

Chapitre1 : Introduction à HTML

I. Qu'est-ce que HTML ?

1. Historique :

HTML est une instance de **SGML (Standard Generalized Markup Language** / langage de balisage standard généralisé). Ce langage conçu dans les années 80, pour les grosses documentations techniques, était difficile d'apprentissage et complexe d'emploi. En particulier, il était inadapté à l'écriture de documents pour Internet. Une version allégée, le **HTML** a donc été développée ; soit un langage facile à apprendre et à utiliser.

2. Définition :

Le **HTML (HyperText Markup Language)** (langage de balisage hypertexte)) est né en 1989 sous l'impulsion de Tim Berners Lee (physicien et informaticien britannique), qui a révolutionné le monde en inventant le World Wide Web (un système de pages web publiques interconnectées à travers Internet pour le partage d'informations à l'échelle mondiale).

N.B : Le web et l'internet ne sont pas la même chose : le web est l'une des nombreuses applications bâties au-dessus de l'internet.

HTML est un **langage structuré** (de structuration) utilisé pour structurer une page web ainsi que son contenu. On peut par exemple organiser le contenu en un ensemble de paragraphes, tableaux, images, animation, liste d'éléments, formulaires, liens,...

HTML est un **langage de balisage** qui définit la structure du contenu à l'aide de repères que l'on appelle des **éléments** utilisés pour marquer différentes parties du contenu afin de les faire apparaître ou se comporter d'une certaine façon. Chaque élément est constitué de **balises** et d'**attributs** qui permettent d'apporter des informations sur son contenu.

Le langage **HTML** permet de créer des documents interactifs grâce à des **liens hypertextes**, qui relient votre document à d'autres documents. Ainsi, en cliquant sur une zone de texte (ou une image, un logo) mise en évidence, on peut accéder à un nouveau document situé sur un autre ordinateur en n'importe quel point du globe.

3. Le W3C :

Le **World Wide Web Consortium**, abrégé par le sigle **W3C** (<https://www.w3.org/>) est un organisme de standardisation international qui a été créé à l'occasion du premier standard du HTML : HTML 1.0. Il a pour rôle de promouvoir (càd, mettre à niveau) la compatibilité de diverses technologies telles que HTML, CSS, XML,...

C'est un consortium industriel piloté par certaines universités comme le Massachusetts Institute of Technology (MIT) aux États-Unis, par l'INRIA (Institut National de Recherche en Informatique et en Automatique) en France, l'Université de Keio au Japon, l'ECRIM (European Research Consortium for Informatics and Mathematics) mais également plus de 400 entreprises partenaires comme Mozilla...

4. CSS :

Le W3C a tenté avec succès de maintenir la séparation entre structure et présentation. Donc, il a adjoint au HTML la notion de **feuilles de style** à travers le langage **CSS (Cascading Style Sheets)** d'où HTML décrit la structure d'un document, CSS décrit sa présentation. Ainsi, un même document HTML peut être présenté de différentes façons à l'aide de différentes feuilles de style.

Les CSS forment un langage informatique qui décrit la présentation des documents HTML et XML. CSS répond bien à une volonté de « proposer une alternative à l'évolution du HTML d'un langage de structuration vers un langage de présentation ».

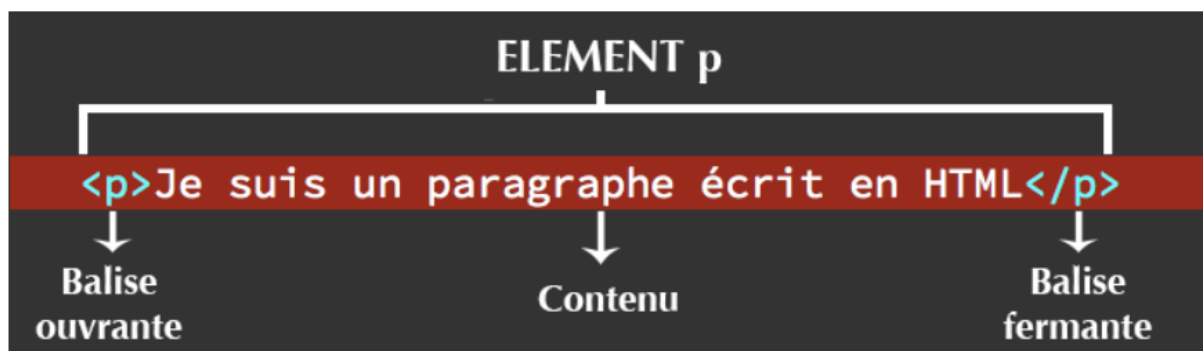
II. Structure et principes de base :

1. Les balises HTML :

Aujourd'hui, il existe plus de 120 éléments HTML différents aux rôles très variés et qui font la richesse de ce langage. Nous allons étudier et nous servir d'une grande partie d'entre eux dans ce cours.

