

# **Le référencement des sites web**

# Table des matières

<b>I. Contexte</b>	<b>3</b>
<b>II. Principes et objectifs</b>	<b>3</b>
<b>III. Exercice : Appliquez la notion</b>	<b>5</b>
<b>IV. Structurer son site web</b>	<b>5</b>
<b>V. Exercice : Appliquez la notion</b>	<b>10</b>
<b>VI. Baliser ses pages et contenus</b>	<b>10</b>
<b>VII. Exercice : Appliquez la notion</b>	<b>14</b>
<b>VIII. Qualité et pertinence du contenu</b>	<b>14</b>
<b>IX. Exercice : Appliquez la notion</b>	<b>16</b>
<b>X. Stratégie de netlinking</b>	<b>16</b>
<b>XI. Exercice : Appliquez la notion</b>	<b>18</b>
<b>XII. Auto-évaluation</b>	<b>19</b>
A. Exercice final .....	19
B. Exercice : Défi .....	21
<b>Solutions des exercices</b>	<b>22</b>
<b>Contenus annexes</b>	<b>29</b>

## I. Contexte

**Durée :** 1 h

**Environnement de travail :** *Visual Studio Code (cf. p.29)*

**Pré-requis :** Connaître la notion de balise, savoir rédiger une page HTML valide

### Contexte

En règle générale, un site web a pour objectif de proposer des services, des produits ou des contenus à destination d'un public. Sa réussite dépend en bonne partie de la quantité et du ciblage du trafic de visiteurs qu'il parviendra à attirer. Il existe de nombreux moyens de générer du trafic sur son site web. Nous allons ici nous intéresser au plus important d'entre eux : le référencement dans les moteurs de recherche.

## II. Principes et objectifs

### Objectifs

- Identifier les canaux d'acquisition de trafic pour un site web
- Connaître le marché des moteurs de recherche
- Comprendre le fonctionnement d'un moteur de recherche

### Mise en situation

Nous allons, dans cette première partie, découvrir les moyens permettant d'attirer des visiteurs sur un site Internet, et nous nous concentrerons sur les moteurs de recherche, qui constituent la source principale de trafic pour la majorité des sites web.

### Les sources de trafic d'un site Internet

Pour accéder à un site web, on peut tout simplement taper l'URL du site (ex : <https://fr.wikipedia.org>) dans la barre d'adresse du navigateur. Mais dans la majorité des cas, les internautes découvrent un site en suivant un lien trouvé sur un autre site web, dans un e-mail ou un SMS, une publication sur les réseaux sociaux ou encore avec un moteur de recherche. Toutes ces sources de trafic peuvent aider à faire connaître un site web, et il ne faut en négliger aucune, car elles sont souvent complémentaires.

La proportion de visiteurs provenant de chaque source sera différente en fonction du type de site et de sa thématique, ou encore de sa stratégie marketing. Toutefois, pour la plupart des sites, les moteurs de recherche représentent la principale source de trafic.

Pour être bien placé sur un moteur de recherche, il y a deux méthodes :

- La première consiste à payer pour que le site apparaisse dans les liens sponsorisés (SEA),
- Et la seconde est de bien concevoir le site pour lui donner toutes les chances d'apparaître dans les résultats de recherche naturels. Cette activité s'appelle le référencement naturel, ou SEO.

### Définition SEO / SEA

**SEO** est l'acronyme de *Search Engine Optimization*, en français Optimisation pour les Moteurs de Recherche. Cette activité consiste à réaliser toutes les actions possibles sur un site web en vue d'améliorer son positionnement dans les résultats naturels des moteurs de recherche, et par conséquent le trafic qu'il recevra gratuitement de ces derniers.

Le SEO s'oppose au **SEA** (*Search Engine Advertising*), qui consiste à positionner son site web dans les liens sponsorisés d'un moteur de recherche, en payant pour les visiteurs envoyés par le moteur.

L'image ci-dessous illustre l'anatomie d'une page de moteur de recherche et permet d'identifier les différences entre les liens sponsorisés et les résultats naturels.

The image shows a search engine results page for the query "ordinateur portable". The search bar at the top contains the query and a magnifying glass icon. Below the search bar, there are tabs for "Tous", "Shopping", "Images", "Actualités", "Vidéos", "Plus", "Paramètres", and "Outils". The results section shows "Environ 124 000 000 résultats (0,51 secondes)". The first row of results is labeled "Afficher les produits correspondants à ordinateur por..." and "Annonce sponsorisée". It contains five product listings: Surface Book 3, HUAWEI MateBook D 15, MacBook Pro 13.3" Core i7, Surface Laptop 3, and HUAWEI MateBook D 14. The second row is labeled "Annonce" and "www.intel.fr", showing "Nouveaux Ordies Portables 2020 - Choisissez Le Bon Processeur". The third row is labeled "www.fnac.com" and "Informatique", showing "Ordinateur portable - Achat PC Portable | fnac". The fourth row is labeled "www.darty.com" and "Informatique", showing "Ordinateur portable - Livraison Gratuite ou dans 1h en ...". Blue brackets on the right side of the image group the first row as "Annonces et liens sponsorisés" and the subsequent rows as "Résultats Naturels".

## Complément Les parts de marché des moteurs de recherche

Bien qu'il existe une multitude de moteurs de recherche, le plus populaire d'entre eux détient plus de 90 % des parts de marché en France. Il s'agit bien évidemment de Google, qui occupe depuis des années une place hégémonique sur le marché de la recherche dans la plupart des pays du monde. C'est pour cette raison que la plupart des professionnels du référencement focalisent principalement leurs efforts sur le positionnement dans les résultats de Google.

Le tableau suivant détaille les chiffres par type d'appareils pour le marché français en 2020, fournis par StatCounter<sup>1</sup>.

Moteur	Tous appareils	Desktop	Mobile
Google	92,21%	85,33%	97,52%
Bing	3,73%	7,97%	0,45%
Yahoo	1,43%	2,66%	0,44%
Ecosia	1,25%	1,86%	0,82%
Qwant	0,76%	1,45%	0,22%
DuckDuckGo	0,38%	0,56%	0,23%

<sup>1</sup> <https://gs.statcounter.com/>

**Fondamental** Le fonctionnement d'un moteur de recherche

Tous les moteurs de recherche ont un fonctionnement assez similaire. Ils fabriquent un index de pages web à l'aide de robots qui parcourent le Web en suivant tous les liens qu'ils trouvent, et aspirent le contenu de toutes les pages qu'ils découvrent. Ces contenus sont ensuite analysés à l'aide d'algorithmes très complexes pour identifier les thématiques et mots-clés associés.

Ainsi, lorsqu'un internaute saisit une recherche, le moteur trouve dans son index les pages en rapport avec les termes de recherche saisis. Pour classer ces pages et présenter en premier lieu celles qui ont le plus de chances de correspondre à la recherche de l'internaute, le moteur utilise un certain nombre de critères, parmi lesquels :

- La pertinence du contenu de la page par rapport aux termes de recherche,
- Les liens pointant vers la page et leur contexte,
- L'expérience utilisateur offerte par le site.

C'est donc principalement sur ces éléments que va porter le travail de référencement d'un site web : **améliorer les contenus, les liens et l'expérience utilisateur** en fonction des critères d'analyse des moteurs de recherche.

Il n'existe pas de recette magique qui garantisse de figurer en première page sur des recherches très concurrentielles. Il y a un ensemble de techniques qui, si elles sont bien appliquées, permettent incontestablement d'améliorer le classement de son site.

**Syntaxe** À retenir

- Le SEO consiste à réaliser toutes les actions possibles sur un site web en vue d'améliorer son positionnement dans les résultats naturels des moteurs de recherche.
- Il s'oppose au SEA, qui consiste à positionner son site web dans les liens sponsorisés d'un moteur de recherche, en payant pour les visiteurs envoyés par le moteur.
- C'est Google qui détient la majorité des parts de marché des moteurs de recherche, en France et dans le monde.
- Lorsque l'on travaille le référencement d'un site web, il faut se concentrer sur l'amélioration des contenus, les liens et l'expérience utilisateur. C'est sur ces critères que se basera un moteur de recherche pour trier naturellement la pertinence de contenus lorsqu'un internaute effectue une recherche sur un thème donné.

