STRUCTURE D'UN DOCUMENT HTML

1. Le langage HTML

L'HTML (Hypertext Markup Language) est le langage universel utilisé pour communiquer sur Internet.

Il a été inventé en 1990 par un scientifique nommé Tim Berners-Lee.

Ce n'est pas un langage de programmation mais un langage de mise en page. Il permet de présenter des informations sur Internet. Ce que vous voyez quand vous regardez une page sur internet est l'interprétation par votre navigateur (Internet Explorer, Firefox, Chrome, Safari...) du code HTML.

L' HTML est un langage utilisé pour construire des documents hypertextes pouvant être visualisé sur le web ou par n'importe quelle application capable d'interpréter l' HTML.

On définit la notion de documents hypertextes comme étant des documents ayant un lien entre eux.

Car sur le web, on navigue de page en page grâce à ces liens hypertextes.

L'HTML est définit comme un langage interprété. Un langage interprété n'est pas exécuté directement sur la machine mais par un autre programme appelé interprète. Les navigateurs sont ces interprètes.

1.1 Les éléments et les balises

Les éléments donnent la structure d'un document HTML. Ils indiquent comment vous voulez que votre navigateur présente votre site web. En genéral, les éléments se composent d'une balise ouvrante, d'un contenu et d'une balise fermante.

Les balises sont les étiquettes que vous utilisez pour marquer le début et la fin d'un élément.

Toutes les balises ont le même format , car elle commencent toujours par un signe inferieur « < » et finissent par un signe extérieur « > ».

Ex:

Voici un exemple avec un texte en gras

Voici un exemple avec un texte en gras

Seul le texte « avec un texte en gras » a subit une mise en forme car il est entouré d'une balise ouvrante et d'une balise fermante .

La lettre « b » est l'abréviation de « bold » (gras).

C'est pour cette raison que le langage HTML est qualifié de langage à balises puisque la mise en page du texte s'effectue sur le fait de placer des balises aux endroits du texte à mettre en page.

En somme, il suffit de prendre un texte brut qui que l'on balise pour mettre en forme. Ensuite le navigateur interprète ces balises afin de mettre en forme le texte. Ainsi le visiteur du site internet accède au rendu final du texte mis en page.

2. Structure du HTML

HTML est simple et logique. Le navigateur lit le code HTML comme vous, de haut en bas et de gauche à droite. Un document HTML commence donc par ce qui vient en premier dans la page et finit par ce qui vient en dernier.

Voici un exemple de page Web : « HelloNFA053 !! » .Ce fichier porte le nom hello_nfa053.html et contient le terxte suivant :

```
_<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"</pre>
1
2
         "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
3
         <html xmlns="http://www.w3.org/1999/xhtml" lang="fr">
4
5
    - <head>
6
          <meta http-equip="Content-Type" content="text/html; charset=utf-8"/</pre>
7
          <title>Hello NFA053!!</title>
8
9
     </head>
10
11
    =<body>
12
     Bonjour NFA053!!
13
     </body>
14
15
     </html>
16
```

Décrivons le contenu de cette page :

La première ligne correspond a la DTD (definition du Type de Document) appelé aussi DOCTYPE.

Contrairement à ce qu'on pourrait croire, elle ne sert pas aux navigateurs Web actuels à comprendre la *structure* d'une page (X)HTML : ceux-ci "lisent" ce code (X)HTML sans se servir de la DTD pour le déchiffrer, en se servant seulement des "règles" contenus dans leur propre "moteur".

En revanche, la DTD est indispensable aux validateurs (X)HTML comme celui du W3C, qui en ont besoin pour savoir à quelles règles le document est supposé obéir. Certes, il est possible de forcer la validation en l'absence de DTD, en utilisant l'interface étendue du validateur du W3C.

Mais un document même valide et dénué de DTD posera un problème de rendu dans les navigateurs modernes.

2.1 Le Document HTML minimum

Nous avons vu qu'une page HTML était en fait un fichier texte, enrichi d'un certain nombre de codes ou commandes, appelés balises.

Ces balises sont toujours exprimées sous la forme d'un mot clé, encadré par les caractères "<" et ">". Exemple : <BALISE>.

