



Département
Éducation
et Technologie

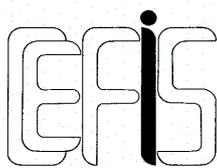
- Hypertexte
- Page Web
- Document HTML

Publier sur le Web
HTML, vous avez dit HTML ?

Guy Vastersavendts

5.60

Novembre 1999



Centre pour la
Formation à
l'Informatique dans
le Secondaire



HTML ... vous avez dit HTML ?

Ces notes étant rédigées pour une formation intitulée "Publier sur le Web", commençons donc par regarder une page Web. Prenons-en une (presque) au hasard, sur le site du CeFIS par exemple.

<http://www.det.fundp.ac.be/cefis/centdoc.html>



Centre de documentation



🔍 Recherche dans le centre de documentation (provisoirement inaccessible)

Le CeFIS gère un centre de documentation qui regroupe, pour les mettre à la disposition de toute personne intéressée, et plus particulièrement des enseignants, plus de 1300 ouvrages, manuels, articles, logiciels, ainsi qu'une quinzaine de revues, concernant informatique et enseignement. Plus largement, le département Éducation et Technologie dispose d'un fond de plus de 3000 documents ayant trait aux sciences de l'éducation. Toutes ces ressources peuvent être consultées ou empruntées.

Parallèlement, les enseignants du CeFIS, comme tous ceux du département Éducation et Technologie, se tiennent à la disposition des enseignants pour répondre aux questions techniques ou pédagogiques liées à l'enseignement de l'informatique ou à l'utilisation des environnements informatisés au sein des écoles.

Enfin, nous avons commencé la mise en place d'un serveur à vocation pédagogique. L'idée est de mettre à disposition des enseignants des matériaux utilisables développés par ailleurs et de faciliter des contacts et des échanges. Nous souhaitons faire de ce serveur non une vitrine ou un recensement de ressources présentes ailleurs, mais une source de documents et d'outils réellement intégrée dans les formations dispensées, qu'elles s'organisent au CeFIS ou au sein des écoles.

Aller à : [FUNDP](#) / [Département Education et Technologie](#) / [CeFIS](#)

Cette page a un aspect assez classique. Il y a des images, un titre, du texte et des liens.

La question suivante vous a déjà sûrement traversé l'esprit : qu'est-ce qu'il y a derrière cette page ? Pour le savoir, il suffit dans le navigateur de demander l'affichage de la source de cette page. Cette commande nous permet de voir l'envers du décor, ce qu'il y a derrière cette page. Voici ce qu'on y trouve dans notre exemple :

```

<html>
<head>
<title>CeFIS Centre de documentation</title>
<meta name="GENERATOR" content="Microsoft FrontPage 3.0">
</head>
<body bgcolor="#FFFFFF">
<table width="761">
<tr>
<td align="center" width="192"><imgsrc="http://www.fundp.ac.be/dessins/sigle.gif"
align="bottom" width="136" height="112"></td>
<td align="center" width="349"><h6><font size="6" color="#000000">Centre de
documentation</font>
</h6>
</td>
<td align="center" width="208"><br>
<font size="1"><tt><font color="#0084D6"><strong>CE</strong></font>nre pour
la<strong><font color="#0084D6">F</font></strong>ormation<br>
à l'<font color="#0084D6"><strong>I</strong></font>nformatique<br>
dans le <font color="#0084D6"><strong>S</strong></font>econdaire</tt></font></td>
</tr>
<tr>
<td colspan="3" width="757"><hr>
</td>
</tr>
</table>
<table border="0" width="100%" cellspacing="0" cellpadding="0">
<tr>
<td width="15%"></td>
<td width="67%"><u><b><p ALIGN="JUSTIFY"></b></u><font size="2"> Recherche
dans le centre de documentation</font>
<font color="#FF0000">(provisoirement inaccessible)</font></p>
<p ALIGN="JUSTIFY">Le CeFIS gère un centre de documentation qui regroupe, pour
les mettre à la disposition de toute personne intéressée, et plus particulièrement des
enseignants, plus de 1300 ouvrages, manuels, articles, logiciels, ainsi qu'une quinzaine de
revues, concernant informatique et enseignement. Plus largement, le département Éducation
et Technologie dispose d'un fond de plus de 3000 documents ayant trait aux sciences de
l'éducation. Toutes ces ressources peuvent être consultées ou empruntées.</p>
<p ALIGN="JUSTIFY">Parallèlement, les enseignants du CeFIS, comme tous ceux du
département Éducation et Technologie, se tiennent à la disposition des enseignants pour
répondre aux questions techniques ou pédagogiques liées à l'enseignement de l'informatique
ou à l'utilisation des environnements informatisés au sein des écoles.</p>
<p ALIGN="JUSTIFY">Enfin, nous avons commencé la mise en place d'un serveur à
vocation pédagogique. L'idée est de mettre à disposition des enseignants des matériaux
utilisables développés par ailleurs et de faciliter des contacts et des échanges. Nous
souhaitons faire de ce serveur non une vitrine ou un recensement de ressources présentes
ailleurs, mais une source de documents et d'outils réellement intégrée dans les formations
dispensées, qu'elles s'organisent au CeFIS ou au sein des écoles.</td>
</tr>
</table>
<hr>
<p>Aller à : <a href="http://www.fundp.ac.be/fundp.html">FUNDP</a> / <a
href="http://www.det.fundp.ac.be">Département Education et Technologie</a> / <a
href="index.html">CeFIS</a></p>
</body>
</html>