**Exercice : Appliquez la notion**

[solution n°1 p.23]

Parmi les propositions ci-dessus, lesquelles sont justes ?

- ☐ Le référencement naturel consiste à afficher son site dans les liens sponsorisés sur la première page d'un moteur de recherche
- ☐ Google est le moteur de recherche le plus utilisé en France
- ☐ Les moteurs de recherche tiennent compte, entre autres, du contenu d'une page, des liens entrants et de l'expérience utilisateur pour déterminer son classement
- ☐ Il existe une technique infaillible pour positionner son site en tête des résultats
- ☐ SEO signifie *Search Engine Optimization*, appellation anglaise du référencement naturel

**IV. Structurer son site web**

## Objectifs

- Concevoir l'arborescence de son site web avec des rubriques et pages
- Optimiser le maillage des liens entre les pages du site
- Optimiser les URL de ses pages
- Sitemap et robots.txt

## Mise en situation

Le référencement de son site est une question que l'on doit se poser bien avant de commencer à le développer. Une bonne partie de la réussite de son référencement tient à la structure du site lui-même, et il est important d'y réfléchir en amont, pendant la phase de conception.

### Attention

Les moteurs de recherche ne référencent pas à proprement parler des sites web, mais des pages.

### Méthode Structurer ses rubriques et ses pages

Chaque page va donc pouvoir se positionner sur certains mots-clés ou certaines expressions de recherche. Pour optimiser leur référencement, il faut découper son site web de manière à ce que la thématique ou les mots-clés de chaque page soient cohérents. Idéalement, chaque page doit correspondre à un sujet précis et unique.

- Cela permet au visiteur de trouver facilement ce qu'il cherche : il ne faut pas oublier que l'expérience utilisateur est prise en compte par les moteurs de recherche.
- Cela permet également aux moteurs de recherche de bien comprendre les thématiques développées sur le site, et de les mettre en relation pour mieux les référencer.

Par exemple, si notre site web présente des activités, on pourra créer des rubriques "Activités sportives", "Activités culturelles", etc. pour y organiser nos pages. Si une page traite du Tambourin (oui, c'est un sport), et qu'elle a été classée dans "Activités Sportives", cela limitera les risques de confusion avec l'instrument de musique, par exemple !

### Conseil

Bien entendu, dans certains cas, la structure peut être plus complexe, et si l'on a de nombreuses rubriques, il peut être nécessaire de faire des sous-rubriques et de les regrouper dans des catégories correspondant à des thématiques plus générales.

### Attention

Si une page n'a pas de lien entrant, elle ne sera pas découverte par les moteurs de recherche.

### Méthode Maillage des liens entre pages

Les robots des moteurs de recherche explorent le Web en suivant les liens de page en page. Au sein d'un site, il est nécessaire de mettre en place des liens hypertextes entre les différentes pages afin de permettre aux robots de les trouver.

- Il est essentiel de mettre en place des liens depuis une page rubrique vers les pages de contenu de la rubrique. Réciproquement, il faut ajouter un lien retour vers la page rubrique, par exemple en ajoutant un fil d'Ariane en haut de chaque page.
- Les liens entre les pages d'un même niveau ont aussi leur importance, et même entre deux pages qui ne font pas nécessairement partie de la même rubrique. Une bonne pratique est de mettre un lien dans le texte d'un article, sur un terme développé dans une autre page de votre site.

**Exemple**

Pour illustrer cela, sur la première ligne de ce paragraphe, nous aurions pu mettre un lien sur les mots "liens hypertextes" vers le cours sur les liens hypertextes.

**Conseil**

Le texte des liens doit toujours correspondre à la thématique de la page de destination : en règle générale, on reprend le titre de la page, ou une expression qui correspond à son contenu.

**Attention**

Les moteurs de recherche s'appuient également sur les URL pour identifier la thématique des pages.

**Méthode Optimiser et réécrire ses URL**

Une bonne pratique consiste à se calquer sur l'arborescence du site en ajoutant dans l'URL le nom de la rubrique et celui de la page, ou bien un terme essentiel.

Comme les espaces ne sont pas des caractères autorisés dans les URL, une bonne pratique consiste à séparer les mots par un caractère tel que - ou \_ et les éléments d'arborescence par des slash /.

**Exemple**

Par exemple : `www.monsite.com/activites_sportives/decouvrir_le_tambourin.html` plutôt que `www.monsite.com/article.php?id=1`.

**Complément**

Pour mettre ceci en pratique sur un site dynamique, il faut avoir recours à la réécriture d'URL.

**Définition Sitemap**

Le `sitemap` est une cartographie du contenu d'un site web, généralement placée à la racine du site afin d'aider les moteurs de recherche à bien explorer l'intégralité de son contenu. Il s'agit généralement d'un fichier au format XML, mais il peut aussi être en texte simple ou au format RSS.

## Syntaxe et formatage d'un Sitemap au format XML

Le formatage d'un sitemap XML fonctionne avec les balises XML. Selon le standard actuel « Sitemaps<sup>1</sup> », trois balises sont obligatoires : `<urlset>`, `<url>` et `<loc>`.

Le plan Sitemap doit :

- Commencer par une balise d'ouverture `<urlset2>` et terminer par une balise de fermeture `</urlset>`.
- Spécifier l'espace de nom (standard de protocole) dans la balise `<urlset>`
- Inclure pour chaque URL une entrée `<url3>` en tant que balise XML parent.
- Inclure une entrée enfant `<loc4>` pour chaque balise parent `<url>`

Toutes les autres balises sont facultatives. La prise en charge de ces balises facultatives peut varier d'un moteur de recherche à un autre. Pour plus d'informations, reportez-vous à la documentation des différents moteurs de recherche.

Exemple	Exemple Sitemap
	<pre> 1 &lt;code&gt; 2 &lt;?xml version="1.0" encoding="UTF-8"?&gt; 3 &lt;urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"&gt; 4 5   &lt;url&gt; 6     &lt;loc&gt;http://www.example.com/&lt;/loc&gt; 7     &lt;lastmod&gt;2005-01-01&lt;/lastmod&gt; 8   &lt;/url&gt; 9 10  &lt;url&gt; 11    &lt;loc&gt;http://www.example.com/page1&lt;/loc&gt; 12    &lt;lastmod&gt;2005-01-04&lt;/lastmod&gt; 13  &lt;/url&gt; 14 15  &lt;url&gt; 16    &lt;loc&gt;http://www.example.com/page2&lt;/loc&gt; 17    &lt;lastmod&gt;2005-01-05&lt;/lastmod&gt; 18  &lt;/url&gt; 19 &lt;/urlset&gt; 20 &lt;/code&gt; </pre>

Il existe bien évidemment énormément d'outils qui permette de générer un sitemap. Le site XML Sitemaps<sup>5</sup> est une bonne référence.

La solution idéale étant d'avoir un sitemap qui se met à jour automatiquement à chaque fois que vous publiez ou que vous mettez à jour un contenu. Par exemple sur WordPress, l'extension Yoast SEO intègre un sitemap par exemple ou encore en utilisant des extensions spécialisées, comme Google XML Sitemaps<sup>6</sup>.

Il est également conseillé d'indiquer le lien de votre sitemap dans le fichier robots.txt situé à la racine du site, en ajoutant ce type de ligne de code à la fin du fichier robots.txt par exemple :Sitemap: https://www.votresite.com/sitemap.xml

1 <https://www.sitemaps.org/fr/protocol.html>

2 <https://www.sitemaps.org/fr/protocol.html#urlsetdef>

3 <https://www.sitemaps.org/fr/protocol.html#urldef>

4 <https://www.sitemaps.org/fr/protocol.html#locdef>

5 <https://www.xml-sitemaps.com/>

6 <https://fr.wordpress.org/plugins/google-sitemap-generator/>



**Définition Robots.txt**

Le fichier `robots.txt` permet de signaler aux moteurs de recherche les pages ou parties de notre site que l'on ne souhaite pas qu'ils explorent ni qu'ils référencent. C'est également un fichier qui doit être posé à la racine du site web. Il permet notamment d'éviter que les moteurs ne référencent des pages contenant du contenu dupliqué ou des pages non pertinentes de votre site.