Pour la plupart des balises, il existe une balise de fermeture associée, reprenant le même nom, mais précédée du caractère "/". Exemple : </BALISE>. La commande spécifiée s'applique donc uniquement au texte situé entre le couple de balises ainsi formé. Exemple:

<HTML> ... </HTML>

Notons que :

- une balise peut indifféremment être indiquée en minuscules ou en majuscules,
- le formatage "manuel" du document (espaces, sauts de lignes,...) est toujours ignoré.
- Par exemple : <HTML>...</HTML> est interprété de la même façon par le navigateur web que la syntaxe sur plusieurs lignes indiquée ci-dessus.

Nous venons de découvrir, sans le savoir, un des éléments de la structure d'une page HTML : toute page doit en effet débuter par la balise <HTML> et se finir par </HTML>.

Entre ces deux balises, on définit deux zones : l'en-tête, spécifié par les commandes <HEAD> et </HEAD>, ainsi que le corps, délimité par: <BODY> et </BODY>. Ce qui donne, comme structure de base :

<HTML>

<HEAD>

...

</HEAD>

<BODY>

. . .

</BODY>

</HTML>

Dans l'en-tête, on ne met généralement qu'une seule information, le Titre du document qui sera affiché en haut de la fenêtre du navigateur et qui apparaît dans les bookmarks ou favoris (listes d'URL gérées par un navigateur, une sorte d'annuaire).

Ce Titre est indiqué entre les balises <TITLE> et </TITLE>. Exemple:

<TITLE>Ceci est le Titre</TITLE>.

Dans le corps, on met en fait tout le document à afficher (texte, définition des images etc.). Signalons l'existence d'une balise de commentaires, qui peut être utilisée partout dans les documents HTML, définie comme suit :

<!-- Ceci est un commentaire --!>.

Récapitulatif:

| <html></html> | Ceci est le début d'un document de type HTML |
|---|---|
| | Ceci est la fin d'un document de type HTML |
| <head></head> | Ceci est le début de la zone d'en-tête (prologue au document |
| | proprement dit contenant des informations destinées au browser) |
| | Ceci est la fin de la zone d'en-tête |
| <title></td><td>Ceci est le début du titre de la page</td></tr><tr><td></title> | Ceci est la fin du titre de la page |
| <body></body> | Ceci est le début du document proprement dit. |
| | Ceci est la fin du document proprement dit. |

3. DTD ET BALISES META

3.1 LES DTD

Différentes DTD existent et sont en vigueurs actuellement .Je vous conseille d'utiliser la DTD suivante qu'il faut recopier en premier lors de la création d'un nouveau document HTML :

```
<!DOCTYPE html>
```

Si vous souhaitez avoir plus d'information sur les différentes DTD, je vous invite à aller sur le site www.alsacreations.com, qui explique les différences entre chaque DTD :

http://www.alsacreations.com/article/lire/560-dtd-doctype-html-xhtml-comment-choisir.html

3.2 LES BALISES META

En dehors du titre de la page (<title>...</title>), les informations contenu dans l'entête (<head>...</head>) d'un document HTML (<html>...</html>) sont les balises <meta>.

Toutes les informations contenu dans cette entête ne seront pas affichées dans le navigateur car en de hors des informations qu'elles peuvent apporter au navigateur (titre de la page, type d'encodage utilisé :charset), ces « META-TAGS» servent principalement aux robots des moteurs de recherche.

Ces robots lise les informations contenues dans les « META-TAG » pour renseigner le moteur de recherche (Titre de la page, description du site, mots-clefs, auteur ..)

