il est donc utile pour l'indexation des pages du site mais plus généralement pour suivre le référencement de votre site.

```
- <urlset>
- <url>
- <loc>
  http://www.e-monsite.com/pages/tutoriels/configuration-avancee-du-site/les-redirections-sur-vos-sites.html
- </loc>
- </url>
- <url>
  http://www.e-monsite.com/pages/tutoriels/trucs-et-astuces-pour-debuter/recuperer-les-urls-des-fichiers-de-l-espace-de-stockage.html
- </loc>
- </urlset>
```

<https://www.e-monsite.com/medias/images/sitemap-1.png>

En tant que fichier XML, le sitemap respectera une certaine structure comme indiquée dans l'exemple suivant :

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2
3 <urlset xmlns="http://www.google.com/schemas/sitemap/0.84">
4
5 <url>
6
7 <loc>https://balafon225.net</loc>
8
9 <lastmod>2016-05-12</lastmod>
10
11 <changefreq>daily</changefreq>
12
13 <priority>1.0</priority>
14
15 </url>
16
17 </urlset>

```

XML va donc utiliser un ensemble de balise, comme avec le HTML, pour lister toutes les pages de votre site web :

- **<Urlset>** : il s'agit d'un ensemble d'URLs qu'on va décrire dans le fichier sitemap
- **<Url>** : il s'agit d'une URL dont on va donner les détails
- **<Loc>** : la location (mot anglais, qui signifie l'emplacement en français) de l'URL
- **<Lastmod>** : la date de dernière modification de la page
- **<Changefreq>** : la fréquence de modifications de la page
- **<Priority>** : la priorité de cette page pour vous. 1.0 étant le maximum.

Le fichier sitemap permet d'indiquer la liste complète des pages de votre site, ainsi que des informations comme :

- Si cette page change souvent,
- Si elle est importante (prioritaire) pour vous,
- etc.

## Méthode

Il existe plusieurs manière manière de disposer d'un sitemap pour son site.

le sitemap, pour être lue par les robots doit être situé à la racine du site web (le htdocs, le www, ou souvent le public\_html) mais pas dans un autre dossier.

- Vous pouvez écrire vous même le contenu de votre sitemap. Cette méthode, est fastidieuse et coûteuse en ressources, ce qui n'est pas conseillé ;
  - Certains éditeur WYSIWYG permettent de créer des sitemap par simple clic (Adobde Dreamweaver);
  - Certains hébergeur propose la génération automatique des sitemap des sites que vous hébergez chez eux ;
  - Vous pouvez utiliser des outils en ligne pour générer votre sitemap
- A ce niveau, il existe plein de sites Internet pour créer votre sitemap XML à partir du nom de domaine du site du site
- <https://www.xml-sitemaps.com> génère gratuitement jusqu'à 500 pages
  - <http://www.web-site-map.com/> celui-ci est également gratuit avec une limite supérieure, de 3500 pages.
  - <https://xmlsitemapgenerator.org/sitemap-generator.aspx> qui inclut jusqu'à 2000 pages dans le sitemap.

Même si cela n'a pas grande importance, vous pourrez ensuite modifier le sitemap généré automatique une fois que vous l'avez téléchargé sur votre site.



## Complément

Le sitemap permet de donner une liste exhaustive des pages et fichiers à indexer dans les moteurs de recherche.

Si, par exemple, on laisse Google trouver par lui-même les pages à indexer, il peut très bien en oublier. Ou ne pas toutes les repérer. Le sitemap va donc les répertorier pour lui !

L'intégration avec la **Google Search Console** est très facile et permet d'accélérer le référencement naturel d'un site.

Enfin, créer un fichier sitemap est super facile.

La page suivante vous donne plus d'information sur le sitemap et ses intérêts : <https://www.yakaferci.com/genere-sitemap/>

#### 4. 4. Les webmaster tools

les outils de gestion des sites et plate-formes web communément appelés Webmaster Tools sont des plate-forme de suivi des sites mis en place pour la plupart par les moteurs de recherches à destination de concepteurs de sites web pour gérer leur sites dans l'environnement du moteur de recherche.

Parmi ceux-ci, on peut citer les plus rependus :

- Google Webmaster Tools ou **Google Search Console**
- **Bing Webmaster Tools**

Ce sont deux produits distincts mais qui repose sur le mêmes principes de fonctionnement a savoir :

- Créer un compte gratuit pour webmaster ;
- Déclarer la propriété du site à la plateforme (par échange de clé ou de fichier que vous devez déposer a la racine de votre site)
- Suivre les activités des robots sur votre site
- Suivre les activités des visiteurs sur votre site.

Pour plus d'information et la manipulation pratique de Google Search Console : <https://facemweb.com/creation-site-internet/google-webmaster-tools-search-console> et/ou <https://www.anthedesign.fr/referencement/google-webmaster-tools-site-web/>

Vous pouvez accéder aussi au **Bing Webmaster Tools** à l'adresse <https://www.bing.com/toolbox/webmaster>