Le fichier `robots.txt` se trouve au niveau de la racine de votre site internet. Pour vérifier sa présence sur votre site, vous tapez dans la barre d'adresse de votre navigateur : <http://www.adressedevotresite.com/robots.txt>.

Si le fichier est :

- présent, il s'affichera et les robots suivront les instructions présentes dans le fichier.
- absent, une erreur 404<sup>1</sup> s'affiche et les robots considèrent qu'aucun contenu n'est interdit.

Le fichier `robots.txt` est créé manuellement ou généré par défaut par la majorité des CMS comme WordPress au moment de leur installation. Mais il est également possible de créer votre fichier pour les robots avec des outils en ligne<sup>2</sup>.

Pour la création manuelle, vous utilisez un simple éditeur de texte tel que Notepad tout en respectant à chaque fois :

- une syntaxe et des instructions,
- un nom de fichier : `robots.txt`,
- une structure : une instruction par ligne et aucune ligne vide.

**La syntaxe et les instructions du fichier robots.txt**

Les fichiers `robots.txt` utilisent les instructions ou commandes suivantes :

- **User-agent** : les user-agents sont les robots des moteurs de recherche, par exemple Googlebot pour Google ou Bingbot<sup>3</sup> pour Bing.
- **Disallow** : `disallow` est l'instruction qui interdit aux user-agents l'accès à une url ou à un dossier.
- **Allow** : `allow` est une instruction autorisant l'accès à une url placée dans un dossier interdit.

**Exemple**

```
1 # Groupe 1
2 User-agent: Googlebot
3 Disallow: /nogooglebot/
4
5 # Groupe 2
6 User-agent: *
7 Allow: /
8
9 Sitemap: http://www.example.com/sitemap.xml
```

Ici, ce fichier indique que :

- Le user-agent nommé "Googlebot" n'est pas autorisé à explorer le dossier <http://example.com/nogooglebot/>, ni ses sous-répertoires.
- Tous les autres user-agents peuvent accéder à l'ensemble du site. Cette partie peut être omise, car l'accès est supposé complet par principe, et cette précision ne change rien au résultat.
- Le fichier `sitemap` du site se trouve à l'adresse <http://www.example.com/sitemap.xml>.

1 <https://www.anthedesign.fr/creation-de-sites-internet/erreur-404-not-found/>

2 <http://www.yellowpipe.com/yis/tools/robots.txt/index.php>

3 <http://robots-txt.com/ressources/robots-txt-bing/>

### Méthode

Le site de support de Google<sup>1</sup> propose une documentation complète sur le format et le contenu attendus pour le `sitemap` et le fichier `robot.txt`.

### Syntaxe À retenir

- Une bonne partie de la réussite de son référencement tient à la structure du site. Ainsi, il est important de structurer correctement ses rubriques et les pages qu'elles contiennent, tout comme il est important d'effectuer des liens entre elles.
- La réécriture des URLs est également importante et permettent d'identifier la thématique d'une page, pour un robot comme pour un humain.
- Le fichier `sitemap` permet de cartographier le contenu d'un site web.
- Le fichier `robots.txt` permet quant à lui de signaler aux robots quelles parties on souhaite qu'ils indexent ou non.

## V. Exercice : Appliquez la notion

### Question

[solution n°2 p.23]

Dans le but d'améliorer notre référencement et de protéger l'accès à certains répertoires de notre site web, nous allons écrire un fichier `robots.txt` qu'on placera à la racine du site. Ce fichier :

- Interdira au robot d'indexation Google (Googlebot) l'accès aux répertoires :  
**admin**  
**documentation**  
**deleted**
- Interdira totalement l'accès du site au robot d'indexation Yahoo (Yahoo! Slurp)
- Autorisera l'accès aux autres `user-agent` à tout le contenu

## VI. Baliser ses pages et contenus

### Objectifs

- Les balises `meta`
- Balisage hiérarchique du contenu
- Les balises `microdata`

### Mise en situation

Une fois que la structure de votre site est en place et permet une bonne compréhension de la structure de votre contenu par les moteurs de recherche, nous entrons dans le vif du sujet : les pages de contenu, qui doivent, elles aussi, être structurées afin d'aider les moteurs à identifier les sujets développés dans vos contenus.

<sup>1</sup> <https://support.google.com/webmasters/answer/183668?hl=fr>

**Définition** Les balises <meta>

Les balises <meta> sont présentes dans la balise <head> des pages HTML. Elles ne s'affichent pas dans le navigateur des visiteurs, mais ont une grande importance pour le référencement et la présentation des pages sur les moteurs de recherche.

Les plus importantes sont les balises `title` et `description`. Elles correspondent :

- au titre de votre page,
- et à sa description.

**Fondamental**

Les balises <meta> sont reprises par les moteurs de recherche dans l'affichage des résultats :

- La balise `title` est reprise systématiquement,
- Et, si la balise `meta description` contient les mots-clés tapés par l'internaute, elle sera affichée en dessous.

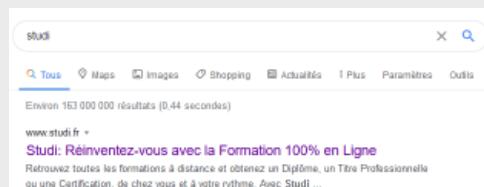
**Exemple**

Voici une illustration avec les balises présentes sur la page d'accueil de STUDI et le résultat sur Google :

```

1 <html>
2   <head>
3     <title>Studi: Réinventez-vous avec la Formation 100% en Ligne" </title>
4     <meta name="description" content="Retrouvez toutes les formations à distance et obtenez un
    Diplôme, un Titre Professionnel ou une Certification, de chez vous et à votre rythme. Avec
    Studi, Progressez, Changez, Evolvez. Nous sommes là." />
5     ...
6   </head>
7   ...

```

**Complément**

Il existe de nombreuses autres balises `meta`, et notamment la balise `robots`, qui contient des directives pour les robots des moteurs de recherche (alternative au fichier `robots.txt`), ou encore la balise `canonical` qui permet de déclarer le nom de domaine principal de votre site pour éviter le *duplicate content*, si les mêmes contenus sont disponibles sur plusieurs domaines ou sous-domaines (ex : `monsite.com` et `www.monsite.com`).

**Remarque**

On rencontre parfois une balise `keywords`, qui autrefois était utilisée pour indiquer aux moteurs de recherche les mots-clés développés dans la page, mais elle est aujourd'hui ignorée, et il est inutile de l'intégrer sur votre site web.

### Définition Balisage hiérarchique du contenu

Les balises hiérarchiques de contenu sont probablement les plus importantes pour structurer le contenu d'une page. Elles permettent d'identifier les titres des différentes parties des contenus en sections et sous-sections.

Ces balises sont de la forme `<hX>`, où `X` est le niveau de profondeur de la section dans l'arborescence de la page.

- Le titre d'une page étant l'élément racine de cette hiérarchie, il doit être placé dans une balise `<h1>`.
- Les titres des principales sections du contenu de la page seront placés dans des balises `<h2>`, les sous-titres de sections dans `<h3>`, et ainsi de suite.

### Exemple

```
1 <html>
2   <body>
3     <h1>Le référencement naturel</h1>
4     <h2>Bien structurer son site web</h2>
5     <h3>Arborescence des rubriques et des pages</h3>
6     <h3>Maillage des liens entre pages</h3>
7     <h3>Optimisation et réécriture d'URL</h3>
8     <h2>Baliser ses pages et contenus</h2>
9     <h3>Les balises meta</h3>
10    <h3>Balisage hiérarchique du contenu</h3>
11    <h3>Les balises Microdata</h3>
12  </body>
13 </html>
```

### Conseil

Il est admis que le contenu des balises `<h1>`, `<h2>` et suivantes a une importance pour les moteurs de recherche.

Le texte de ces balises doit être soigneusement choisi pour décrire en quelques mots le contenu qu'elles précèdent, en reprenant les mots-clés principaux que l'on souhaite mettre en avant.