A partir de l'exemple précédent, nous allons l'enrichir de balises META afin d'y décrire son contenu auprès des robots des moteurs de recherche.

```
C!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
     "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
    <html xmlns="http://www.w3.org/1999/xhtml" lang="fr">
- <head>
     <meta http-equip="Content-Type" content="text/html; charset=utf-8"/>
     <meta http-equip="Content-Language" content="fr"/>
     <meta name="description" content="Description de votre site Web">
     <meta name="keywords" content="les,mots,clefs,utlises,pour,la,recherche,internet">
     <meta name="authots" content="nom de l'auteur">
     <meta name="robots" content="all">
     <title>Hello NFA053!!</title>
 </head>
-<body>
 Bonjour NFA053!!
 </body>
  </html>
```

<meta http-equiv="content-type" content="text/html; charset=utf-8"/</pre>

Cette balise indique au moteur de recherche mais aussi au navigateur le codage utilisé pour ce fichier. Le codage UTF8 est recommandé pour son universalité. Il permet notamment de limiter le problème des caractères spéciaux tel que les accents qui doiventêtre codés si on utilise la table *iso-latin1* par exemple.

<meta http-equiv="content-language" content="fr">

Cette balise permet de définir la langue de cette page au moteur de recherche. lci nous avons affaire à une page « fr » donc écrite en langue française.

<meta name="description" content="Description de ma premiere page Web">

Cette balise permet d'indiquer la description de cette page que l'on désire voir apparaître dans le moteur de recherche.

<meta name="keywords" content="Mes, mots, clef, separés, par, des, virgules">

Cette balise indique les mots clefs par ordre d'importance que le moteur de recherche doit mettre en relation avec le contenu de notre page HTML.

<meta name="author" content="nom de l'auteur">

Cette balise vous permet de préciser le nom de l'auteur.

<meta name="robots" content="instructions pour les robots">

Cette balise vous permet d'indiquer aux moteurs de recherche si vous souhaitez que votre site soit ou non, indexé par eux.

Les instructions possibles sont les suivantes :

All (valeur par défaut) : le site sera entièrement référencé (pages et liens hyper textes).

None : vos pages ne sont pas indexées.

Index : seule la page qui définit ce meta-tag sera indexée.

Nolndex : la page n'est pas indexée, mais les liens définis pourront éventuellement être indexés.

Follow: autorise les robots à suivre les liens hyper textes définis.

NoFollow: pour le contraire.

<title>Hello NFA053 !!</title>

Cette balise permet d'indiquer le titre de cette page que l'on désire voir apparaître dans le moteur de recherche.

Même si il est important de renseigner les balises META de vos pages, les moteurs de recherche

devenant de plus en plus « *intelligents* », ces derniers ont de moins en moins besoins de ces informations meta que vous renseignez vous-même, rendant certaines balises META complètement obsolètes (ex : <meta name="keywords" content="...">).

4. Le traitement du texte

4.1 les Titres

En HTML vous avez la possibilité de faire des titres grace à la balise <Hn>.La lettre « n » représentant la taille allant de 1 (pour le plus grand) à 6 (pour le plus petit).

```
1 <h1>ceci estr une balise h1, la taille la plus grande</h1>
```

Voici le résultat :

ceci est une balise h1, la taille la plus grande

Exemple:

```
-<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"</pre>
 2
         "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
         <html xmlns="http://www.w3.org/1999/xhtml" lang="fr">
 3
 4
 5
    - <head>
 6
 7
          <title>exo titre</title>
 8
 9
     </head>
10
11
    E<body>
12
13
14
      <h1>ceci est une balise h1, la taille la plus grande</h1>
15
      <h2>ceci est une balise h2</h2>
16
      <h3>ceci est une balise h3</h3>
      <h4>ceci est une balise h4</h4>
17
18
      <h5>ceci est une balise h5</h5>
19
      <h6>ceci est une balise h6</h6>
20
21
      </body>
      </html>
22
```

ceci est une balise h1, la taille la plus grande

ceci est une balise h2

ceci est une balise h3

ceci est une balise h4

ceci est une balise h5

ceci est une balise h6

4.2 Les paragraphes

Le paragraphe se détermine grâce à la balise

```
ceci est un paragraphe
```

Il est tout a fait possible d'aligner le texte de manière justifié, à gauche, au centre ou à droite grâce à la propriété *align*.

C'est ce qui est fait dans l'exemple suivant :

4.3 Les listes

Le langage HTML permet de définir deux styles de listes : Les listes numérotées et les listes nonnumérotées.

