

R1.02 - WEB - CHAP 1 - INTRODUCTION HTML

Enseignants (S1) : Amira Ben Hadid, Nicolas Buyle-Bodin, Stéphane Leroux

Références :

- Le cours s'appuie sur le site de Pierre-Antoine Champin:
<http://champin.net/enseignement/intro-web>
- Ce travail est sous licence Creative Commons Attribution-ShareAlike 3.0



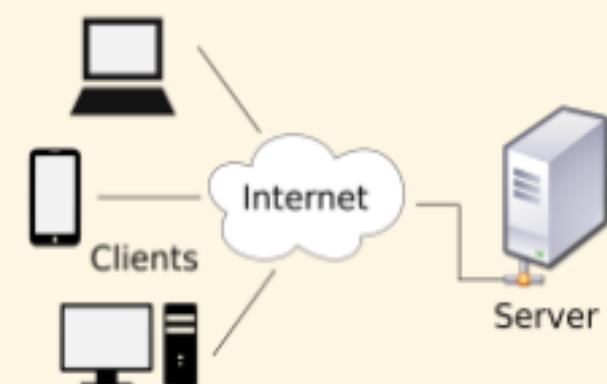
LE WEB EN S1

- Des séances:
 - 8 séances de TD (en groupe entier),
 - 4 séances de TP (en demi-groupe) dont 1 TP noté,
 - 2 séances de PE (projet encadré)
 - 1 DS individuel sur feuille (commun aux 5 groupes)
- Des notes :
 - une note pour la ressource R1.02 Web (DS individuel (coeff 2) + TP (coeff 1))
 - une note pour la SAE 1.05 Recueil des besoins (à partir du TP noté)
 - éventuelle mise en place de bonus dans certains groupes

COMPOSANTS DU WEB - 1/ HTTP

Quelques définitions :

- Le protocole HTTP pour **HyperText Transfer Protocol** (<https://httpwg.org/specs/>) décrit comment les données du Web sont échangées entre les machines.
- Une **Ressource** désigne une unité d'information (document, image, vidéo, ...) accessible sur le Web.
- Un **Serveur** est un ordinateur "contenant" des ressources. Il est toujours connecté à internet.
- Un **Client** est un ordinateur/smartphone/tablette/... utilisé pour afficher des ressources.
- Un **Navigateur** est un logiciel qui affiche les pages web.



source image : <http://commons.wikimedia.org/wiki/File:Client-server-model.svg>

COMPOSANTS DU WEB - 2/ HTML

- HTML pour Hypertext Markup Language : langage Hypertexte de balisage décrit comment les données doivent être interprétées par le navigateur
- HTML est le langage principal du Web, mais ce n'est pas le seul :
 - CSS (Cascading Style Sheets : feuilles de style en cascade) décrit la présentation des pages
 - Javascript : un langage de script du Web
 - ...

Ce semestre nous verrons surtout HTML5 et CSS3.

UTILITÉ DU HTML

Pourquoi apprendre le HTML alors que presque personne aujourd'hui n'écrit du HTML directement ? Les gens utilisent des traitements de texte et des générateurs.

- les générateurs ne permettent pas de tout faire
- le HTML généré a souvent besoin d'être retouché à la main
- vous pouvez être amenés à écrire des générateurs
- vous pouvez être amenés à écrire des programmes qui consomment du HTML (pour le traiter ou l'afficher)

POURQUOI HTML+CSS ?

Le HTML n'est pas (plus) un langage de présentation.

- HTML décrit la structure logique des pages
- CSS décrit leur présentation

Contrairement aux documents composés avec les traitements de texte, dont le produit fini est un PDF, le WYSIWIG (What You See Is What You Get) n'a pas vraiment de sens en HTML puisque la plupart du temps l'affichage d'un document HTML dépend en partie du contexte d'affichage.

La séparation entre fond et forme permet cette adaptation de l'affichage des sites à des clients de types différents (écrans d'ordinateur, tablette,...)

