

# Mettre du fond : HTML

Ce document est une version rapide du cours de SNT de seconde, à destination des élèves n'ayant que peu travaillé sur HTML/CSS.

## 1. C'est quoi une page Web

Commençons par placer du vocabulaire ainsi qu'un bref aperçu de l'histoire du web.

1. Donner une définition de **World Wide Web** :

 Done



2. Donner une définition d'**Internet** :

 Done



3. Donner une définition de **système hypertexte** :

 Done



4. Donner une définition de **Uniform Ressource Locator** (URL) :

 Done




5. Que signifie http et https ?

 Done



6. Expliquez ce qu'est **Standard Generalized Markup Language** :

 Done



## 2. Un langage de balise : HTML

Une page web, dans sa partie logique, **c'est un fichier contenant du texte**, ce texte étant enrichi par un système de **balises ouvrantes** - par exemple `<p>` et fermantes - par exemple `</p>`, permettant de donner un sens particulier (autrement dit une *sémantique*) au texte encadré par ces balises (titres, chapitres, paragraphes, illustrations...).

Le langage utilisé qui contient le texte et les balises s'appelle **HTML** ( d'où le nom de l'extension de fichier `.html` ), qui est un acronyme pour *Hyper Text Markup Language*, soit langage de balisage hypertexte. Il est à noter que HTML *n'est pas un langage de programmation*, mais simplement un langage de description du texte, privilégiant la mise en avant du sens avant la forme.

Il existe de nombreux langages de balisages, permettant de donner une forme particulière à du texte ou à des chaînes de caractères. Par exemple, **HTML** pour le Web, **L<sup>A</sup>T<sub>E</sub>X** pour les documents scientifiques et les formules mathématiques, **XML** pour organiser des données, ou même **Markdown**, utilisé dans les Notebooks et pour la conception de ce site.

Un fichier `html` se présente donc sous la forme suivante :

```

<!DOCTYPE html>

<html lang="fr">
  <head>
    <meta charset="utf-8" >
    <title>Une simple page web &#129409; </title>

  </head>

  <body>
    <h1>Ma première page web</h1>
    <h2>Le code html</h2>
    <!-- ceci est un commentairen comme en Python, il n'est pas interprété par le navigateur, et n'est là
que pour les codeurs ou les curieux -->
    <p> Pour faire une page web, il faut taper du code <strong>HTML</strong>.
    <p> Afin de faire fonctionner celui-ci, il faut placer des <strong>balises</strong> qui encadrent des
portions du texte :
    <p>
    <ul>
      <li> Une balise ouvrante, utilisant les symboles &lt; et &gt; , qui encadrent un mot clé ;</li>
      <li> Une balise fermante, identique à la balise ouvrante, mais un slash / précède le mot clé ;</li>
      <li> éventuellement des <strong>attributs</strong> viennent compléter le mot clé, afin de préciser
un peu plus la balise, par exemple en donnant un identifiant particulier, ou l'url d'un lien.</li>
    </ul>

    <h2> Les différentes balises</h2>
    <p> Il existe de nombreuses balises en HTML. On peut trouver des explications sur leur fonctionnement en
suivant <a href="https://developer.mozilla.org/fr/docs/Apprendre/HTML/Balises_HTML">ce lien</a></p>
    <p> Une autre possibilité est de télécharger la fiche suivante (une cheat sheet) : </p>
    <img src = "https://media.cheatography.com/storage/thumb/rmathieu13_html.750.jpg" alt ="Image de cheat
sheet HTML" title="Cheat Sheet HTML" >
  </body>

</html>

```

### Notepad++

Pour l'édition de fichier de type `.html`, il ne faut **surtout pas** utiliser un traitement de texte classique, comme Word ou LibreOffice. En effet ceux-ci rajoutent de manière cachée des informations <sup>2</sup> dans le fichier en édition, ce qui risque de présenter de sérieux problèmes lors de l'affichage de la page web par un navigateur.

Dans la suite de ce chapitre, nous utiliserons majoritairement le logiciel **Notepad++**, qui est un éditeur de texte orienté codage, reconnaissant la plupart des langages informatiques. Ce logiciel est libre, disponible sur les ordinateurs du Lycée, et peut-être téléchargé en suivant [ce lien](#). Notez qu'il existe aussi en version portable, et que vous pouvez donc l'installer sur une clé USB.

### une première page web

1. Ouvrez le logiciel **Notepad++**
2. Créez un fichier vide nommé `pagesimple.html` dans un dossier `NSI/Web`.
3. Copiez dans le corps de ce fichier le contenu HTML présenté au dessus.
4. