

2 : LA BALISE META

Explications :

Les balises META servent à placer des **métadonnées** (metadata) dans une page HTML. On placera ces informations dans l'élément **head**, et elles ne seront pas affichées sur la page.

À qui sont destinées ces informations ? A tous les outils susceptibles de les exploiter.
Et notamment :

- Aux navigateurs web ;
- Aux moteurs de recherche ;
- Plus largement, à tous les outils d'indexation, c'est-à-dire tous les outils qui analysent vos pages pour y identifier diverses informations.

Anatomie d'une balise META :

Lorsqu'on parle d'une balise META, on pense en général à un élément meta qui utiliserait des attributs et valeurs précises. On parle par exemple de « META Description » pour la balise suivante :

```
index.html
```

```
<meta name=' 'description' ' content=' 'Une description de la page...' ' />
```

Meta description

Australian Shepherd Dog Breed Information and Pictures

www.dogbreedinfo.com/australianshepherd.htm ▼

All about the Australian Shepherd, info, pictures, breeders, rescues, care, temperament, health, puppy pictures and more.

Voici un deuxième exemple avec une META Content-Type et une META Keywords :

```
index.html
```

```
<meta http-equiv=' 'Content-Type' ' content=' 'text/html; charset=ISO-8859-1' ' />
```

```
<meta name=' 'keywords' ' content=' 'réservation, chambre, shrek' ' />
```

Le nom que l'on place juste après le mot « META » correspond au mot-clé utilisé dans un attribut précis :

suivant les cas, il s'agit de l'attribut **name** ou de l'attribut **http-equiv**.

On trouvera toujours dans la balise un deuxième attribut, l'attribut **content**, qui contiendra la valeur associée au mot-clé.

Est-ce que c'est vraiment utile ?

Les balises META sont un mécanisme qui permet d'associer des métadonnées à un document (une page web). **Dans l'absolu, ce mécanisme est utile.** La vraie question, c'est celle de **l'utilité de chaque balise META.**

Nous disions que les métadonnées insérées grâce aux balises META sont destinées à divers outils capables de les exploiter. Cela a deux conséquences :

- Si aucun outil n'utilise une information précise, alors donner cette information est inutile.
- S'il existe des outils qui exploitent une information précise mais que cette exploitation ne nous intéresse pas, ou que l'information n'est pas pertinente pour le site sur lequel on travaille, on passera également son chemin.

Dans l'ensemble les balises META ont peu d'importance pour le référencement. Mais puisque c'est au cas par cas, qu'en est-il des cas particuliers ? Passons en revue les balises META courantes, en indiquant à chaque fois leurs utilités.

Les incontournables :

META Content-Type :

```
index.html  
  
<meta charset=' 'utf-8' '>
```

<meta charset> va permettre d'indiquer l'encodage du document HTML, **"utf-8"** est un codage de caractères informatiques conçu pour coder l'ensemble des caractères du « répertoire universel de caractères codés ».

Il est recommandé de la placer au tout début de la section **<head>**, surtout avant **<title>**, car ce dernier peut déjà être concerné par l'encodage des caractères.

Utiliser cette balise META n'est pas indispensable. En effet, les informations qu'elle donne devraient logiquement être envoyées par le serveur web, au moment où il envoie la page HTML au navigateur. Mais redoubler cette information dans la page HTML elle-même permet d'avoir une information sur l'encodage des caractères utilisé (UTF-8 ou ISO-8859-1 ou d'autres encore), ce qui s'avèrera très utile dans les cas suivants :

- Le serveur web, mal configuré, n'envoie pas les informations sur l'encodage. Le navigateur web va alors tenter de les trouver dans le code HTML de la page.
- Dans tous les cas où la page HTML est un fichier local, ouvert directement par l'utilisateur et pas consulté depuis un serveur web. Par exemple pour la documentation en HTML fournie avec un logiciel, pour une page d'information sur un CD ou DVD, lorsque l'utilisateur enregistre une page web dans ses documents, etc.

META Description :

Comme vu ci-dessus sur la capture d'écran, cette META permet de fournir une description courte (deux à trois phrases courtes) de la page web. Cette description est notamment utilisée (dans certains cas) par les principaux moteurs de recherche, pour fournir le court texte d'aperçu des pages de résultats, ou bien par les réseaux sociaux pour définir le paragraphe de texte associé au lien partagé.

Pour rappel, voici un exemple de balise META Description :

```
index.html
```

```
<meta name=' 'description' ' content=' 'Page de réservation de chambre  
de l'hôtel Le Marais. Toujours à des tarifs abordables.' ' />
```

On conseille généralement de ne pas utiliser la même description pour les différentes pages d'un site. Dans l'idéal, chaque page aura sa propre description habilement composée. Voici quelques conseils pour la rédaction de cette description :

- Elle doit être unique, et pas dupliquée sur d'autres pages du site ;
- Elle doit refléter fidèlement le contenu de la page ;
- Elle doit être attractive pour l'utilisateur dont le regard « scanne » une page de résultats de moteur de recherche ;
- Elle doit reprendre les principaux mots-clés de la page et être plutôt brève (on recommande de ne pas dépasser 200 caractères, bien qu'il ne s'agisse pas d'une règle stricte).

On ne dispose pas toujours d'un contenu spécifique pour la META Description de chaque page. Dans ce cas, trois stratégies possibles :

1. Ne pas utiliser la balise META Description du tout ;
2. Utiliser une META Description seulement pour certaines pages clé, ou uniquement pour la page d'accueil ;
3. Extraire un contenu des contenus disponibles en base de données, par exemple le texte d'accroche pour une page d'article ou le début de la description du produit pour une page produit.

Notez qu'une page ne sera pas « pénalisée » par un moteur de recherche parce qu'elle n'utilise pas de META Description. Mais si vous n'en utilisez pas, faites attention à ne pas utiliser la balise du tout, et évitez les balises sans contenu.

Les indications données aux robots d'indexation :

Vous pouvez utiliser les balises META pour donner des indications aux robots d'indexation des moteurs de recherche (indications qu'ils seront libres de suivre ou pas, mais les principales sont bien appliquées).

On utilisera pour cela la META Robots :

```
index.html
```

```
<meta name=' 'robots' ' content=' 'commande1, commande2, commande3' ' />
```

Les commandes sont des mots prédéfinis, séparés par des virgules. On peut aussi utiliser plusieurs balises META à la suite pour les différentes commandes :

```
index.html
```

```
<meta name=' robots' content=' commande1' />
<meta name=' robots' content=' commande2' />
<meta name=' robots' content=' commande3' />
```

Restreindre l'indexation et les données affichées en page de résultat :

Voici un ensemble de commandes pour restreindre l'indexation :

- **noindex**: pas d'indexation de la page;
- **nofollow**: ne pas transmettre le potentiel de positionnement de la page;
- **noarchive**: ne pas afficher de lien «En cache» dans les pages de résultat;
- **none**: raccourci pour noindex, nofollow, noarchive.

À cela se rajoutent quelques directives, souvent propres à un moteur donné, telles que :

- **noodp**: ne pas utiliser les descriptions de l'Open Directory Project dans les pages de résultat (Google...);
- **nosnippet**: ne pas utiliser de description ou d'extrait dans les pages de résultats, donc utiliser le titre de page (et dans certains cas les sitelinks) uniquement (Google...);

Il est aussi possible d'adresser les différentes directives à un moteur de recherche donné, plutôt qu'à tous les moteurs de recherche. Pour cela, on remplace le mot-clé robots par le nom du **robots** d'indexation du moteur visé. Par exemple :

```
index.html
```

```
<meta name=' Googlebot' content=' directives pour Google Search' />
<meta name=' Slurp' content=' directives pour Yahoo! Search' />
<meta name=' MSNbot' content=' directives pour Live Search' />
```

Alternatives à la META Robots :

Pour adresser des directives aux moteurs de recherche, il existe deux autres solutions : le fichier **robots.txt** et l'en-tête HTTP **X-Robots-Tag**.

Le fichier **robots.txt** permet d'interdire l'accès à certaines pages ou certains dossiers aux robots d'indexation qui en respectent les directives (c'est le cas des moteurs de recherche majeurs). Il se place à la racine du domaine ou sous-domaine. Les fonctionnalités disponibles sont plus limitées qu'avec une balise META Robots, mais les directives d'un **robots.txt** ont l'avantage de s'appliquer de manière globale à un sous-domaine complet. Utiliser ce fichier est bien souvent suffisant pour gérer l'indexation ou la non-indexation des différentes pages d'un site.