RESPONSIVE DESIGN

CONTENT IS LIKE WATER



“ You put water into a cup it becomes the cup.
You put water into a bottle it becomes the bottle.
You put it in a teapot, it becomes the teapot. ”

— Josh Clark - Seven deadly mobile myths

Source image : Wikipedia

LES OUTILS AU SEMESTRE 1

- un éditeur de texte (par exemple Notepad++) pour éditer les fichiers HTML et CSS
- un navigateur moderne pour visualiser les pages HTML
- Le W3C* fournit un service de validation en ligne : <http://validator.w3.org/> Son utilisation vous est vivement recommandée.
- Documentation :
 - le cours en ligne : <http://champin.net/enseignement/intro-web>
 - Dev Docs : <https://devdocs.io/> (agrégateur de documentations) (anglais)
 - MDN Web Docs : <https://developer.mozilla.org/fr/docs/Web> (anglais/français)

* W3C : World Wide Web Consortium est l'organisme de standardisation du web. Il implique 49 employés (dont P-A Champin)

UN LANGAGE À BALISES

- HTML est un format textuel
- Le texte normal sera affiché tel quel
- Les **balises** sont des codes qui ne sont pas affichés mais qui servent à indiquer le rôle du texte balisé et donc *indirectement* comment il sera présenté.

En anglais, les balises sont appelées **tags**. "Langage à balise" se traduit par *markup language*.

EXEMPLE

Si vous écrivez dans l'éditeur de texte :

```
Qu'est-ce qu'une <em>balise</em> ?
```

... vous verrez dans le navigateur :

Qu'est-ce qu'une *balise* ?

Le mot « balise » est marqué comme important (em = emphase), c'est pourquoi il est affiché (par défaut) en italique.

STRUCTURE DES BALISES

Qu'est-ce qu'une ``balise`` ?

- Tout texte entre les **chevrons** (en anglais *brackets*) < et > est considéré une balise
- `` est une balise *ouvrante*, `` est la balise *fermante* qui lui correspond; Elles délimitent le texte concerné.

Les éditeurs de texte reconnaissent les balises et les mettent automatiquement en couleur pour faciliter la lecture du code HTML.

EMBOÎTEMENT DES BALISES

On peut appliquer plusieurs balises au même texte, à condition de respecter la règle d'emboîtement : toute balise B ouverte à l'intérieur d'une balise A doit également être fermée à l'intérieur de A.

Exemples :

```
<A>Emboîtement <B> correct </B></A>
```

```
<A>Emboîtement <B> incorrect </A></B>
```

STRUCTURE EN ARBRE

De la règle d'emboîtement, il découle que les balises confèrent une structure d'arbre au document :

```
<X>Lorem <Z>ipsum</Z> dolor </X>sit.
```

```
<X>
  | Lorem | <Z> | ipsum | </Z>
  | dolor
</X>
sit.
```

Lorem ipsum est une suite de mots sans signification utilisée d'abord en imprimerie pour calibrer une mise en page.

ATTRIBUTS

Certaines balises, en plus du contenu textuel qu'elles délimitent, ont besoin d'informations supplémentaires appelés **attributs**.

Le format est **nom=valeur** à l'intérieur des chevrons de la balise ouvrante et après son nom.

Exemple :

```
<a href="http://champin.net/">P-A. Champin</a>
```

n'affichera que le texte « P-A. Champin » sous la forme d'un lien qui renvoie vers l'adresse indiquée dans l'attribut **href**.

BALISES VIDES

Certaines balises particulières n'attendent pas de contenu textuel. Ces balises vides n'ont donc pas de balise fermante correspondante

Exemple : la balise `img` pour insérer une image :

```

```

n'a pas de contenu donc pas de balise fermante.

On peut indiquer explicitement l'absence de balise fermante en ajoutant la barre oblique à la fin de la balise ouvrante.

GESTION DES ESPACES

Si vous écrivez

```
Gestion  
des espaces.
```

Vous verrez :

```
Gestion des espaces.
```

Explication : HTML considère tous les caractères d'espacement (espace, retours à la ligne,...) comme des séparations entre les mots et les affiche comme un *simple espace*.

On va voir dans la suite comment faire afficher des retours à la ligne.