### Définition Microdata

Les **microdata**, ou micro-données, sont des attributs sémantiques qui peuvent être ajoutées aux balises d'une page web pour permettre aux moteurs de recherche de mieux interpréter son contenu, et de le mettre en relation avec d'autres contenus.

Ces attributs suivent une nomenclature universelle que l'on peut retrouver sur le site <https://schema.org>.

### Microdata avec HTML5

Avec l'arrivée de nouveaux éléments comme `<time>` et `<article>`, la sémantique est apportée par les éléments HTML, mais pas seulement : trois attributs principaux sont utilisés pour intégrer des Microdata dans HTML5 :

- `itemscope` : Cet attribut définit la portée du vocabulaire déclaré par `itemtype` et englobe généralement un objet.
- `itemtype` : Cet attribut définit l'URL du vocabulaire qui spécifie le modèle de données utilisé pour décrire les objets.
- `itemprop` : Cet attribut est utilisé afin d'ajouter des propriétés à un élément.

Ils sont complétés par `itemref` qui permet de faire référence à un type déjà défini dans la page et `itemid` qui fait référence à l'identifiant unique d'un objet.

**Exemple** Un exemple d'utilisation pour la présentation d'un livre

```

1 <div itemscope itemtype="http://schema.org/Book">
2   <h3 itemprop="name">Le Seigneur des Anneaux</h3>
3   <p itemprop="author">J.R.R. Tolkien</p>
4   <p itemprop="genre">Genre : Heroic Fantasy</p>
5   <p>Prix: <span itemprop="price" content="9.50">9.50 €</span></p>
6   <meta itemprop="priceCurrency" content="EUR" />
7 </div>

```

**Conseil**

L'implémentation des microdata n'est pas indispensable, mais elle constitue un plus pour garantir la bonne compréhension des contenus et du type de chaque élément.

Ces données pourront être mises en avant dans les résultats de recherche, notamment pour l'affichage des prix ou encore des notations dans les *snippets* de résultats.

**Définition** Snippet

Ce que les professionnels du référencement appellent **snippet** correspond à la présentation d'une page dans les résultats. Il est au minimum constitué d'un titre, d'une description et d'un lien, mais peut être enrichi avec de nombreux éléments : image, prix, notations (les petites étoiles jaunes), fil d'Ariane... On parle alors de snippets enrichis, comme l'illustre l'exemple suivant :

**Conseil**

Les snippets enrichis offrent plus de visibilité aux contenus sur la page de résultats et permettent de se démarquer dans la liste de résultats des moteurs, et d'attirer ainsi plus de clics.

**Complément**

Pour simplifier l'intégration des microdatas, Google fournit un outil d'aide au balisage des pages<sup>1</sup> et un outil de test d'insertion des microdata afin de vérifier le respect de l'implémentation des valeurs de schema.org : <https://developers.google.com/structured-data/testing-tool/>. Vous pourrez vérifier soit directement une URL soit directement un bout de code.

**Syntaxe** À retenir

- Les balises `meta` les plus importantes à ajouter son `title` et `description`.
- L'ordre de présentation des titres avec les balises `hx` est important pour structurer le contenu.
- Il est possible de rajouter des attributs spécifiques sur les éléments HTML pour aiguiller le moteur de recherche sur le type de données dont il s'agit.

<sup>1</sup> <https://www.google.com/webmasters/markup-helper/u/0/?hl=fr>

## VII. Exercice : Appliquez la notion

### Question

[solution n°3 p.23]

L'objectif est de référencer le mieux possible un site qui fait une liste des meilleurs films avec, pour chaque film, son titre, sa note sur 10 et son réalisateur.

À l'aide de l'outil Google de balisage<sup>1</sup>, générez le code au format **microdonnées**, qui donnera plus de contexte à notre liste de films.

Le type de données qu'on souhaite représenter sur la page est **Films**. Pour commencer, collez le code HTML en sélectionnant l'onglet "HTML" dans le bas de la page.

```

1  <ul>
2    <li>
3      <h4>Citizen Kane</h4>
4      <p>8.3</p>
5      <p>Orson Welles</p>
6    </li>
7    <li>
8      <h4>Le Parrain</h4>
9      <p>8.5</p>
10     <p>Francis Ford Coppola</p>
11   </li>
12   <li>
13     <h4>Pulp Fiction</h4>
14     <p>8.9</p>
15     <p>Quentin Tarantino</p>
16   </li>
17 </ul>

```

### Indice :

Après avoir indexé toutes les données de la liste, il est nécessaire de changer le format d'export en **microdonnées**.



## VIII. Qualité et pertinence du contenu

### Objectifs

- Rédaction de contenus optimisés
- Travail des champs lexicaux
- Proposer un contenu unique et original

<sup>1</sup> <https://www.google.com/webmasters/markup-helper/u/0/>

## Mise en situation

Bien que toute la structuration de votre site et de vos pages soit importante et nécessaire, la réussite de votre référencement passera inévitablement par un contenu de qualité, unique et actualisé.

Il y a quelques années, on trouvait encore en tête des résultats des moteurs de recherche des pages contenant des textes manifestement écrits avant tout pour les moteurs de recherche : phrases n'ayant aucun sens truffées de mots-clés et de synonymes, répétition des mêmes mots-clés tout au long de la page... Ces techniques ont fait leur temps, et si nous les appliquons aujourd'hui, notre site serait tout bonnement mal référencé.

### Attention Rédaction de contenus pertinents

On n'écrit pas pour les moteurs de recherche, mais pour l'internaute, avec pour objectif premier de lui apporter l'information qu'il recherche, et la plus pertinente possible.

Il y a au moins deux raisons à cela :

- Les moteurs mesurent le taux de rebond de nos pages, c'est-à-dire le nombre de personnes qui quittent notre page immédiatement après y être entrées. Si notre contenu n'intéresse pas le visiteur, notre taux de rebond va grimper en flèche et le référencement de nos pages en sera pénalisé.

Si, au contraire, notre contenu est captivant, l'internaute restera sur notre site, notre taux de rebond sera faible et nos pages seront mieux référencées sur les moteurs.

- Avec l'intégration de l'intelligence artificielle, les moteurs tendent de plus en plus à apprécier le contenu comme un utilisateur humain, et si l'on fait en sorte d'intéresser l'internaute, il y a toutes les chances que l'IA du moteur de recherche le perçoive et ait une bonne appréciation de nos contenus.

### Méthode Travail des champs lexicaux

Pour référencer une page sur un mot-clé ou une expression donnée, il peut être tentant de les répéter à plusieurs reprises dans les titres et contenus de la page, au risque d'être redondant et d'apporter de la lourdeur au contenu. Comme nous l'avons vu précédemment, cette méthode est à **proscrire**.

Il existe une alternative bien plus efficace : explorer le champ lexical pour intégrer dans les différents contenus de la page des mots-clés ou expressions en rapport avec les termes de recherche visés.

Cette pratique présente un triple intérêt :

- L'analyse des moteurs de recherche tient compte de la présence de mots-clés et expressions faisant partie du champ lexical d'un terme pour le positionnement de la page sur celui-ci,
- La page va pouvoir se positionner sur des mots-clés connexes, sur lesquels il y a peut-être une concurrence moindre, et de multiplier les requêtes sur lesquelles le site apparaîtra,
- L'utilisateur aura l'impression que le contenu couvre le sujet plus largement, et restera plus volontiers sur le site, ce qui peut améliorer le taux de rebond. Les poètes et auteurs de littérature ont souvent recours aux champs lexicaux pour immerger le lecteur dans leur univers, pourquoi ne pas appliquer cette technique pour nos contenus ?

### Définition Duplicate content (contenu dupliqué)

Le **duplicate content**, que l'on traduit par « contenu dupliqué », correspond à un contenu qui est présenté sur plusieurs pages d'un même site, ou repris d'un site sur un autre.

**Attention** Proposer un contenu unique et original

Les moteurs de recherche mènent une guerre au duplicate content, pour dissuader du plagiat et encourager les contenus uniques et originaux.

