
Cours Dreamweaver CS5 - Partie 13

Balise Meta

Pierre Tomy Le Boucher

Sommaire

Balises Meta.....	3
C'est quoi une balise META?.....	3
Anatomie d'une balise META.....	3
La grande question: est-ce que c'est vraiment utile?.....	4
Les incontournables.....	4
Élément title (utile et obligatoire, mais pas une balise META).....	4
META Content-Type (très utile).....	5
META Description (utile).....	6
Informations de classement.....	7
META Keywords (obsolète).....	7
META Generator (utile dans certains cas).....	7
META Author, META Copyright (utilité faible).....	8
Les META Dublin Core (utilité faible).....	8
Les équivalents d'en-têtes HTTP (http-equiv).....	8
Faut-il utiliser les META http-equiv?.....	10
Rappel: évitez les META Refresh!.....	10
Les indications données aux robots d'indexation.....	10
Demander aux robots d'indexer vos pages.....	11
Restreindre l'indexation et les données affichées en page de résultat.....	11
Alternatives à la META Robots.....	11
Balises META liées à un outil précis, et cas particuliers.....	12
META verify-v1 (usage particulier).....	12
META Revisit ou Revisit-After (inutile, mythe du Web).....	12
META MS Smart Tags (inutile).....	13

Balises Meta

Il y avait avant, il a maintenant et il y a sûrement plus tard, alors !!!

Avant !!!!!, le référencement était lié aux balises méta, mais ça s'était avant, alors maintenant c'est

C'est quoi une balise META?

(merci à alsacréation du prof. Nirentet)

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? Eh bien, à 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

Lorsque l'on parle de «balises META», on parle en réalité de l'élément meta en HTML. Si on tient à parler de balises, on devrait plutôt parler de «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:

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

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

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

```
<meta name="keywords" content="péniche, écluses, paquebot, crise, décroissance" />
```

Vous l'aurez peut-être remarqué, 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.

Une balise META est donc une paire « mot-clé-valeur », transcrite en HTML de la manière suivante:

```
<meta name|http-equiv="Mot-clé" content="Valeur" />
```

(Attention: la notation ci-dessus indique qu'il faut utiliser soit l'attribut « name », soit l'attribut « http-equiv », mais pas de manière arbitraire. Chaque « balise META » utilisera l'un ou l'autre attribut.)

La grande question: est-ce que c'est vraiment utile?

On peut lire ici ou là que les balises META sont « obsolètes ». Juste à côté, dans un forum consacré au référencement, on pourra lire des messages qui conseillent de « renseigner les balises META ». Et si on a le malheur de demander si c'est vraiment utile, on aura droit à des réponses laconiques du genre: « Ça ne sert pas à grand chose, mais autant le faire. » La réponse la plus juste que l'on fera ici est la suivante: « **ça dépend** ».

En effet, 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 disons que les métadonnées insérées grâce aux balises META sont destinées à divers outils capables de les exploiter. Cela a deux conséquences:

1. Si aucun outil n'utilise une information précise, alors donner cette information est inutile.
2. 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.

(Le premier point peut éventuellement être discuté, et certains préconisent le principe de précaution: « on ne sait jamais, ça pourrait être exploité à l'avenir par tel ou tel type d'outil ». Pourquoi pas, mais attention à ne pas perdre trop de temps sur du « on ne sait jamais » alors qu'il y aurait énormément de problèmes avérés sur lesquels travailler.)

Ce qui est certain, c'est que dans l'ensemble les balises META ont peu d'importance pour le référencement. N'en faites donc pas une obsession ou une maladie ! ;) Voilà pour le principe. Mais puisque c'est du cas par cas, qu'en est-il des cas particuliers? Nous vous proposons de passer en revue les balises META courantes, en indiquant à chaque fois leur utilité.

Les incontournables

Élément *title* (utile et obligatoire, mais pas une balise META)

De nombreux sites parlent de « META Title » pour désigner cet élément, bien qu'il ne s'agisse absolument pas d'une balise META. Il s'agit d'un élément HTML différent de « meta », qui s'écrit généralement ainsi:

```
<title>Titre de la page</title>
```

Le contenu de cet élément n'apparaît pas directement dans la page, mais est utilisé de diverses manières:

- par le navigateur web, dans le titre de la fenêtre, les onglets, comme titre des marque-pages (favoris), etc.;
- par les moteurs de recherche et divers outils, comme titre d'un résultat de recherche par exemple.

Cet élément est par ailleurs obligatoire en HTML.