STRUCTURE GLOBALE D'UN DOCUMENT HTML

```
<!DOCTYPE html>
<!-- version HTML5--&gt;
&lt;html lang="fr"&gt;
    <!-- contenu en français --&gt;
    &lt;head&gt;
        &lt;meta charset="UTF-8" /&gt; <!-- recommandé* --&gt;
        &lt;title&gt;Titre du document&lt;/title&gt;    <!-- obligatoire --&gt;
    &lt;/head&gt;
    &lt;body&gt;
        ...
    &lt;/body&gt;
&lt;/html&gt;</pre>
```

Tout document HTML comprend deux parties :

- la tête `<head>` qui contient les métadonnées non affichées
- le corps `<body>` qui contient le contenu à proprement parler du document.

* l'indication d'encodage est une balise sans contenu avec un attribut. Elle permet d'afficher les caractères spéciaux sans problème.

TITRES ET PARAGRAPHES

Un document typique est une séquence :

- de titres (ou en-tête, en anglais *heading*),
- de paragraphes.

Les titres ont différents niveaux d'importance, repérables à leur typographie (et parfois à leur numérotation).

TITRES ET PARAGRAPHES EN HTML

- Les paragraphes sont délimités par la balise `<p> ... <\p>`
- Les titres sont délimités par les balises `<h1>, <h2>, ..., <h6>` du plus important au moins important.

```
<h1>Sujet</h1>
<h2>Thèse</h2>
<p>Paragraphe d'introduction</p>
<h3>Argument 1</h3>
<p>...</p>
<h3>Argument 2</h3>
<p>...</p>
<p>...</p>
<h2>Anti-thèse</h2>
<h3>Contre-argument 1</h3>
<p>...</p>
<h2>Synthèse</h2>
<p>...</p>
```

Pas d'indentation dans l'éditeur car les balises ne sont pas contenues les unes dans les autres. A comparer avec la structure globale d'un document.

COHÉRENCE [1/2]

L'enchaînement des niveaux de titre doit être cohérent :

- le premier titre devrait toujours être de niveau 1,
- un titre ne devrait pas monter de plus d'un niveau par rapport au précédent.

« Si le <h2> ne se différencie pas assez du <h1> visuellement, peut-on mettre un <h3> à la place ? »
--> non, mauvaise pratique.

Le HTML doit être cohérent avec la structure du document : utiliser le css pour modifier l'aspect du h2

COHÉRENCE (2/2)

Ces règles de cohérence ne sont pas normatives, leur non-respect n'est pas signalé par le validateur, mais il est tout de même à proscrire.

C'est donc à vous d'être attentifs à bien les respecter.

Exercice 1

Lire les consignes sur moodle et utiliser la feuille distribuée. Faire les questions 1) à 6)

SECTIONS

Les titres définissent en fait une structure de plus haut niveau :

- chaque titre indique le début d'une *section*,
- qui se termine au prochain titre de même niveau ;

Une section contient donc

- un titre,
- éventuellement des paragraphes,;
- éventuellement une ou plusieurs section(s) de niveau suivant.

Depuis HTML5, on est encouragé à utiliser la balise `<section>`. Lorsqu'on étudiera CSS, on comprendra mieux l'intérêt de rendre cette structure explicite.

ILLUSTRATION

```
<h1>Sujet</h1>
<section>
  <h2>Thèse</h2>
  <p>Paragraphe d'introduction</p>
  <section>
    <h3>Argument 1</h3>
    <p>...</p>
    <p>...</p>
    <h3>Argument 2</h3>
    <p>...</p>
    <p>...</p>
  </section>
  <h2>Anti-thèse</h2>
  <h3>Contre-argument 1</h3>
  <p>...</p>
  <h3>Contre-argument 2</h3>
  <p>...</p>
  <h3>Contre-argument 3</h3>
  <p>...</p>
</section>
```

SECTIONS SPÉCIALISÉES

Les balises suivantes représentent des sections avec une sémantique particulière.