**Rappel**

Inutile, donc, d'espérer nourrir nos pages de textes copiés/collés sur Wikipédia ou autres sites. D'une part, cela peut nous causer des problèmes (plaintes pour atteinte au droit d'auteur), mais en plus les moteurs de recherche l'identifieront et ne référenceront pas ces pages. Il ne faut pas non plus espérer tromper leur vigilance en changeant certains mots : au-dessus d'un certain niveau de similitude, le contenu sera catalogué comme duplicate content.

Il arrive parfois qu'il soit nécessaire et pertinent de présenter le même contenu sur plusieurs pages d'un même site, ou sur plusieurs sites différents, ou encore que ceci soit le fait d'une contrainte technique (le site est accessible sur `monsite.com` et sur `www.monsite.com`).

Dans ce cas, il existe plusieurs solutions : indiquer aux robots de ne pas indexer certaines pages (`robots.txt` ou balise `meta noindex`), déclarer la page originale sur les copies (balise `canonical`).

Mais en dehors de ces cas de figure, si l'on veut assurer un bon référencement de nos contenus, il est essentiel de présenter des textes uniques et originaux sur nos pages.

**Syntaxe** À retenir

- Le contenu reste l'élément le plus important pour réussir le référencement de son site : un contenu bien écrit et original sera apprécié par les internautes, et donc par les moteurs de recherche.
- Plutôt que de répéter certains mots, il est préférable de mettre à profit le champ lexical du thème sur lequel porte notre contenu pour diversifier le texte et améliorer la qualité générale du site.
- Le contenu dupliqué est généralement pénalisé par les moteurs de recherche. Le robot peut détecter le copié-collé provenant d'un autre site et déterminer que vous n'êtes pas à l'origine du contenu, et que par conséquent votre site a moins d'intérêt à être visité.

## Exercice : Appliquez la notion

[solution n°4 p.24]

Exercice

Le succès du référencement passe par...

- ☐ Le fait de répéter des mots-clés liés au thème du site
- ☐ D'écrire du contenu de qualité

Exercice

Il est préférable...

- ☐ D'avoir du contenu original et de créer des liens vers les ressources externes au besoin
- ☐ De copier les paragraphes de texte sur son site pour garder les utilisateurs plus longtemps sur son contenu

## X. Stratégie de netlinking



## Objectifs

- Obtenir de bons liens vers votre site
- Partenariats et échanges de liens
- Stratégies abusives et sanctions

## Mise en situation

Dernier pilier de notre stratégie de référencement, et probablement le plus complexe : il s'agit maintenant d'obtenir des liens venant d'autres sites web vers nos pages afin de générer du trafic entrant, mais surtout d'améliorer la popularité de notre site aux yeux des moteurs.

## Les liens entrants

Les moteurs utilisent les liens pointant vers une page pour évaluer sa popularité. Pour schématiser, on pourrait dire qu'un lien entrant représente un vote, et plus la page sur laquelle le lien est présent est efficace, plus le vote a de poids pour mesurer la popularité de la page de destination du lien. Et les moteurs de recherche utilisent ces votes pour déterminer les pages les plus populaires sur un sujet, et trier ainsi les résultats.

Ce schéma simpliste était valable dans les années 2000, mais aujourd'hui l'analyse des liens entrants par les moteurs de recherche est beaucoup plus fine. Le texte du lien, la thématique de la page sur laquelle il est intégré, et le contexte dans lequel il est placé, et d'autres facteurs sont pris en compte dans la valeur du lien, et donc l'impact qu'il pourra avoir sur votre référencement.

### Fondamental

Le lien idéal serait le suivant :

- sur une page populaire traitant d'une thématique connexe ou complémentaire,
- dans un paragraphe traitant d'un sujet proche de celui de la page cible,
- pointant vers une page traitant spécifiquement d'un sujet plutôt que vers votre page d'accueil (sauf si tout votre site traite du sujet en question),
- avec un texte reprenant une expression ou un terme pertinent par rapport à votre contenu, mais pas nécessairement le mot-clé précis sur lequel vous souhaitez vous référencer (nous verrons plus tard pourquoi).

### Méthode Comment obtenir de tels liens, en pratique ?

Si l'on a déjà des sites sur des thématiques connexes, on peut utiliser ses propres sites pour commencer.

Vous pouvez également proposer un contenu dont l'intérêt est tel que, naturellement, vos visiteurs vont en parler sur les forums, blogs et réseaux sociaux. Ces liens-là valent de l'or, car ce sont ceux que les moteurs de recherche valorisent le plus.

Enfin, vous n'êtes pas le seul webmaster ou éditeur de site à vouloir faire connaître son site, et vous allez pouvoir mettre en place des partenariats avec d'autres sites web.

### Méthode Partenariats et échanges de liens

Bien que les moteurs affectionnent particulièrement les liens "naturels", qui font suite à l'intérêt particulier d'un visiteur pour une page web, ils ne peuvent proscrire les échanges de liens, et une bonne partie des liens pris en compte pour le référencement des sites sont issus de tels partenariats.

L'échange de liens doit toutefois être bien pensé pour optimiser son efficacité pour chacun des sites. Ils doivent pour cela répondre aux critères vus dans le paragraphe précédent.

Pour obtenir de tels partenariats, on peut faire appel à notre réseau de connaissances, ou prospecter dans les forums de webmasters, notamment ceux dédiés au référencement (ceux-ci proposeront généralement de nombreuses ressources pour approfondir les concepts présentés dans ce cours).

Une méthode très en vogue actuellement est le *guest blogging*, qui consiste à s'entendre avec un influenceur pour publier sur son blog un article pertinent sur un sujet en lien avec votre site, en prenant soin d'y insérer avec parcimonie quelques liens pertinents vers vos pages.

#### Attention Stratégies abusives et sanctions

Le référencement n'est pas une science exacte, et ce qui était vrai hier ne l'est plus aujourd'hui. Il y a 20 ans, il était assez facile de référencer un site web en ayant recours à des pratiques qui, aujourd'hui, sont totalement proscrites. Comme, par exemple, la création de réseaux de sites ou de pages satellites sans contenu, n'ayant pour but que de générer un maximum de liens vers les pages à référencer, ou abuser de la répétition de mots-clés.

Ces pratiques, que l'on appelle **spamdexing**, pourraient marcher quelque temps, mais conduiront assurément à une pénalisation et un déréférencement de vos pages.

#### Complément

D'autres pratiques sont interdites, notamment l'achat de liens, qui consiste à payer l'éditeur d'un site pour qu'il ajoute un lien vers notre site afin d'influencer le référencement. Il n'est pas interdit d'acheter des liens pour apporter du trafic, c'est d'ailleurs le principe des liens sponsorisés que pratiquent tous les moteurs de recherche. Mais dans ce cas, il est indispensable d'ajouter l'attribut `nofollow` sur les liens, pour qu'ils ne soient pas pris en compte par les moteurs de recherche, comme ci-dessous :

```
1 <a href="https://www.monsite.com" rel="nofollow">Mon site</a>
```

Les moteurs de recherche ont développé des algorithmes très performants pour identifier toutes ces fraudes, et pénaliser les sites qui y ont recours.

#### Syntaxe À retenir

- La popularité d'un site web est également estimé par le robot aux nombre de liens de sites externes qui pointent vers le site. C'est ce qu'on appelle un lien "entrant". Pour générer des liens entrants, il est possible de se baser sur ses autres sites web pour faire de la publicité au nouveau site.
- Il est possible de procéder à un "échange de liens" avec le gestionnaire d'un autre site, mais il est préférable que ce lien ait un sens, par exemple si les contenus des deux sites sont complémentaires.

## Exercice : Appliquez la notion

[solution n°5 p.24]

### Exercice

Pour s'assurer du bon référencement de son site, il est important...

- ☐ D'ajouter un maximum de liens vers des sites externes
- ☐ D'avoir un maximum de sites externes qui proposent un lien vers son site

### Exercice

L'échange de liens est une pratique consistant à...