Anatomie d'un élément HTML :

Voyons cet élément de paragraphe plus en détails :



Les parties principales de notre élément sont :

1. **La balise ouvrante** : celle-ci se compose du nom de l'élément en question (ici **p**), entre deux **chevrons** (le premier ouvrant **<** et le second fermant **>**). Cela indique le début de l'élément. Ici cela indique plus particulièrement où le paragraphe commence.
 2. **La balise fermante** : à la différence de la balise ouvrante, une *barre oblique (slash)* est placée avant le nom de l'élément. Cela indique la fin de l'élément. Dans notre exemple, c'est l'endroit où le paragraphe s'arrête.
 3. **Le contenu** : le contenu de l'élément. Pour cet exemple, il s'agit uniquement de texte.
 4. **L'élément** : l'ensemble formé par la balise ouvrante, la balise fermante et le contenu constitue l'élément.
- Un élément HTML peut être soit constitué d'une paire de balises (ouvrante et fermante) et d'un contenu, soit d'une balise unique qu'on dit alors « **orpheline** » ou « **auto-fermée** » ; c'est à dire qu'elle n'a qu'une seule balise ouvrante, et aucune balise fermante. Cela va être le cas pour certains éléments qui ne possèdent pas de contenu textuel comme l'élément **
** par exemple qui sert simplement à créer un retour à la ligne en HTML.
Les éléments n'ayant pas de contenu, on les qualifie **d'éléments vides**.

- Notez la présence du `<` avant le second chevron `>`. Les balises auto-fermantes en sont pourvues.
- Le navigateur client (Internet Explorer, Netscape, Mozilla, Firefox,...) ne fera pas apparaître **les balises (ou tags ou marqueurs)** mais interprètera leurs effets. L'interprétation de certaines balises peut varier d'un navigateur à l'autre.

2. Les attributs HTML :

Les balises peuvent posséder un ou plusieurs **attributs** qu'on va alors placer au sein de la balise ouvrante de ceux-ci. Pour certains éléments, les attributs vont être facultatifs tandis que pour d'autres ils vont être obligatoires pour garantir le bon fonctionnement du code HTML.

Les attributs vont en effet venir compléter les éléments en définissant plus précisément une balise, en spécifiant son action, en modifiant son comportement par défaut, ou bien en spécifiant des informations indispensables (comme l'adresse d'une image intégrée à la page, ou bien celle d'un lien externe) pour répondre aux critères souhaités par les utilisateurs.

Un attribut contient toujours une valeur, qu'on peut cependant parfois omettre dans le cas des attributs ne possédant qu'une seule valeur (car la valeur est alors considérée comme évidente). Certains attributs n'ont pas de valeur, comme **required**.

La syntaxe utilisée est la suivante :

```
<balise attribut1="valeur1" attribut2="valeur2" attribut3="valeur3"> Cours HTML</balise>
```

Les attributs qui ont une valeur observent ces règles :

1. Il y a un espace entre le nom de l'attribut et le nom de l'élément (ou le précédent attribut si l'élément a déjà plusieurs attributs).
2. Le nom de l'attribut est suivi par un signe égal (=).
3. La valeur de l'attribut est délimitée par des doubles (") ou simples (') quotes.
4. La séquence d'attributs n'a pas d'importance.
5. La paire attribut/valeur définit ce qu'on appelle **une propriété**.

3. Les commentaires :

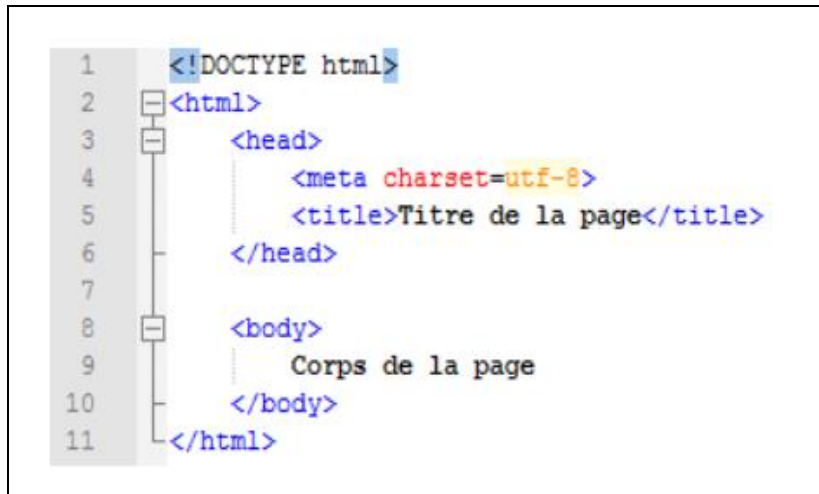
Les commentaires vont être des lignes de texte que l'on va écrire au milieu de notre code, afin de donner des indications sur ce que fait le code en question.