Les listes numérotées

Une liste numérotée est délimitée par les balise et entre lesquelles chaque élément de la liste est précédé de la balise

Les Listes à puces

Une liste non numérotée est définie de la même façon, sauf qu'elle est encadrée par les balises et .

Exemple:

```
C!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
    <html xmlns="http://www.w3.org/1999/xhtml" lang="fr">
- <head>
     <title>exo liste</title>
 </head>
=<body>
户<0L>
-<LI>élément 1
⊟<LI>élément 2
⊟<LI>élément 3
 </OL>
□ <UL>
= <LI>élément 1
-<LI>élément 2
</UL>
 </body>
</html>
```

Voici le résultat obtenu :

- 1. élément 1
- 2. élément 2
- 3. élément 3
- élément 1
- élément 2
- élément 3

Il est possible aussi d'imbriquer des listes du même type ou de type différents.

Exemple:

```
<UL>
<LI>élément 1:
<OL>
<LI>sous-élément 1
<LI>sous-élément 2
<LI>sous-élément 3
</OL>
<LI>élément 2
<LI>élément 2
<LI>élément 3
</UL>
```

On obtient alors:

- élément 1:
 - 1. sous-élément 1
 - 2. sous-élément 2
 - 3. sous-élément 3
- élément 2
- élément 3

Les listes de tabulations

Il est possible de définir des listes spéciales, permettant de faire des glossaires, en utilisant les balises <DL> et </DL> pour encadrer la liste, <DT> pour spécifier un élément et <DD> pour indiquer sa définition.

Exemple:

4.4 FILET HORIZONTAL

Il existe le filet horizontal (<HR/>), très pratique pour séparer les chapitres, cela donne plus de clarté au document HTML. Il est très utilisé aussi pour marquer la fin du document. Dans ce cas on trouve en général des informations sur le copyright.

```
<body>
Maître corbeau, sur un arbre perché, Tenait en son bec un fromage.
Maître renard par l'odeur alléché , Lui tint à peu près ce langage :
«Et bonjour Monsieur du Corbeau.
Que vous êtes joli! que vous me semblez beau!
Sans mentir, si votre ramage
Se rapporte à votre plumage,
Vous êtes le phénix des hôtes de ces bois»
A ces mots le corbeau ne se sent pas de joie;
Et pour montrer sa belle voix,
Il ouvre un large bec laisse tomber sa proie.
Le renard s'en saisit et dit: "Mon bon Monsieur,
Apprenez que tout flatteur
Vit aux dépens de celui qui l'écoute:
Cette leçon vaut bien un fromage sans doute."
Le corbeau honteux et confus
Jura mais un peu tard , qu'on ne l'y prendrait plus.
Copyright
</body>
</html>
```

Ce qui donne:

Maître corbeau, sur un arbre perché, Tenait en son bec un fromage. Maître renard par l'odeur alléché, Lui tint à peu près ce langage: «Et bonjour Monsieur du Corbeau. Que vous êtes joli! que vous me semblez beau! Sans mentir, si votre ramage Se rapporte à votre plumage, Vous êtes le phénix des hôtes de ces bois» A ces mots le corbeau ne se sent pas de joie; Et pour montrer sa belle voix, Il ouvre un large bec laisse tomber sa proie. Le renard s'en saisit et dit: "Mon bon Monsieur, Apprenez que tout flatteur Vit aux dépens de celui qui l'écoute: Cette leçon vaut bien un fromage sans doute." Le corbeau honteux et confus Jura mais un peu tard, qu'on ne l'y prendrait plus.

Copyright

On remarque que la balise <hr/> possède un « / » dans son écriture.Le fait d'ecrire <hr/> est un raccourci d'éciture des balises ouvrantes et fermantes <hr></hr>.

Comme nous n'avons aucun texte à mettre entre la balise ouvrante et fermante alors on peut utiliser le raccourci <hr/>qui indique que cette balise est à la fois ouvrante et fermante.

4.5 Saut de ligne

La balise
br> force le passage d'une ligne. Même si les paragraphes passent des lignes automatiquement, on peut avoir envie d'espacer un peu plus nos paragraphes.