- ☐ Faire un partenariat avec un gestionnaire de site pour référencer mutuellement vos sites web
- ☐ Acheter de l'espace publicitaire sur Google pour être dans les résultats les plus hauts

## XII. Auto-évaluation

### A. Exercice final

#### Exercice 1

[solution n°6 p.25]

##### Exercice

La meilleure stratégie pour structurer son site web consiste à...

- ☐ Mettre toutes les pages à la racine du site, avec un lien vers chaque page sur la page d'accueil du site
- ☐ Créer une page par thématique, et inclure tous les sujets sous la forme de sous-parties de la page
- ☐ Créer une rubrique par thématique, chaque rubrique étant divisée en plusieurs pages traitant chacune un sujet précis
- ☐ Créer une page spécifique pour chaque mot-clé

##### Exercice

Par quel(s) moyens les robots des moteurs de recherche peuvent-ils découvrir les pages de votre site ?

- ☐ En suivant des liens depuis d'autres sites
- ☐ En suivant des liens depuis d'autres pages de votre site
- ☐ Grâce au fichier `sitemap.xml` présent à la racine de votre site
- ☐ Grâce au fichier `robots.xml` présent à la racine de votre site

##### Exercice

Sur mon site, la page `https://www.monsite.com/produit.php?id=3` présente le modèle de clavier de marque Techlogic et de référence KB123. Parmi les propositions suivantes, laquelle constitue le meilleur exemple de réécriture d'URL ?

- ☐ `https://www.monsite.com/claviers_sans_fil/techlogic/kb123.html`
- ☐ `https://www.monsite.com/produits/kb123.html`
- ☐ `https://www.monsite.com/KB123/clavier.html`
- ☐ `https://www.monsite.com/claviers_sans_fil_techlogickb123.html`

##### Exercice

Une balise `h1` est utilisée pour encadrer...

- ☐ Le titre d'un paragraphe
- ☐ Le titre d'une section de la page
- ☐ Le titre principal de la page
- ☐ Le titre du site

##### Exercice

Pour que les moteurs de recherche considèrent que ma page est pertinente pour un mot-clé, où ai-je intérêt à intégrer ce mot-clé ?

- ☐ Dans le titre de la page
- ☐ Dans le contenu de la page, au sein d'un contenu en lien avec ce mot-clé
- ☐ Dans la balise meta `keywords`, avec tous les mots-clés sur lesquels je veux positionner ma page
- ☐ Dans le fichier `robots.txt`

#### Exercice

Quel élément de ma page est généralement repris comme titre dans les résultats de recherche ?

- ☐ La balise `title`
- ☐ La balise meta `description`
- ☐ La balise `h1`
- ☐ Les premiers mots affichés sur la page

#### Exercice

Comment se définit le taux de rebond sur une page web ?

- ☐ Le nombre de clics par visiteur sur les liens d'une page web
- ☐ La proportion de visiteurs qui a visité cette page web, puis a visité d'autres pages sur le même site
- ☐ La proportion de visiteurs qui est resté sur la page pour consulter son contenu
- ☐ La proportion de visiteurs qui a quitté rapidement la page sans consulter son contenu

#### Exercice

Parmi les pratiques suivantes, lesquelles représentent un danger pour mon référencement sur un moteur de recherche ?

- ☐ Acheter des liens sponsorisés sur un moteur concurrent
- ☐ Proposer des contenus dupliqués sur mon site
- ☐ Obtenir des liens en faisant du guest blogging
- ☐ Développer un réseau de pages satellites sans contenu pertinent
- ☐ Pratiquer la répétition excessive de mots-clés dans les contenus de mes sites

#### Exercice

Pour maximiser les chances de référencer mon site sur la requête "Planter des tomates", comment dois-je rédiger mes pages de contenu ?

- ☐ Je dois inclure cette requête dans tous les titres et sous-titres de ma page, et répéter de nombreuses fois les mots "planter" et "tomates" dans mes textes, en les mettant en gras
- ☐ Je dois essayer de me mettre à la place du robot du moteur de recherche et faire des phrases courtes et simples pour qu'il les comprenne bien
- ☐ Je dois essayer de me mettre à la place du visiteur qui a tapé cette requête dans Google et essayer de lui proposer un contenu qui explique de manière pertinente et bien illustrée comment planter des tomates

- ☐ Je dois copier/coller le texte de la page qui est en première position sur la requête "Planter des tomates" et changer quelques mots pour éviter les problèmes de droits d'auteur

#### Exercice

Une page web ne peut être référencée que sur un seul mot-clé ou une seule expression.

- ☐ Vrai
- ☐ Faux

#### Exercice

Les liens qui apportent une valeur ajoutée à mes pages en termes de référencement naturel sont...

- ☐ Les liens naturels, que les visiteurs de mon site ajoutent spontanément dans des posts de forum, sur des blogs ou réseaux sociaux
- ☐ Les liens sponsorisés
- ☐ Les liens "Nofollow"
- ☐ Les liens intégrés dans des contenus (blogs ou sites) partenaires, dont la thématique est proche de celle de mon site

### B. Exercice : Défi

Vous allez maintenant mettre en œuvre certaines des bonnes pratiques acquises durant ce cours.

Vous devez optimiser un site web très simple constitué de 3 pages et présentant les activités sportives de la ville (fictive) de Saint-Martin.

L'objectif est de mettre en application les bonnes pratiques vues dans ce cours, et notamment :

- La structuration du contenu (balises `meta` et `h1`, `h2`...)
- Le bon maillage des liens entre les pages.

#### Question 1

[solution n°7 p.27]

Voici le code source de la page d'accueil (**index.html**) : modifiez-en la structure et le maillage entre les pages.

```

1 <!doctype html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5   <title>Activités de la ville de Saint Martin</title>
6   <meta name="description" content="Découvrir, apprendre, échanger : la ville de Saint Martin
  propose un large choix de clubs sportifs et associations culturelles." />
7 </head>
8 <body>
9   <div class="titre">Activités de la ville de Saint Martin</div>
10  <p>Bienvenue sur le site de la ville de Saint Martin. Découvrir, apprendre, échanger : la
  ville de Saint Martin propose un large choix de clubs sportifs et associations culturelles.
  </p>
11  <div class="soustitre">Types d'activités</div>
12  <a href="activites_sportives.html">Activités sportives</a>
13  <a href="activites_culturelles.html">Activités Culturelles</a>
14 </body>
15 </html>

```

## Question 2

[solution n°8 p.27]

Voici le code d'une page rubrique d'activités, par exemple les activités sportives (**activites\_sportives.html**) : modifiez-en la structure et le maillage entre les pages.

```
1 <!doctype html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5   <title>Activités sportives de la ville de Saint Martin</title>
6 </head>
7 <body>
8   <div class="navigation">Activités de la ville de Saint Martin > Activités sportives</div>
9   <div class="titre">Activités sportives de la ville de Saint Martin</div>
10  <p>Bienvenue sur le site de la ville de Saint Martin. Découvrir, apprendre, échanger : la
ville de Saint Martin propose un large choix de clubs sportifs.</p>
11  <div class="soustitre">Sports individuels</div>
12  <a href="tennis_club_saint_martin.html">Tennis club de Saint Martin</a>
13  <a href="golf_de_saint_martin">Golf de Saint Martin</a>
14  <div class="soustitre">Sports collectifs</div>
15  <a href="football_club_saint_martin.html">Football club de Saint Martin</a>
16  <a href="saint_martin_basket.html">Saint Martin Basket</a>
17 </body>
18 </html>
```

## Question 3

[solution n°9 p.28]

Et, enfin, voici le code d'une page présentant une activité, par exemple le tennis club (**tennis\_club\_saint\_martin.html**) : modifiez-en la structure et le maillage entre les pages.

```
1 <!doctype html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5   <title>Tennis Club de Saint Martin</title>
6 </head>
7 <body>
8   <div class="navigation">Activités de la ville de Saint Martin > Activités sportives > Tennis
Club</div>
9   <div class="titre">Tennis Club de Saint Martin</div>
10  <p>Bienvenue sur la page du Tennis Club de la ville de Saint Martin. Découvrir, apprendre,
échanger : la ville de Saint Martin propose de pratiquer le Tennis grâce à son club municipal.
</p>
11  <div class="soustitre">Présentation</div>
12  <p>Le club de Tennis municipal de Saint Martin vous offre quatre courts en plein-air et un
court couvert ...</p>
13  <div class="soustitre">Coordonnées</div>
14  <div>Président du Club</div>
15  <p>Nicolas Durand<br/>
16  10 rue des fleurs
17  Saint Martin</p>
18  <p>Téléphone : 01 23 45 67 89</p>
19  <div>Entraîneur</div>
20  <p>Gérard Lebon<br/>
21  1 avenue des platanes
22  Saint Martin</p>
23  <p>Téléphone : 01 23 98 76</p>
24 </body>
25 </html>
```

## Solutions des exercices

**Exercice p. 5 Solution n°1**

Parmi les propositions ci-dessus, lesquelles sont justes ?