```

Effrayant ? Oui et non. Vu comme ça d'une manière brutale, cela peut décourager. Mais, nous allons découvrir petit à petit les clés de lecture (et donc d'écriture) de ce genre de document et nous rendre compte finalement que c'est moins compliqué qu'il n'y paraît au premier abord.

HTML signifie HyperText Markup Language ou en français langage de marquage hypertexte.

SGML signifie Standard Generalized Markup Language.

Il s'agit d'une page écrite en HTML, le formatage du document se fait à l'aide de codes de marquage appelés balises (tag en anglais). On les reconnaît dans le texte car elles sont entourées des signes < et >. Le HTML est un langage puissant dérivé du SGML, utilisé depuis longtemps dans les logiciels de publication assistée par ordinateur (PAO). Le langage HTML a été défini expressément pour être utilisé sur le Web et sa définition évolue continuellement.

Les pages écrites en HTML sont décodées et affichées par des logiciels appelés navigateurs (browsers en anglais). Le navigateur reçoit le texte, comme celui qui précède, et affiche la page telle qu'on peut la voir en haut. Le navigateur est donc un logiciel qui transforme votre machine en un exécutant afficheur de pages HTML. Le navigateur "lit" chaque ligne du texte et "interprète" chaque balise. A chaque balise HTML rencontrée, le navigateur exécute la primitive (fonction) correspondante, par exemple la balise de mise en gras provoquera l'affichage en gras du texte balisé. Une des fonctions globales des navigateurs est donc d'interpréter et d'appliquer les balises HTML. Pour rappel, l'autre fonction principale d'un navigateur est d'envoyer des requêtes aux serveurs.

Les pages en HTML peuvent être écrites à partir de logiciels simples, appelés éditeurs de texte, qui produisent du texte brut sans aucun formatage, comme par exemple le "bloc-notes". Il existe aussi des logiciels spécialisés, appelés éditeurs HTML et destinés à faciliter l'écriture de page HTML. Nous aurons l'occasion d'en parler dans la suite.