Exemple:

```
Maître corbeau, sur un arbre perché, Tenait en son bec un fromage.
Maître renard par l'odeur alléché , Lui tint à peu près ce langage : <br/> <br/>
«Et bonjour Monsieur du Corbeau.
Que vous êtes joli! que vous me semblez beau!<br/>
Sans mentir, si votre ramage
Se rapporte à votre plumage, <br/>
Vous êtes le phénix des hôtes de ces bois><br/>
A ces mots le corbeau ne se sent pas de joie; <br/>
Et pour montrer sa belle voix,
Il ouvre un large bec laisse tomber sa proie. <br/>
Le renard s'en saisit et dit: <br/> "Mon bon Monsieur,
 Apprenez que tout flatteur
Vit aux dépens de celui qui l'écoute: <br/>
Cette leçon vaut bien un fromage sans doute. "<br/>
Le corbeau honteux et confus
Jura mais un peu tard , qu'on ne l'y prendrait plus.
Copyright
</body>
</html>
Résultat :
Maître corbeau, sur un arbre perché, Tenait en son bec un fromage.
Maître renard par l'odeur alléché , Lui tint à peu près ce langage : <br/> <br/>
«Et bonjour Monsieur du Corbeau.
Que vous êtes joli! que vous me semblez beau!<br/>
Sans mentir, si votre ramage
Se rapporte à votre plumage, <br/>
Vous êtes le phénix des hôtes de ces bois><br/>
A ces mots le corbeau ne se sent pas de joie; <br/>
Et pour montrer sa belle voix,
Il ouvre un large bec laisse tomber sa proie. <br/>
Le renard s'en saisit et dit: <br/> "Mon bon Monsieur,
 Apprenez que tout flatteur
Vit aux dépens de celui qui l'écoute: <br/>
Cette leçon vaut bien un fromage sans doute. "<br/>
Le corbeau honteux et confus
Jura mais un peu tard , qu'on ne l'y prendrait plus.
Copyright
</body>
</html>
```

4.6 Mise en forme du Texte et Police

Mise en forme du texte

Il est possible de mettre en gras, souligné, en italique ou encore mettre en télétype un texte.

```
<b>texte en gras</b>
<br/>
<br/>
<i>texte en italique</i>
<br/>
<br/>
<u>texte souligné</u>
<br/>
<br/>
<tt>texte télétype</tt>
<br/>
```

Résultat :

texte en gras

texte en italique texte souligné texte télétype

Combinaison de texte en gras et souligné

Pensez bien à mettre les balises dans le bon ordre.

Ecriture correcte

```
<b><u>Combinaison de texte en gras et souligné</u>
```

Ecriture incorrecte

<u>Combinaison de texte en gras et souligné</u>

4.7 Mise en forme du paragraphe

Il existe d'autres balises de mise en forme des paragraphes ou d'une partie du paragraphe. Certaines génèrent un passage à la ligne automatique, et d'autre non .

```
<h3><u>Le corbeau et le renard</u></h3>
<cite>balise cite citation Le Corbeau et le Renard </cite>
<blockquote>balise blockquote citation de jean de La Fontaine</blockquote>
<br/>
<br/>
<br/>
<br/>
<em>balise em mise en italique</em><br/>
<tt>balise tt style machine a ecrire</tt><br/>
<strong>balise strong mise en valeur (traduit en gras)</strong><br/>
<sup>balise sup place un texte en exposant</sup><br/>
<sub>balise sub place un texte en indice</sub><br/>
<sub>balise sub place un texte en indice</sub><br/>
<sub>balise</pr>
```

Résultat obtenu :

Le corbeau et le renard

balise cite citation Le Corbeau et le Renard

balise blockquote citation de jean de La Fontaine

balise em mise en italique
balise tt style machine a ecrire
balise strong mise en valeur (traduit en gras)
balise sup place un texte en exposant
balise sub place un texte en indice

Maître corbeau, sur un arbre perché, Tenait en son bec un fromage.

4.8 Les attributs de mise en page

La technique est simple. Il suffit d'ajouter un attribut à la balise et ainsi spécialiser la commande.