- ☐ Le référencement naturel consiste à afficher son site dans les liens sponsorisés sur la première page d'un moteur de recherche  
*Le référencement naturel (SEO) consiste à positionner son site dans les résultats naturels de recherche, contrairement au référencement payant (SEA), qui permet d'afficher son site dans les liens sponsorisés moyennant paiement des visites envoyées par le moteur de recherche.*
- ☒ Google est le moteur de recherche le plus utilisé en France  
*La part de marché de Google en France en 2020 est de l'ordre de 90 %, il domine largement le marché de la recherche sur Internet, que ce soit en France ou ailleurs dans le monde.*
- ☒ Les moteurs de recherche tiennent compte, entre autres, du contenu d'une page, des liens entrants et de l'expérience utilisateur pour déterminer son classement  
*Ces trois facteurs sont pris en compte par les moteurs, parmi d'autres indicateurs. Pour bien positionner une page, il est important de travailler ces aspects.*
- ☐ Il existe une technique infaillible pour positionner son site en tête des résultats  
*Évidemment, non, il n'existe pas de telle technique, et c'est bien logique : si elle existait, tout le monde l'appliquerait, et chaque site serait premier, ce qui est tout bonnement impossible. Il existe des techniques qui permettent d'améliorer le positionnement de son site web, sans pour autant garantir la première place.*
- ☒ SEO signifie Search Engine Optimization, appellation anglaise du référencement naturel  
*SEO signifie bien Search Engine Optimization, qui se traduit par Référencement Naturel. L'appellation anglaise est plus explicite : il s'agit bien d'optimiser son site pour les moteurs de recherche, et non d'attendre que celui-ci se positionne "naturellement" sans rien faire !*

**p. 10 Solution n°2**

```

1 User-agent: Googlebot
2 Disallow: /admin/
3 Disallow: /documentation/
4 Disallow: /deleted/
5
6 User-agent: Yahoo! Slurp
7 Disallow: /
8
9 User-agent: *
10 Allow: /
11
12 Sitemap: http://www.example.com/sitemap.xml

```

**p. 14 Solution n°3**

```

1 <ul>
2   <li itemscope itemtype="http://schema.org/Movie">
3     <h4 itemprop="name">Citizen Kane</h4>
4     <p itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
5       <span itemprop="ratingValue">8.3</span>
6     </p>
7
8     <span itemprop="director" itemscope itemtype="http://schema.org/Person">

```

```

9      <p itemprop="name">Orson Welles</p>
10    </span>
11  </li>
12  <li itemscope itemtype="http://schema.org/Movie">
13    <h4 itemprop="name">Le Parrain</h4>
14    <p itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
15      <span itemprop="ratingValue">8.5</span>
16    </p>
17
18    <span itemprop="director" itemscope itemtype="http://schema.org/Person">
19      <p itemprop="name">Francis Ford Coppola</p>
20    </span>
21  </li>
22  <li itemscope itemtype="http://schema.org/Movie">
23    <h4 itemprop="name">Pulp Fiction</h4>
24    <p itemprop="aggregateRating" itemscope itemtype="http://schema.org/AggregateRating">
25      <span itemprop="ratingValue">8.9</span>
26    </p>
27
28    <span itemprop="director" itemscope itemtype="http://schema.org/Person">
29      <p itemprop="name">Quentin Tarantino</p>
30    </span>
31  </li>
32 </ul>

```

#### Exercice p. 16 Solution n°4

##### Exercice

Le succès du référencement passe par...

- ☐ Le fait de répéter des mots-clés liés au thème du site
- ☒ D'écrire du contenu de qualité

##### Exercice

Il est préférable...

- ☒ D'avoir du contenu original et de créer des liens vers les ressources externes au besoin
- ☐ De copier les paragraphes de texte sur son site pour garder les utilisateurs plus longtemps sur son contenu

#### Exercice p. 18 Solution n°5

##### Exercice

Pour s'assurer du bon référencement de son site, il est important...

- ☐ D'ajouter un maximum de liens vers des sites externes
- ☒ D'avoir un maximum de sites externes qui proposent un lien vers son site

##### Exercice

L'échange de liens est une pratique consistant à...



- ☒ Faire un partenariat avec un gestionnaire de site pour référencer mutuellement vos sites web
- ☐ Acheter de l'espace publicitaire sur Google pour être dans les résultats les plus hauts

**Exercice p. 19 Solution n°6****Exercice**

La meilleure stratégie pour structurer son site web consiste à...

- ☐ Mettre toutes les pages à la racine du site, avec un lien vers chaque page sur la page d'accueil du site
- ☐ Créer une page par thématique, et inclure tous les sujets sous la forme de sous-parties de la page
- ☒ Créer une rubrique par thématique, chaque rubrique étant divisée en plusieurs pages traitant chacune un sujet précis
- ☐ Créer une page spécifique pour chaque mot-clé

**Exercice**

Par quel(s) moyens les robots des moteurs de recherche peuvent-ils découvrir les pages de votre site ?

- ☒ En suivant des liens depuis d'autres sites
- ☒ En suivant des liens depuis d'autres pages de votre site
- ☒ Grâce au fichier `sitemap.xml` présent à la racine de votre site
- ☐ Grâce au fichier `robots.xml` présent à la racine de votre site

**Exercice**

Sur mon site, la page <https://www.monsite.com/produit.php?id=3> présente le modèle de clavier de marque Techlogic et de référence KB123. Parmi les propositions suivantes, laquelle constitue le meilleur exemple de réécriture d'URL ?

- ☒ [https://www.monsite.com/claviers\\_sans\\_fil/techlogic/kb123.html](https://www.monsite.com/claviers_sans_fil/techlogic/kb123.html)
- ☐ <https://www.monsite.com/produits/kb123.html>
- ☐ <https://www.monsite.com/KB123/clavier.html>
- ☐ [https://www.monsite.com/claviers\\_sans\\_fil\\_techlogickb123.html](https://www.monsite.com/claviers_sans_fil_techlogickb123.html)

**Exercice**

Une balise `h1` est utilisée pour encadrer...

- ☐ Le titre d'un paragraphe
- ☐ Le titre d'une section de la page
- ☒ Le titre principal de la page
- ☐ Le titre du site

**Exercice**

Pour que les moteurs de recherche considèrent que ma page est pertinente pour un mot-clé, où ai-je intérêt à intégrer ce mot-clé ?

- ☒ Dans le titre de la page
- ☒ Dans le contenu de la page, au sein d'un contenu en lien avec ce mot-clé
- ☐ Dans la balise meta `keywords`, avec tous les mots-clés sur lesquels je veux positionner ma page
- ☐ Dans le fichier `robots.txt`

### Exercice

Quel élément de ma page est généralement repris comme titre dans les résultats de recherche ?

- ☒ La balise `title`
- ☐ La balise meta `description`
- ☐ La balise `h1`
- ☐ Les premiers mots affichés sur la page

### Exercice

Comment se définit le taux de rebond sur une page web ?

- ☐ Le nombre de clics par visiteur sur les liens d'une page web
- ☐ La proportion de visiteurs qui a visité cette page web, puis a visité d'autres pages sur le même site
- ☐ La proportion de visiteurs qui est resté sur la page pour consulter son contenu
- ☒ La proportion de visiteurs qui a quitté rapidement la page sans consulter son contenu

### Exercice

Parmi les pratiques suivantes, lesquelles représentent un danger pour mon référencement sur un moteur de recherche ?

