

Le langage HTML

I. La notion de « balise » :

L'**HyperText Markup Language**, généralement abrégé HTML, est le **langage de balisage** conçu pour représenter les pages web. On peut le traduire par « langage de balisage des liens hypertextes ».

Si le World Wide Web et le protocole HTTP furent créés en 1989, on admet qu'il s'est stabilisé sous sa forme actuelle en 1993. On utilise aujourd'hui la version 5, normalisée par le W3C (**World Wide Web Consortium**, organisme de standardisation à but non lucratif, fondé en octobre 1994 et chargé de promouvoir la compatibilité des technologies).

HTML a été conçu pour baliser (ou marquer) simplement le texte, notamment pour y ajouter des hyperliens. A l'origine, on utilisait un minimum de balises, comme dans le document HTML ci-contre.

```
<TITLE>Exemple de HTML</TITLE>
Ceci est une phrase avec un <A HREF=cible.html>hyperlien</A>.
<P>
Ceci est un paragraphe où il n'y a pas d'hyperlien.
</P>
```

Le titre se trouve entre les balises <TITLE> et </TITLE> et le paragraphe se trouve entre les balises <P> et </P>. L'hyperlien est entre les balise et .

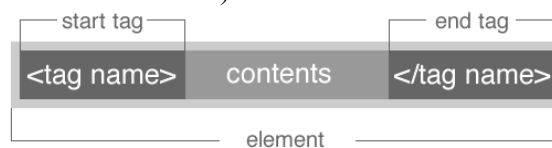
Aujourd'hui, on peut baliser différents types de données : des tableaux, des images, des vidéos, etc.

Syntaxe des balises :

Une **balise** est composée avec des caractères spéciaux : <, > et /. Elles sont interprétées par les navigateurs pour composer un **élément** HTML.

Décomposition des éléments HTML :

Les objets à baliser sont toujours entourés par deux balises : une **balise ouvrante** et une **balise fermante** (celle avec un slash).



Exemples :

Les balises mon_image.jpeg marquent une image.

Les balises <audio>mon_son.mp3</audio> marquent un son.

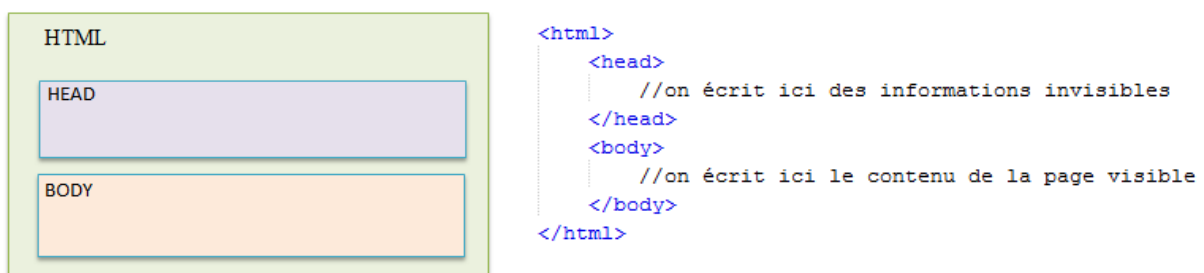
II. Structure d'une page HTML :

Les pages HTML sont standardisées pour pouvoir être ouvertes par n'importe quel navigateur et sur n'importe quelle machine.

Elles commencent toujours par la balise ouvrante <html> et se terminent toujours par la balise fermante </html>.

Ensuite, elles sont constituées de deux parties :

- ❖ L'en-tête, appelé « HEAD », contient des informations invisibles pour la personne qui consulte la page. On y place des métadonnées comme l'auteur de la page, des mots-clés pour les moteurs de recherche, les feuilles de style pour la mise en page, les scripts des pages interactives (hors programme), etc.
L'en-tête commence toujours par la balise ouvrante <head> et se termine toujours par la balise fermante </head>.
- ❖ Le corps, appelé « BODY », contient les informations visibles de la page.
Le corps de page commence toujours par la balise ouvrante <body> et se termine toujours par la balise fermante </body>.

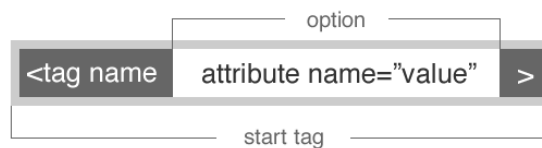


A gauche, un schéma simplifié de la structure d'une page Web. A droite, la structure avec les balises du langage HTML.

III. Précisions sur les balises :

Un élément HTML peut avoir des **attributs** pour affiner son sens.

Ces attributs sont spécifiés dans la balise de début. Ils se composent d'un nom et une valeur, séparés par un caractère "=".



Exemple : <h2 id="titre_1">Ceci est un sous-titre</h2>

La balise <h2> indique un titre de deuxième niveau auquel on a attribué l'identifiant titre_1.

A retenir :

- ✓ **Les pages Web sont codées dans le langage HTML.**
- ✓ **Le langage HTML est un langage à balises qui permet aux hyperliens de cibler des ressources (via les URL).**
- ✓ **Les pages HTML sont modifiables grâce à un éditeur et s'ouvrent grâce à un navigateur.**