Balise **signification**

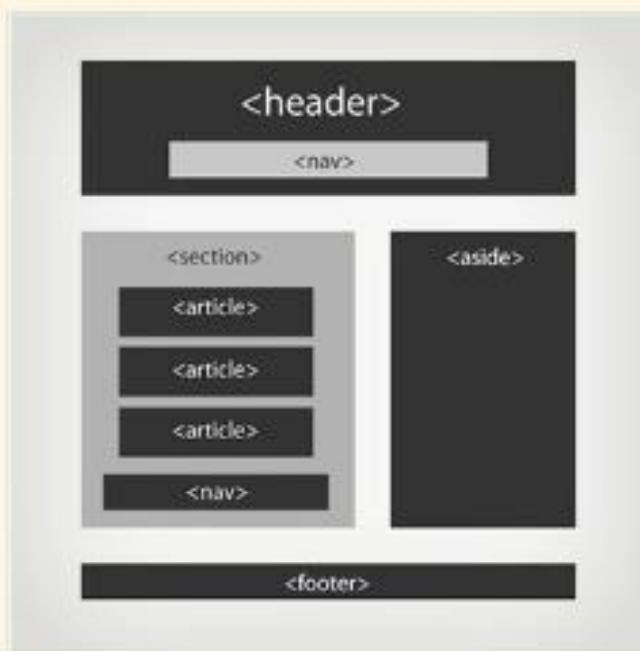
article contenu auto-suffisant

aside encart

header en-tête de la section parent (ou du doc)

footer pied de la section parent (ou du doc)

nav section contenant des liens de navigation



BALISES UTILES DANS LES TITRES ET PARAGRAPHES

Balise	signification	effet par défaut
< em >	emphase (important)	italique
< strong >	emphase forte (très important)	gras
< dfn >	définition	aucun
< cite >	titre d'ouvrage	italique
< q >	citation	guillemets
< code >	extrait de code informatique	police : monospace

Dans l'éditeur, vous tapez :

```
Selon <cite>Wikipedia\</cite>, la <dfn>prose</dfn> est :  
<q>la forme <em>ordinaire</em> de l'expression verbale.</q>
```

vous verrez dans le navigateur :

Selon Wikipedia, la prose est : "la forme ordinaire de l'expression verbale."

BALISES HISTORIQUES

Certaines balises datant des premières versions de HTML, où la séparation fond/forme n'était pas encore de rigueur, sont valides et acceptées par le validateur, mais il faut leur préférer une balise plus sémantique.

, <i>, <u>,
 ne doivent plus être utilisées.

LISTES NON NUMÉROTÉES

Une liste est un paragraphe d'un type particulier, contenant une énumération d'éléments.

```
<ul>
  <!-- ul pour unordored list -->
  <li>sucre</li>
  <!-- li pour list-item -->
  <li>céréales</li>
</ul>
```

- sucre
- céréales

LISTES NUMÉROTÉES

```
<ol>
  <!-- ol pour ordored list -->
  <li>sucre</li>
  <!-- li pour list-item -->
  <li>céréales</li>
</ol>
```

1. sucre
2. céréales

La numérotation est automatique mais on peut la modifier dans la feuille de style.

Terminer l'exercice 1

SENSIBILISATION À L'ACCESSIBILITÉ

Pour la prochaine séance, regarder la vidéo d'introduction à l'accessibilité du W3C :

<https://www.w3.org/WAI/videos/standards-and-benefits/fr> (4 minutes)

Point à observer (en plus du contenu) :

- la qualité des sous-titres.
- comparer avec le sous-titrage automatique d'une des vidéos de présentation du BUT sur le site de l'IUT : <https://iut.univ-lyon1.fr/formation/but/info-doua/but-informatique-doua>

COMPLEMENTS

Pour aller plus loin

Les diapositives suivantes présentent des compléments au cours.

Les contenus ne sont pas exigibles en examen

Elles peuvent servir sur des exercices plus avancés comme le TP de synthèse.

COMPLÉMENT SUR LES SECTIONS HTML

L'emboîtement des sections suffit à déterminer leur niveau, donc l'information portée par le numéro de titre est redondante. En HTML5, on peut n'utiliser que `<h1>` à tous les niveaux, les navigateurs compatibles adapteront la typographie en fonction de la « profondeur ».

Cependant, ceci n'est pas compatible avec les anciens navigateurs. De plus, les moteurs de recherche recommandent une seule balise `h1` par page pour un meilleur référencement.