Soyons clair ! Notre propos n'est certainement pas de décortiquer le langage HTML de A à Z, ni d'en analyser toutes les finesses, mais bien d'en dégager les caractéristiques et les éléments principaux. Ceci dans un double but, d'une part mieux comprendre la nature et le fonctionnement de pages Web, d'autre part mieux comprendre (et donc mieux maîtriser) ce qui se passe quand on produit une page HTML avec d'autres logiciels qu'un simple éditeur de texte. Ou en d'autres mots, faire de vous des créateurs intelligents de pages Web.

Définition de la structure d'un document

Le "/" (slash) indique toujours la fin d'une balise.

Tout document HTML commence par la balise `<HTML>`, qui indique au navigateur qu'il s'agit d'un document au format HTML, et se termine par `</HTML>`. Chaque document HTML doit aussi comporter un entête marqué par `<HEAD>` `</HEAD>`. Il sert à préciser certaines caractéristiques du document, comme par exemple le titre. Après l'entête se trouve le corps du document marqué par `<BODY>` `</BODY>`. Le corps contient les textes et les objets qui doivent être affichés par le navigateur. Même si certains navigateurs permettent qu'on déroge à cette règle, il vaut mieux la respecter.

Le document HTML minimum est donc :

```
<HTML>
<HEAD>
</HEAD>
<BODY>
</BODY>
</HTML>
```

Ce document n'affiche strictement rien, mais est syntaxiquement correct.

Entête HTML

L'élément le plus important de l'entête est le titre, balisé par `<TITLE>` `</TITLE>`. Le titre du document joue un rôle important pour celui qui consulte le document, il apparaît d'ailleurs dans le navigateur et dans les résultats affichés par les moteurs de recherche. Nous avons donc intérêt à toujours mettre un titre significatif à un document.

Texte HTML	Page Web
<pre><HTML> <HEAD> <TITLE>Formation CeFIS : publier sur le Web</TITLE> </HEAD> <BODY> </BODY> </HTML></pre>	On ne voit toujours rien sur cette page.

Il existe d'autres balises utilisables dans l'entête, mais qui nous intéressent moins dans le cadre de cette introduction.

Corps HTML

Le corps constitue la partie principale d'un document HTML, toutes les informations à afficher apparaissent dans cette partie. Le corps contient des titres, des paragraphes, des commentaires, des liens hypertextes, des références, des images, ...

Titres

La balise H signifie Heading

Les titres permettent d'organiser les idées dans un texte. HTML fournit 6 niveaux de titres, balisés de la manière suivante :

`<H1></H1>`

...

`<H6></H6>`

Chaque niveau de titre possède ses propres caractéristiques : police, taille, ...

Par exemple,

`<H1>Ceci est un titre de niveau H1</H1>`

donnera à l'écran :

Ceci est un titre de niveau H1

Contrairement à ce qui se passe avec un traitement de texte, l'auteur d'un document HTML n'est pas totalement maître de l'apparence de son document dans un navigateur. En effet, cette apparence, police et taille des caractères, dépend des paramètres de configuration du navigateur.

Paragraphes

La balise P signifie Paragraph.

Un document HTML peut être découpé en paragraphes comme dans un traitement de texte. Cela se fait avec la balise `<P>` `</P>`. Lorsque le navigateur lit la balise de paragraphe, il saute une ligne avant d'insérer le texte ou l'objet qui suivent cette balise.

Voici un exemple simple de l'utilisation de cette balise :

`<P>Nous commençons ici un nouveau paragraphe.</P>`

La balise `<P>` supporte plusieurs attributs, dont l'attribut `ALIGN` qui peut avoir comme valeur `LEFT`, `CENTER`, `RIGHT`, `JUSTIFY`.

`<P ALIGN=CENTER>Je suis placé au milieu</P>`

Cette balise place le texte au milieu (horizontalement) de l'écran ou pour être plus précis, lorsque le navigateur lit cette balise, il place le texte au milieu.

Commentaires

Dans un document HTML, les commentaires sont insérés, à n'importe quel endroit, de la manière suivante :

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

Les commentaires peuvent être utilisés pour indiquer la date de mise à jour du document, pour insérer des remarques dans le texte, ...