Quoiqu'ils alourdissent le temps de chargement d'une page, les commentaires semblent être nécessaires dans certains cas. En plus, ils n'apparaissent pas dans la fenêtre du navigateur.

La syntaxe est la suivante :

```
<!-- Voici un commentaire HTML -->
<!--
Voici un commentaire HTML
Qui peut se placer sur plusieurs lignes
-->
```

4. La structure de base d'un fichier html :



a. Définition de type de document :

La première balise de votre document doit être la balise de **DTD (Document Type Definition)** qui va servir uniquement au bon comportement du document. Comme son nom l'indique, le **doctype** sert à indiquer le type du document. Ainsi, dans la balise de l'élément doctype, on va préciser le langage utilisé, à savoir le HTML dans notre cas.

b. L'élément HTML :

Après la DTD, il suffit d'ouvrir l'élément qui englobe tout le reste de la page qu'on appelle aussi **l'élément racine**. Cet élément est composé d'une paire de balises ouvrante `<html>` et fermante `</html>`.

c. Les éléments Head et Body :

A l'intérieur de l'élément **html**, nous devons à nouveau indiquer deux parties fondamentales d'un document HTML qui sont une seule balise `<head>` et une seule balise `<body>`.

- L'élément **head** est un élément d'en-tête qui va servir de conteneur pour tout ce qu'on veut inclure dans une page HTML qui n'est *pas* du contenu à afficher à l'écran. Au sein de l'élément **head**, nous allons devoir à minima indiquer deux éléments :
 - ✓ L'élément **title** va nous permettre d'indiquer **le titre de la page** en soi. Ce titre apparaîtra dans l'onglet de votre navigateur sur lequel la page est chargée.
 - ✓ L'élément **meta** sert à transmettre des méta-informations sur la page à tous les outils susceptibles de les exploiter notamment les navigateurs, les moteurs de recherche et tous les outils qui analysent vos pages afin d'identifier diverses informations. Cet élément possède de nombreux attributs différents. Le type d'informations transmis va dépendre de l'attribut que l'on va préciser. Par exemple :
 - `<meta charset="utf-8">` : Cet élément définit le type d'encodage des caractères du document. Il s'agit d'un jeu de caractères universel qui inclut à peu près tous les caractères des langues humaines pour que le navigateur puisse afficher les caractères de texte correctement.
 - `<meta name="author" content="Mme. AMARA">`

- `<meta name="keywords" content="Introduction à HTML">`
- `<meta name="description" content="Cette page fournit une introduction au langage HTML">`
- `<meta name="date" content="2020-09-01">`

Il est utile d'inclure ces deux méta-éléments dans votre page pour définir son auteur et donner une courte description de son contenu.

- L'élément **body** contient tous les éléments définissant les contenus « visibles » de la page, c'est-à-dire les contenus à destination des utilisatrices et utilisateurs web lorsqu'ils visitent la page et notamment les différents textes présents dans la page, les images, des vidéos, des pistes audio, des jeux, etc.

5. La règle d'imbrication :

L'imbrication est l'une des fonctionnalités du HTML qui fait toute sa force. Cependant, on ne peut pas imbriquer des éléments HTML n'importe comment mais il faut suivre des règles précises.

Lorsque plusieurs balises sont imbriquées, une règle simple s'applique : une balise ouverte à l'intérieur d'une autre doit obligatoirement être refermée avant la fermeture de sa balise mère. Autrement dit, le premier élément déclaré doit toujours être le dernier refermé, tandis que le dernier ouvert doit toujours être le premier fermé.

Le schéma ci-dessous présente toutes les situations valides d'imbrication d'éléments :

```
<balise ouvrante élément A>
  <balise ouvrante élément B>
  </balise fermante élément B>

  <balise ouvrante élément C>
  </balise fermante élément C>

  <balise orpheline élément D>
</balise fermante élément A>
```

Bibliographie

<https://www.pierre-giraud.com/html-css-apprendre-coder-cours/structure-minimale-page/>

https://developer.mozilla.org/fr/docs/Learn/Getting_started_with_the_web/HTML_basics