<balise attribut= "valeur de l'attribut ">

Nous avons vu précédemment qu'on pouvait aligner les paragraphes de la sorte en utilisant l'attribut « align » :

4.9 Choix et Taille de la police

On peut modifier la taille et le type de police comme cet exemple :

```
<body>
<h3><u>Utilisation de la Police</u></H3>
<font size="1">Taille</font>
<font size="2">de</font>
<font size="3">police</font>
<font size="4">variant de 1 à 7</font>
<font size="5">de plus </font>
<font size="6">en plus</font>
<font size="7">grande</font>
<br/>
<hr/>
<h3><u>Utilisation de la taille relative</u></h3>
<font>la taille de la police est standard</font>
<font size="-1">police dont la taille est inférieure de 1</font>
<font size="+1">police dont la taille est supérieure de 1</font>
<br/>
<hr/>
<h3><u>Exemples de Polices</u></h3>
<font face="arial">Police Arial</font>
<br/>
<font face="Times">Police Times</font>
<font face="courrier">Police Courrier</font>
<font face="verdana">Police Verdana</font>
<br/>>
\langle hr/ \rangle
<font face="courrier"size ="1">Police Courrier de taille 1</font>
</body>
</html>
```

Voici le résultat :

Utilisation de la Police

Taille de police variant de 1 à 7 de plus en plus grande

Utilisation de la taille relative

la taille de la police est standard police dont la taille est inférieure de 1 police dont la taille est supérieure de 1

Exemples de Polices

Police Arial Police Times Police Courrier Police Verdana

Police Courrier de taille 1

5. LES COMMENTAIRES

On peut ajouter des commentaires à n'importe quel endroit d'un code HTML.Ce commentaire vous aidera a vous repérer dans votre fichier HTML.

Voici un exemple : < !--ceci est un commentaire-->

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"
   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
        <!--debut du fichier HTML-->
        <html xmlns="http://www.w3.org/1999/xhtml" lang="fr">
        <!--debut de l'entete-->
        <head>
        <title>Hello NFA053!!</title>

</head>
</!--fin de l'entete-->
<!--debut du corps du document-->
<body>
Bonjour NFA053!!
</body>
<!--fin du corps du document-->
</html>
```

Il est vivement conseillé d'utiliser les commentaires. Ils vous servent à vous repérer dans votre code lors de son écriture mais aussi lors de sa reprise plusieurs temps après ou encore par un autre webmaster qui serait amené à utiliser votre code.

5.1 VALIDATION DU CODE PAR LA W3C

Il est important une fois votre page HTML terminée, de vérifier si son code ne comporte pas d'erreur. Une erreur n'est pas forcément visible dans le rendu du navigateur. De plus cette erreur est plus ou moins lisible sur tel ou tel navigateur.

L'oublie d'une fermeture de balise, une erreur dans un attribut, etc.. Ces exemples d'erreur sont très courantes et donc il est conseillé de vérifier son code sur le site de la W3C.

Le **World Wide Web Consortium**, abrégé par le sigle **W3C**, est un organisme de standardisation à but non-lucratif, fondé en octobre 1994 comme un consortium chargé de promouvoir la compatibilité des technologies du World Wide Web telles que HTML, XHTML, XML, RDF, CSS, PNG, SVG et SOAP.

Le W3C n'émet pas des normes au sens Européen, mais des recommandations à valeur de standards industriels. C'est à partir de ces recommandations que les navigateurs sont développés par les éditeurs. Ainsi plus vous respectez les recommandations de la W3C, plus vos pages auront des chances d'être compatibles avec un maximum de navigateurs.

http://validator.w3.org

Il suffit de valider vos fichiers en ligne en uploadant votre fichier directement de votre ordinateur.



This validator checks the <u>markup validity</u> of Web documents in HTML, XHTML, SMIL, MathML, etc. If you wish to validate specific content such as <u>RSS/Atom feeds</u>

Dès que vous avez cliqué sur le bouton check, le site de la W3C va vous indiquer si votre fichier est conforme ou non.

Si il ne l'ai pas il vous indiquera les erreurs trouvés dans votre page web.