Sauts de ligne

La balise BR signifie Break.

Le texte placé entre 2 balises de paragraphe `<P>` `</P>` est formaté automatiquement par le navigateur. Les retours à la ligne dépendent de la définition de l'écran et de la taille de la fenêtre du navigateur. L'auteur peut imposer un saut de ligne pour créer une certaine mise en page, en utilisant la balise `
`, qui est une balise unique.

Lignes horizontales

La balise HR signifie Horizontal Rule.

Une façon simple d'augmenter l'impact visuel du document consiste à insérer une ligne horizontale sur toute la largeur du document. L'insertion de cette ligne horizontale s'obtient avec la balise `<HR>`.

Styles physiques

La balise B signifie Bold, I signifie Italic, U signifie Underline et TT signifie Teletype.

Les styles physiques sont utilisés pour indiquer au navigateur le format précis à afficher. Il s'agit essentiellement de format de caractères, comme par exemple :

```
<B>Gras</B>
<I>Italique</I>
<U>Souligné</U>
<TT>Police à espacement fixe</TT>
```

Le terme style physique est employé ici par opposition à style logique. Les styles physiques donnent à l'auteur la possibilité de fixer lui-même certaines caractéristiques à son texte. Lorsque l'auteur utilise des styles logiques, c'est le navigateur qui applique ces styles en fonction des paramètres de sa configuration.

Images

Le document en cours de rédaction peut être illustré (agrémenté) d'images, pour cela il faut utiliser la balise ``. A l'intérieur de cette balise, il faut préciser la source de l'image sous forme d'URL. Par exemple :

```
<IMG SRC="http://www.det.fundp.ac.be/cefis/cefis100.gif">
```

ou pour centrer l'image

```
<P ALIGN=CENTER>
```

```
<IMG SRC="http://www.det.fundp.ac.be/cefis/cefis100.gif"></P>
```

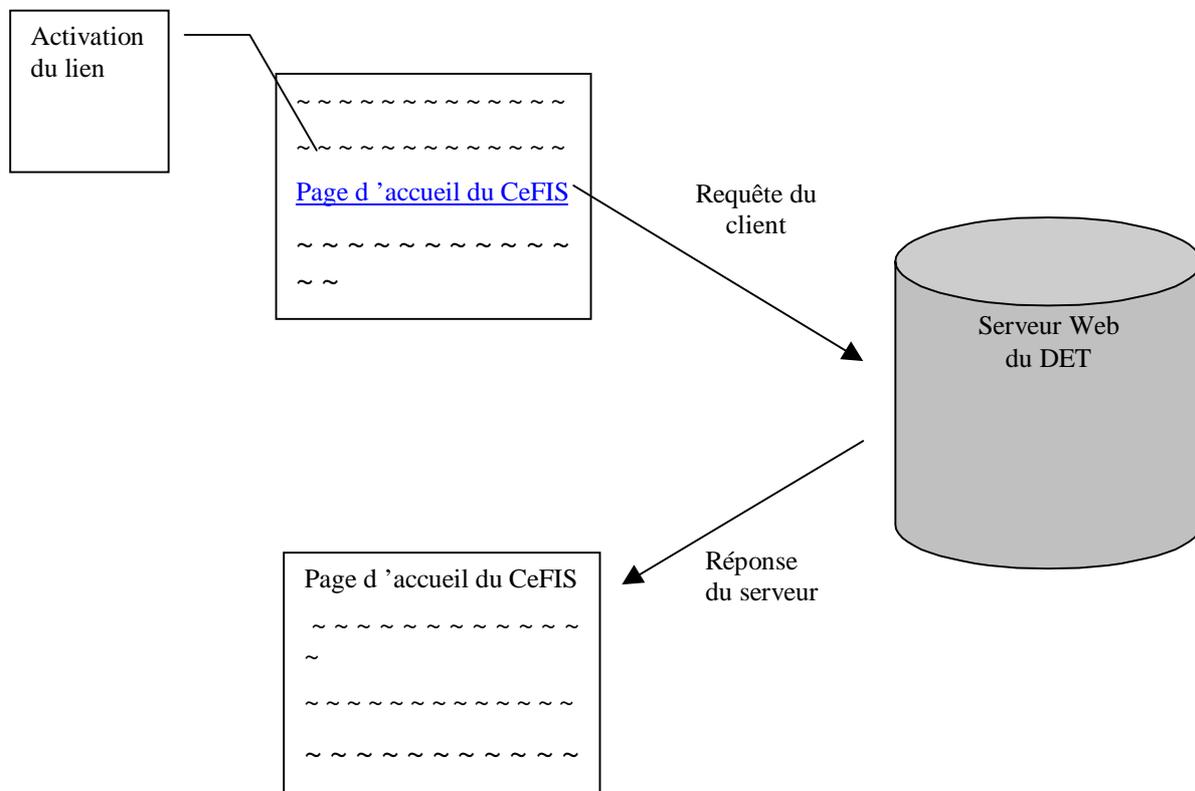
Ces images peuvent être animées. Il s'agit d'un ensemble d'images rassemblées dans un même fichier et dont la succession (animation) est gérée par le navigateur.