Nous n'allons pas exploiter ces outils actuellement mais ils pourront être très utiles un fois que nous aurons abordé l'aspect pratique de l'hébergement web.



## 5. Exercice : Question pour vous !

*[solution n°3 p.29]*

### Exercice

---

1) Les Webmaster Tools sont :

- ☐ mis à la disposition des webmaster par les hébergeurs.
- ☐ les outils de gestion des sites et plate-formes web proposé par le moteurs de recherche
- ☐ gratuits
- ☐ payants
- ☐ des interfaces web auxquelles on accède par URL
- ☐ des applications qu'il faut installer sur son ordinateur

### Exercice

---

2) Google Webmaster Tools permet de :

- ☐ Créer un compte gratuit pour webmaster
- ☐ Faire des échanges de liens avec d'autres sites référencés avec le même outil
- ☐ Déclarer la propriété du site à la plateforme
- ☐ Programmer des recherche automatiques à faire a l'avenir
- ☐ Suivre les activités des robots sur votre site
- ☐ Personnaliser les résultats des recherches
- ☐ Suivre les activités des visiteurs sur votre site
- ☐ Gérer l'indexation et le référencement du site

### Exercice

---

3) L'enregistrement d'un site web dans la base de données d'un moteur de recherche est appelé :

- ☐ Numérisation
- ☐ Recherche du site
- ☐ Indexation
- ☐ Mot-clé
- ☐ SEA

### Exercice

---

4) Complétez la phrase suivante :

Une liste de mots- permet de définir les contenues dans un document.

### Exercice

---

5) Peut-on supprimer un site ou une même une page de l'index de Google ?

- ☐ Oui
- ☐ Non

# Solutions des exercices



## > Solution n°1

Exercice p. 9

Exercice

1) On ne parle de référencement que dans le cadre des sites web

- ☐ Vrai  
☒ Faux

On peut bien référencer un livre dans une bibliothèque !

Exercice

2) Qu'il s'agisse de SEO ou de SEA, les deux pratiques répondent :

- ☐ aux mêmes règles  
☐ aux mêmes stratégies  
☐ aux mêmes techniques  
☐ aux mêmes besoins

Exercice

3) Le SEO peut aussi être payant

- ☐ Vrai  
☒ Faux

Exercice

4) Quel type de référencement est indiqué pour un site web de e-commerce ?

- ☐ Référencement Naturel  
☒ Référencement payant

Exercice

5) L'ensemble des méthodes et techniques qui visent à positionner les pages web de votre site internet dans les premiers résultats naturels des moteurs de recherche dans le but d'améliorer la visibilité des pages en question s'appelle :



- ☒ le SEO
- ☐ le SEA

### > Solution n°2

Exercice p. 13

## Exercise

### 1) Trouvez l'intrus

- ☐ Bing
- ☒ Gmail
- ☐ Google
- ☐ Quant

Les autres sont des moteurs de recherches alors que Gmail non !

## Exercice

2) Un mot clé, c'est :

- ☐ Un dessin montrant un mot avec une clé
- ☒ Mot qu'on utilise pour faire une recherche
- ☐ Un mot avec une clé dessus

## Exercice

3) Si je connais l'adresse URL d'un site, je ne suis pas obligé d'utiliser Google pour accéder à ce site

- ☒ Vrai
- ☐ Faux

## Exercise

4) Que fais le moteur de recherche Google lorsqu'il ne trouve pas exactement ce qu'il recherche ?

- ☐ Il ferme le navigateur
- ☐ Il se plante
- ☒ Il vous propose des résultats de recherche similaires ou proches de ce que vous cherchez
- ☐ Il vous déconnecte de l'Internet et vous envoie une notification

## Exercise

5) Avec lesquelles des connections Internet on ne peut pas faire de recherches ?

- ☐ 3G+
- ☐ ADSL
- ☐ Wimax
- ☐ 4G
- ☐ 3G
- ☒ Aucune, on peut faire des recherches avec toutes ces connexions

### > **Solution n°3**

Exercice p. 25

#### Exercice

1) Les Webmaster Tools sont :

- ☐ mis à la disposition des webmaster par les hébergeurs.
- ☒ les outils de gestion des sites et plate-formes web proposé par le moteurs de recherche
- ☒ gratuits
- ☐ payants
- ☒ des interfaces web auxquelles on accède par URL
- ☐ des applications qu'il faut installer sur son ordinateur

#### Exercice

2) Google Webmaster Tools permet de :

- ☒ Créer un compte gratuit pour webmaster
- ☐ Faire des échanges de liens avec d'autres sites référencés avec le même outil
- ☒ Déclarer la propriété du site à la plateforme
- ☐ Programmer des recherche automatiques à faire a l'avenir
- ☒ Suivre les activités des robots sur votre site
- ☐ Personnaliser les résultats des recherches
- ☒ Suivre les activités des visiteurs sur votre site
- ☒ Gérer l'indexation et le référencement du site

#### Exercice

3) L'enregistrement d'un site web dans la base de données d'un moteur de recherche est appelé :

- ☐ Numérisation
- ☐ Recherche du site
- ☒ Indexation
- ☐ Mot-clé
- ☐ SEA

## Exercise

4) Complétez la phrase suivante :

Une liste de mots-clés permet de définir les thématiques contenues dans un document.

## Exercise

5) Peut-on supprimer un site ou une même une page de l'index de Google ?

- ☒ Oui
- ☐ Non