Enfin, notez que son importance pour le référencement des pages est primordiale. Un élément « title » bien rédigé devrait, dans l'idéal :

- être lisible pour un lecteur humain,
- être informatif ou attrayant pour un lecteur humain,
- refléter le contenu de la page,
- contenir des « mots-clé » pertinents,
- faire moins de 70 caractères (pas une règle absolue, mais Google par exemple coupe les titres trop longs entre 65 et 70 caractères dans les pages de résultat, et les titres trop longs diluent le poids des différents « mots-clé »).

Le sujet est vaste mais nous en resterons ici. Revenons à nos balises META...

META Content-Type (très utile)

Si vous n'en utilisez qu'une, ce sera sans doute celle-là. Elle se présente ainsi...

En HTML 4 et XHTML 1.x :

```
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
```

Il s'agit d'un équivalent, placé dans le code HTML, de l'en-tête HTTP Content-Type, qui permet d'indiquer le type de document et son encodage.

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 utilisés (« 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.

Nous reviendrons sur les différentes META utilisant l'attribut « http-equiv » un peu plus loin.

En HTML 5 :

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

Comme vous pouvez le constater, elle a été considérablement simplifiée avec la venue de HTML5. 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.

META Description (utile)

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 (Facebook, Google+) pour définir le paragraphe de texte associé au lien partagé.

Pour rappel, voici un exemple de balise META Description:

```
<meta name="description" content="Tout sur les apparitions muettes de  
la Banane Masquée dans le cinéma colombien des années 20 et 30." />
```

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-clé » 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:

```
<meta name="description" content="" /> <!-- À éviter -->
```

Dernière chose: fournir une META Description ne garantit pas qu'elle sera effectivement utilisée par les moteurs de recherche dans leurs pages de résultats!

Informations de classement

META Keywords (obsolète)

Utilisée par les premiers moteurs de recherche, cette balise META permettait d'indiquer un certain nombre de « mots-clé » à associer à la page. Petit problème: cette possibilité d'associer des « mots-clé » à une page sans afficher ce contenu dans la page permet de « spammer » les moteurs en utilisant des « mots-clé » qui ne reflètent pas le contenu de la page. C'est pourquoi la plupart des moteurs de recherche:

- soit ne l'ont jamais prise en compte;
- soit ont arrêté de la prendre en compte;
- soit la prennent toujours en compte mais lui donnent un poids très faible.

Il est difficile de dire ce que fait chaque moteur car ils communiquent peu ou pas du tout sur leurs choix en la matière. Fin 2007, cependant, Google annonçait déjà ne pas l'utiliser. Pour ce qui est du format, on utilisera une seule balise META Keywords par page, avec pour contenu une série de termes (avec un ou plusieurs mots) séparés par des virgules:

```
<meta name="keywords" content="tartiflette, saucisson, maroilles, ..." />
```

Si vous souhaitez l'utiliser malgré sa probable inutilité, voici quelques conseils que l'on retrouve souvent:

- ne pas utiliser « trop » de « mots-clé » (le critère est nécessairement vague...);
- ne pas utiliser de « mots-clé » sans rapport avec la page, voire ne pas utiliser de mot-clé que l'on ne retrouverait pas dans le contenu de la page.

Attention également: si vous avez consciencieusement étudié votre stratégie de positionnement et défini un corpus de « mots-clé », il n'est pas forcément très intelligent de les lister dans vos pages, en les exposant ainsi aux regards de vos concurrents.

META Generator (utile dans certains cas)

Certains éditeurs de code et logiciels de gestion de contenu (CMS) placent automatiquement une META Generator dans vos pages. Elle peut ressembler à ceci:

```
<meta name="generator" content="Nom de l'outil (version)" />
```

Cela permet aux éditeurs du logiciel de faire des statistiques: combien de pages sont construites avec leur outil? combien de sites utilisent leur CMS? les sites qui utilisent leur CMS sont-ils rapidement passés à la dernière version majeure? etc.

En tant qu'éditeur ou créateur de site, cette information ne vous sert à rien. Mais si vous utilisez un CMS libre fourni gracieusement par des bénévoles, ça ne vous coûtera pas grand chose de garder cette ligne de code dans vos pages HTML et ça leur fournira quelques informations utiles.

META Author, META Copyright (utilité faible)

Ces balises permettent de fournir le nom de l'auteur d'une page et une information sur les droits d'auteurs (auteur, année, mention de licence ou mention de type «tous droits réservés»). Utile surtout si vous voulez rappeler aux curieux qui lisent le code source des pages que le site a un auteur (personne ou organisation), et que tout n'est pas permis.

Pas exploité par les principaux moteurs de recherche. Peut-être exploité par certains outils très spécifiques. On se passera de ces informations finalement peu utiles si on souhaite garder un code HTML plus compact.

```
<meta name="author" content="Mary Poppins" />
```

```
<meta name="copyright" content="© Supercalifragilisticexpialidocious  
Inc." />
```

On trouve également, citées dans certains articles, des META Designer, META Publisher, etc. Il est peu probable que ces informations soient exploitées par le moindre outil. Et ces informations ont plus leur place sur des pages «À propos» et «Mentions légales» de vos sites. Bien entendu, dans beaucoup de cas on pourrait faire la même remarque pour les META Author et Copyright.

Les META Dublin Core (utilité faible)

Cette série de balises META reprend les types de métadonnées proposées par la « Dublin Core Initiative ». Elles sont décrites dans l'article « Expressing Dublin Core in HTML/XHTML meta and link elements ».

En français on pourra lire l'article publié sur « Openweb: Métadonnées et Dublin Core ». Voici un exemple de code issu de cet article:

```
<meta name="DC.Title" lang="fr" content="Un document en HTML" />
```

```
<meta name="DC.Date.created" scheme="W3CDTF" content="2003-04-03" />
```

```
<meta name="DC.Date.modified" scheme="W3CDTF" content="2003-04-27" />
```

Aux dernières nouvelles, les moteurs de recherche majeurs ignorent les métadonnées « Dublin Core ». Elles n'ont donc pas d'application concrète pour le référencement, mais peuvent servir de base standardisée à utiliser (en HTML, XML ou autre format) pour associer des métadonnées à des documents, selon les besoins. Elles trouvent des applications dans les domaines de l'archivage et de la gestion documentaire.

Les équivalents d'en-têtes HTTP (http-equiv)

Les en-têtes HTTP « (HTTP headers) » font partie d'un des standards fondamentaux du Web: « l'HyperText Transfer Protocol ». Ce sont des métadonnées ou indications qu'un serveur web envoie à chaque fois qu'il envoie un fichier (page web, feuille de styles, image, etc.) à un navigateur web.

Voici à quoi peuvent ressembler les en-têtes HTTP envoyées avec un fichier HTML:


```
Date: Wed, 30 Feb 2009 16:50:59 GMT
Server: Apache
Last-Modified: Mon, 11 Jan 2009 12:28:37 GMT
Etag: "6407f-459e-f44f3740"
Accept-Ranges: bytes
Cache-Control: max-age=600
Expires: Wed, 30 Feb 2009 17:00:59 GMT
Vary: Accept-Encoding
Content-Encoding: gzip
Content-Length: 4516
Content-Type: text/html; charset=UTF-8
Content-Language: fr-CA

200 OK
```

Cela fait pas mal d'informations. Vous reconnaissez peut-être la dernière: 200 OK (tout va bien). Si le fichier n'avait pas été disponible, le serveur aurait pu renvoyer un « 404 Not Found » par exemple. Vous pouvez retrouver toutes ces informations dans les consoles de développement, souvent dans l'onglet Réseau ou Network.

Parmi les autres informations, on retrouve l'en-tête « Content-Type » qui informe le navigateur sur le type de fichier (ici « text/html », mais pour d'autres fichiers on pourrait avoir « text/plain, text/css, image/png, etc. »); et qui donne également une information sur l'encodage des caractères. Attention: ces informations sont prioritaires sur la META Content-Type, qui devrait normalement les reprendre à l'identique lorsqu'elle est utilisée. On retrouve aussi l'en-tête « Content-Language », qui précise la langue principale du document; dans l'exemple ci-dessus, il indique que le public visé lit le français du Canada (québécois).

Si on voulait reprendre à l'identique quelques-uns de ces en-têtes HTTP, on pourrait par exemple écrire:

```
<meta http-equiv="Content-Language" content="fr-CA" />
<meta http-equiv="Cache-Control" content="max-age=600" />
<meta http-equiv="Expires" content="Wed, 30 Feb 2009 17:00:59 GMT" />
```

Faut-il utiliser les META http-equiv?

En dehors de l'information sur l'encodage de caractères donnée par l'en-tête « HTTP Content-Type », il n'est pas utile de dupliquer les informations de vos en-têtes HTTP en utilisant la balise META équivalente.

Il ne s'agit cependant pas d'une règle stricte, et il peut y avoir des exceptions.

Enfin, même s'il est préférable de s'assurer que les bonnes informations sont présentes dans les en-têtes HTTP, dans certains cas il sera peut-être nécessaire d'utiliser les balises META à la place des en-têtes HTTP, par exemple si vous n'avez pas accès à la configuration du serveur et que vous souhaitez tout de même envoyer certaines informations aux navigateurs web.

Rappel: évitez les META Refresh!

Trop de sites encore utilisent une balise « META Refresh » pour forcer un rechargement de la page par le navigateur. Suivant les cas, il s'agit de faire ainsi une redirection de page au bout d'un laps de temps, ou bien de forcer le rafraichissement de page pour renouveler les publicités affichées.

Lorsque le rafraichissement n'est pas logique, attendu par l'utilisateur, il est ressenti comme une perturbation: le site ou le navigateur agissent sans son aval. Lorsque vous êtes en train de lire un article sur un site d'actualité et que vous perdez le fil de votre lecture à cause d'un rafraichissement, c'est pénible. Si en plus vous utilisez un lecteur d'écran, ça devient très problématique car récupérer le fil de votre lecture devient sensiblement plus compliqué.

Les différentes normes et méthodes d'application pour l'accessibilité des sites web proscrivent l'usage des META Refresh.

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: »

```
<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:

```
<meta name="robots" content="commande1" />
```

```
<meta name="robots" content="commande2" />
```

```
<meta name="robots" content="commande3" />
```

Demander aux robots d'indexer vos pages

Inutile de demander aux robots d'indexation d'indexer vos pages: c'est leur comportement par défaut! Ainsi les deux lignes suivantes, en plus d'être équivalentes, sont tout à fait inutiles:

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

```
<meta name="robots" content="all" />
```

Elles illustrent toutefois le principe de fonctionnement de cette balise « META »: on utilise une balise « META », et on y liste un certain nombre d'options prédéfinies, séparées par des virgules.

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 (MSN, Google, Yahoo!);
- noydir: ne pas reprendre les descriptions de l'annuaire Yahoo! dans les pages de résultat (Yahoo!).
- 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, Yahoo!);

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 robot d'indexation du moteur visé. Par exemple:

```
<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'entête « 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.

L'en-tête « HTTP X-Robots-Tag » a été proposé par Google en 2006, et Yahoo! a annoncé son support en 2007. Il permet d'appliquer les mêmes directives que la balise META Robots.

Balises META liées à un outil précis, et cas particuliers

Il n'existe pas de liste exhaustive des balises « META » car tout un chacun est libre de créer un couple « clé-valeur », pour peu qu'il en ait l'usage.

Ainsi, certains outils vous proposent de placer des balises « META » très spécifiques dans vos pages pour répondre à leurs besoins. C'est la documentation de chaque outil, ou les instructions qu'il vous donne, qui vous apprendront à quoi sert la balise en question.

META verify-v1 (usage particulier)

Voici un exemple courant, que vous verrez peut-être sur certains sites (ou plutôt dans leur code source):

```
<meta name="verify-v1" content="chaîne aléatoire" />
```

Cette balise « META » est utilisée par Google pour ses outils Google Webmaster Tools. Elle permet de vérifier qu'un utilisateur est bien le propriétaire d'un site.

META Revisit ou Revisit-After (inutile, mythe du Web)

La « META Revisit », inventée par un petit moteur de recherche régional désormais tombé dans l'oubli « (searchBC, canadien) », n'a jamais été utilisée par d'autres outils. Malgré son inutilité flagrante, de nombreux webmasters s'évertuent à l'utiliser, pensant à tort que les moteurs de recherche l'utilisent dans une certaine mesure. En cherchant « meta revisit-after » sur le Web, vous pourrez lire nombre d'articles qui perpétuent le mythe de cette balise « META ».

Faut-il vraiment donner un exemple de code? Peut-être, oui, juste pour vous aider à reconnaître les avatars du mythe lorsque vous le rencontrerez sur d'autres sites.

La forme la plus connue et la plus reprise utilise le mot-clé «Revisit-After», mais on trouve aussi simplement «Revisit»:

```
<meta name="revisit-after" content="14 days" />
```

```
<meta name="revisit" content="7 days" />
```

Pour préciser (de manière purement indicative) la fréquence de mise à jour de vos contenus, vous pouvez utiliser une Sitemap XML.

META MS Smart Tags (inutile)

La technologie «Smart Tags» a été intégrée par Microsoft dans une version beta d'Internet Explorer. Le principe était que Microsoft pouvait vendre à des annonceurs un mot-clé ou une expression (comme le font de nombreuses régies publicitaires, dont celles de Google et Microsoft), et lorsqu'un utilisateur visitait une page web contenant cette expression, l'expression devenait un lien spécial vers le site de l'annonceur.

Cette très mauvaise idée, largement critiquée, a été abandonnée et n'a jamais fait partie d'une version finale d'Internet Explorer. Microsoft ne propose plus ce service. Malgré cela, on trouve encore de nombreux sites (même récents) avec la balise META censée désactiver ce comportement:

```
<meta name="MSSmartTagsPreventParsing" content="true" />
```

À oublier.