Ergonomie : sauf pour des raisons très particulières, évitez d'utiliser des images animées. En effet, les animations ont tendance à attirer l'attention du lecteur au détriment du contenu.

Liens

Toute la puissance et aussi tout l'intérêt des pages Web résident dans les liens qu'elles contiennent. Les liens donnent la possibilité de lier des textes, des images graphiques, des objets multimédia aux documents. Les textes, les images et les objets peuvent se trouver sur n'importe quel site du Web. On arrive ainsi à la métaphore de cette toile d'araignée mondiale (World Wide Web), qui permet des parcours virtuels dans des documents tout aussi virtuels.



C'est la balise `<A>` `` qui permet de faire ces liens. En voici un exemple :

```
<A HREF="http://www.det.fundp.ac.be/cefis/index.html">Page d'accueil
du CeFIS</A>
```

La référence du document ou de l'objet est donnée sous forme d'URL.

Lorsque le navigateur rencontre cette balise, il exécute la requête correspondant à HREF. Dans ce cas-ci, la requête s'adresse au serveur HTTP (Web) du Département Education et Technologie ; la demande concerne plus précisément un document appelé "index.html" dans le répertoire "cefis" du site. L'expression "Page d'accueil du CeFIS" apparaîtra sous forme de lien hypertexte dans le document. Lorsque le lien est activé, cela déclenche un mécanisme de requête client-serveur, comme nous l'avons déjà vu dans la partie "Utiliser Internet et ses services".

Dans les URL, les références peuvent être précisées :

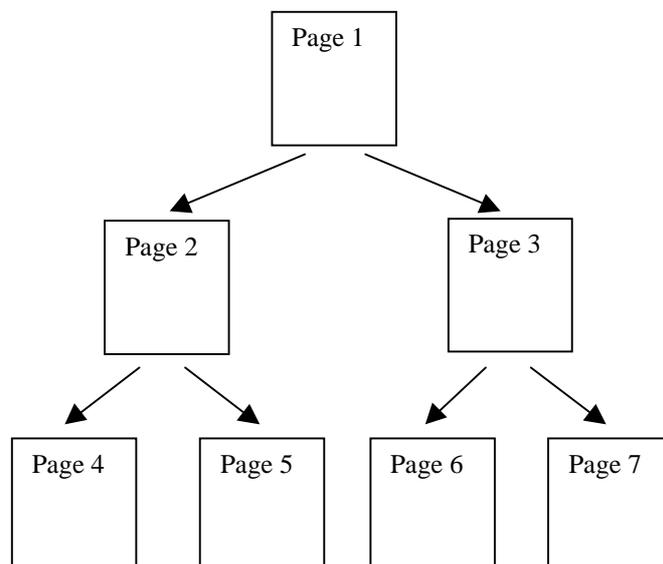
soit d'une manière absolue, quand on note le chemin d'accès complet à la ressource,

exemple : `http://www.det.fundp.ac.be/cefis/index.html`, le chemin d'accès complet comprend le nom du protocole (`http`), le nom du serveur Web (`www`), le domaine dans lequel se trouve ce serveur (`det.fundp.ac.be`), le chemin d'accès sur cette machine (répertoire `cefis`) et le nom du document (`formation.html`).

soit d'une manière relative, quand le chemin est noté par rapport au répertoire en cours,

exemple : `presentation.html`, par rapport à l'exemple précédent, on fait référence à un document `presentation.html` qui se trouve dans le répertoire `cefis`.

Ces liens sont le fondement même d'un site Web. L'auteur de pages Web doit non seulement gérer le contenu de ses pages, mais aussi les liens qui les unissent. C'est ce qu'on appelle l'arborescence du site.



URL signifie Uniform Resource Locator, c'est une notation standardisée pour désigner une ressource sur Internet.

Un site Web est un ensemble de page Web liées logiquement par l'auteur du site.



Remarque : on notera ici une grande différence entre une publication électronique hypertexte et une publication papier. Dans une publication papier, les pages s'enchaînent les unes aux autres de manière linéaire. Le lecteur va donc parcourir (habituellement) la publication de manière séquentielle. Dans une publication électronique hypertexte, l'auteur prévoit des liens entre les pages, mais il ne peut pas prévoir comment le lecteur fera son parcours. Il y aura à peu près autant de parcours que de lecteurs.

Les liens dont nous avons parlé jusqu'ici sont des liens externes, c'est-à-dire qui font référence à un autre document que celui qui est affiché. Cependant, le lien peut aussi porter sur un endroit précis (marque) du même document. Cette marque correspond à une espèce de signet, qui peut être utilisé pour atteindre cet endroit précis dans un document ou quand on veut noter un déplacement à l'intérieur même du document. On utilise le symbole # comme suffixe du signet.

exemple : <http://www.det.fundp.ac.be/cefis/publications.html#REF5.40>
ou
#REF5.40, si on est déjà dans le document concerné.

Le signet est placé dans le document, à l'endroit voulu, grâce à la balise : `Texte du signet`.

Listes

Nous ne parlerons ici que des listes, dites liste à puces. Elles sont désignées par la balise `` `` et chaque élément de la puce est balisé par `` ``. La liste peut être décalée vers la droite en utilisant la balise `<BLOCKQUOTE>` `</BLOCKQUOTE>`.

Texte en HTML	Page Web
<pre><html> <head> <title>Nouvelle page 1</title> </head> <body> <blockquote> <blockquote> 1° élément de la liste 2° élément de la liste </blockquote> </blockquote> </body> </html></pre>	<ul style="list-style-type: none"> • 1° élément de la liste • 2° élément de la liste

Tableaux

Dans les pages HTML, les tableaux sont d'une grande utilité car ils permettent une mise en page précise. On peut ainsi disposer les textes et les images aux endroits voulus.

Les tableaux sont définis par les balises suivantes :

<TABLE> </TABLE>, qui délimite la table et possède 2 attributs intéressants **WIDTH="LARGEUR%"** et **BORDER=ÉPAISSEUR_BORD**.

<TR> </TR>, qui définit une ligne dans le tableau.

<TD> </TD>, qui définit une cellule dans la ligne et qui possède, elle aussi, l'attribut **WIDTH="LARGEUR%"**.

La balise TR signifie Table Row et la balise TD signifie Table Data.

Texte HTML	Page Web		
<pre> <HTML> <HEAD> </HEAD> <BODY> <TABLE WIDTH="100%" BORDER=1> <TR> <TD WIDTH="50%">1° colonne</TD> <TD WIDTH="50%">2° colonne</TD> </TR> </TABLE> </BODY> </HTML> </pre>	<table border="1" style="width: 100%; height: 100%;"> <tr> <td style="width: 50%; text-align: center;">1° colonne</td> <td style="width: 50%; text-align: center;">2° colonne</td> </tr> </table>	1° colonne	2° colonne
1° colonne	2° colonne		

Autres protocoles dans les documents Web

Le Web a été conçu à l'origine comme un système multimédia extensible basé sur l'hypertexte. Un protocole spécifique a donc été développé pour transporter ce genre de documents, le HTTP ou HyperText Transfer Protocol. Les documents Web peuvent cependant se référer à d'autres protocoles, comme par exemple :

- FTP (File Transfer Protocol), qui permet le transfert de fichiers et qui est utilisé dans les téléchargements. Exemple : `Texte du lien`.
- Mailto, qui permet d'ouvrir une session de courrier électronique à partir du navigateur de l'utilisateur. Exemple : `E-mail – Guy Vastersavendts`.
- NNTP (News Network Transfer Protocol), qui permet d'établir un lien avec un groupe de news. Exemple : `Texte du lien` ou `Texte du lien`



Remarque : il est bien entendu que l'hôte auquel on fait référence doit avoir un logiciel-serveur actif correspondant au protocole.

Retour à la case départ

Après ce petit tour d'horizon, vous avez appris un certain nombre de choses sur le langage HTML. Nous avons choisi les éléments essentiels pour vous permettre de démarrer dans votre lecture (décryptage) de pages HTML. Vous en voulez la preuve ! Reprenez la page du CeFIS du début de ce chapitre dans sa version HTML et décryptez-la.

Faut-il faire toutes ses pages Web avec un éditeur de texte ?

La réponse est heureusement non. Pourquoi heureusement ? Le HTML en lui-même, ça n'a pas l'air trop compliqué, mais écrire toutes ces pages Web comme cela, c'est un peu fastidieux. De plus, après avoir écrit une page en HTML pour voir ce que cela donne en réalité, il faut visualiser la page avec un navigateur. La démarche est donc assez longue et pourrait rebuter plus d'un auteur, même s'il est courageux.

Wysiwyg signifie What You See Is What You Get, autrement dit "ce que vous voyez à l'écran est ce que vous obtiendrez comme résultat.

Dans le chapitre suivant, nous allons voir d'autres types d'éditeurs, dont certains sont plus Wysiwyg. Nous n'abandonnerons pas le HTML pour autant. Même en utilisant un éditeur Wysiwyg, nous irons fréquemment voir le code HTML généré par cet éditeur. Ceci nous permettra de comprendre ce qui se passe derrière cet éditeur, mais aussi, le cas échéant, de modifier et de raffiner notre code HTML.

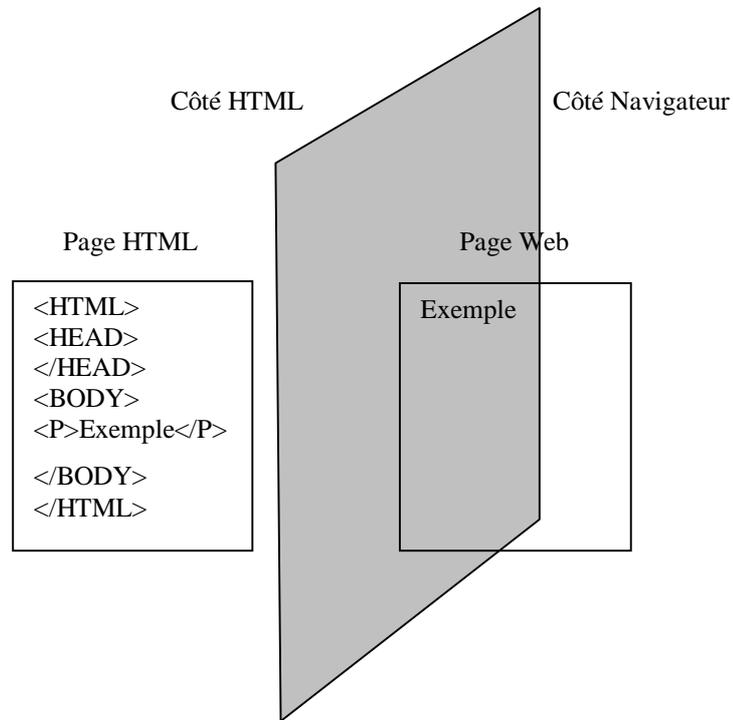
Une bonne connaissance des bases du HTML est de toute façon indispensable pour ajouter des scripts (Java ou VB) dans des pages Web. Nous verrons comment ces scripts introduisent un certain dynamisme dans les pages Web.

Résumons-nous !

HTML est bien un langage, dont la spécificité est d'être un langage de balisage qui permet de mettre en forme du texte. Ce langage comprend donc une série de balises, qui servent à donner une certaine apparence au texte. Le langage HTML comporte aussi des balises qui permettent de créer des liens et de transformer ainsi le texte en hypertexte. Les pages produites avec ce langage portent le nom de pages HTML. Elles sont, en général, stockées sur un serveur Web, qui les fournira à la demande.

De l'autre côté, nous avons un navigateur. Un navigateur est un logiciel spécialisé dans l'interprétation de pages HTML et dans leur affichage. Ce navigateur est aussi capable de faire des requêtes (via Internet) à des serveurs Web. On dira donc que le navigateur affiche des pages Web.

Page HTML ou page Web, cela dépend quel côté on regarde !



Exercices

Les exercices devront être réalisés avec un éditeur de texte (par exemple le "bloc-notes") et un navigateur (par exemple "internet explorer").

Le but pour chaque exercice est d'écrire un document HTML qui apparaîtra dans le navigateur tel qu'il est décrit ou visible ici.

Le contenu du texte de l'exercice peut parfois vous aider.

Exercice 1

Créer une page blanche avec un titre "Publier sur le Web" et le nom de l'auteur".

Exercice 2

Créer une page avec comme texte "Publier sur le Web" et le nom de l'auteur.

Exercice 3

Créer une page qui comporte les éléments suivants :

Ceci est un titre de niveau H1

Ceci est un titre de niveau H2

Exercice 4

Créer une page ayant l'aspect suivant :

Ceci est un paragraphe aligné à gauche.

Ceci est un paragraphe centré.

Ceci est un paragraphe aligné à droite.

Exercice 5

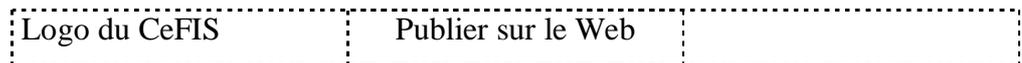
Créer une page qui comporte un titre de niveau 2, une image centrée et une liste à puces.

Exercice 6

Créer une page qui comporte un lien hypertexte avec une autre page et un lien hypertexte vers une marque de cette même page.

Exercice 7

Créer une page ayant l'aspect suivant :



Insérer ici un petit texte, dans lequel certains mots sont en gras, italique et souligné.

Insérer ici une image qui est centrée et qui sert de lien vers une autre page que vous aurez créée.

Aller à la page d'accueil des FUNDP (il s'agit du texte d'un lien qui pointe sur la page d'accueil des FUNDP)

Exercice 8

Créer tout d'abord un petit document avec un traitement de texte (Word par exemple), ensuite ajouter à la page Web précédente un lien qui permettra le téléchargement de ce nouveau document.