- ☐ Acheter des liens sponsorisés sur un moteur concurrent
- ☒ Proposer des contenus dupliqués sur mon site
- ☐ Obtenir des liens en faisant du guest blogging
- ☒ Développer un réseau de pages satellites sans contenu pertinent
- ☒ Pratiquer la répétition excessive de mots-clés dans les contenus de mes sites

### Exercice

Pour maximiser les chances de référencer mon site sur la requête "Planter des tomates", comment dois-je rédiger mes pages de contenu ?

- ☐ Je dois inclure cette requête dans tous les titres et sous-titres de ma page, et répéter de nombreuses fois les mots "planter" et "tomates" dans mes textes, en les mettant en gras
- ☐ Je dois essayer de me mettre à la place du robot du moteur de recherche et faire des phrases courtes et simples pour qu'il les comprenne bien
- ☒ Je dois essayer de me mettre à la place du visiteur qui a tapé cette requête dans Google et essayer de lui proposer un contenu qui explique de manière pertinente et bien illustrée comment planter des tomates
- ☐ Je dois copier/coller le texte de la page qui est en première position sur la requête "Planter des tomates" et changer quelques mots pour éviter les problèmes de droits d'auteur

**Exercice**

Une page web ne peut être référencée que sur un seul mot-clé ou une seule expression.

- ☐ Vrai
- ☒ Faux

**Exercice**

Les liens qui apportent une valeur ajoutée à mes pages en termes de référencement naturel sont...

- ☒ Les liens naturels, que les visiteurs de mon site ajoutent spontanément dans des posts de forum, sur des blogs ou réseaux sociaux
- ☐ Les liens sponsorisés
- ☐ Les liens "Nofollow"
- ☒ Les liens intégrés dans des contenus (blogs ou sites) partenaires, dont la thématique est proche de celle de mon site

**p. 21 Solution n°7**

```

1 <!doctype html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5   <title>Activités de la ville de Saint Martin</title>
6   <meta name="description" content="Découvrir, apprendre, échanger : la ville de Saint Martin
  propose un large choix de clubs sportifs et associations culturelles." />
7 </head>
8 <body>
9   <h1>Activités de la ville de Saint Martin</h1>
10  <p>Bienvenue sur le site de la ville de Saint Martin. Découvrir, apprendre, échanger : la
  ville de Saint Martin propose un large choix de clubs sportifs et associations culturelles.
  </p>
11  <h2>Types d'activités</h2>
12  <a href="activites_sportives.html">Activités sportives</a>
13  <a href="activites_culturelles.html">Activités Culturelles</a>
14 </body>
15 </html>

```

Il fallait remplacer la balise div du titre par une balise h1, et la balises div du sous-titre de la page par une balise h2, afin de structurer correctement le contenu.

**p. 22 Solution n°8**

```

1 <!doctype html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5   <title>Activités sportives de la ville de Saint Martin</title>
6   <meta name="description" content="Découvrir, apprendre, échanger : la ville de Saint Martin
  propose un large choix de clubs sportifs." />
7 </head>
8 <body>
9   <div class="navigation"><a href="index.html">Activités de la ville de Saint Martin</a> >
  Activités sportives</div>
10  <h1>Activités sportives de la ville de Saint Martin</h1>

```

```

11 <p>Bienvenue sur le site de la ville de Saint Martin. Découvrir, apprendre, échanger : la
    ville de Saint Martin propose un large choix de clubs sportifs.</p>
12 <h2>Sports individuels</h2>
13 <a href="tennis_club_saint_martin.html">Tennis club de Saint Martin</a>
14 <a href="golf_de_saint_martin">Golf de Saint Martin</a>
15 <h2>Sports collectifs</h2>
16 <a href="football_club_saint_martin.html">Football club de Saint Martin</a>
17 <a href="saint_martin_basket.html">Saint Martin Basket</a>
18 </body>
19 </html>

```

Il fallait remplacer la balise div du titre par une balise h1, et les balises div des sous-titres de la page par des balises h2 afin de structurer correctement le contenu.

Il fallait également ajouter un lien retour vers la page d'accueil sur le fil d'Ariane (div class="navigation").

Il fallait enfin ajouter la balise meta description manquante.

## p. 22 Solution n°9

```

1 <!doctype html>
2 <html>
3 <head>
4   <title>Tennis Club de Saint Martin</title>
5   <meta name="description" content="Découvrir, apprendre, échanger : la ville de Saint Martin
    propose pratiquer au Tennis Club." />
6 </head>
7 <body>
8   <div class="navigation"><a href="index.html">Activités de la ville de Saint Martin</a> > <a
    href="activites_sportives.html">Activités sportives</a> > Tennis club</div>
9   <h1>Tennis Club de Saint Martin</h1>
10  <p>Bienvenue sur la page du Tennis Club de la ville de Saint Martin. Découvrir, apprendre,
    échanger : la ville de Saint Martin propose de pratiquer le Tennis grâce à son club
    municipal.</p>
11  <h2>Présentation</h2>
12  <p>Le club de Tennis municipal de Saint Martin vous offre quatre courts en plein-air et un
    court couvert ...</p>
13  <h2>Coordonnées</h2>
14  <h3>Président du Club</h3>
15  <p>Nicolas Durand<br/>
    10 rue des fleurs
    Saint Martin</p>
18  <p>Téléphone : 01 23 45 67 89</p>
19  <h3>Entraîneur</h3>
20  <p>Gérard Lebon<br/>
    1 avenue des platanes
    Saint Martin</p>
23  <p>Téléphone : 01 23 98 76</p>
24 </body>
25 </html>

```

Il fallait remplacer la balise div du titre par une balise h1, les balises div des sous-titres de la page par des balises h2, et la balise div englobant la fonction par une balise h3, afin de structurer correctement le contenu.

Il fallait également ajouter un lien retour vers la page d'accueil et un lien retour sur la page rubrique sur le fil d'Ariane (div class="navigation").

Il fallait enfin ajouter la balise meta description manquante.

## Contenus annexes

### 1. Visual Studio Code

#### Objectifs

- Comprendre pourquoi utiliser un éditeur de texte
- Découvrir et installer notre premier éditeur de texte

#### Mise en situation

Afin de créer notre premier site web, il va nous falloir télécharger les bons outils et comprendre en quoi ils nous sont utiles.

Définition	Présentation
------------	--------------

Les pages web HTML sont écrites dans des fichiers textes avec l'extension *.html*.

Ce sont des fichiers contenant du texte, au même titre qu'une image peut se trouver sous l'extension *.jpg* ou encore les chansons que l'on trouve dans des fichiers *.mp3*.

Il faut un logiciel pour les créer et les éditer : pour cela, nous utiliserons un éditeur de texte.

Conseil
---------

On pourrait utiliser un éditeur de texte standard (Bloc-Notes, vi, emacs...), mais, pour des raisons de confort, nous utiliserons un éditeur de texte plus avancé : Visual Studio Code. Il permettra notamment de colorer automatiquement les éléments du code, afficher des aides, etc.

#### Installation

Maintenant, installons notre premier logiciel qui nous permettra d'écrire du HTML : <https://code.visualstudio.com/>.

À la fin de l'installation, ouvrons le logiciel Visual Studio Code :

- Cliquez sur le menu "File" puis "New file", un nouveau fichier vierge apparaît au centre.
- Sauvegardez le nouveau fichier au format *.html*. Par exemple, appelons-le *premiere-page.html*.
- Écrivez une phrase (par exemple, "Voici un fichier HTML") et sauvegardez.
- Double-cliquez sur le fichier pour l'ouvrir dans le navigateur ou effectuez un glisser-déposer vers votre navigateur préféré, le texte va s'afficher !

Les pages web HTML sont des fichiers textes presque comme les autres !

Syntaxe	À retenir
---------	-----------

- Un éditeur de texte est un logiciel qui permet de créer et d'éditer des fichiers notamment des fichiers textes.
- Pour écrire du code HTML, des éditeurs plus adaptés que les éditeurs de texte standard existent, offrant de multiples fonctionnalités.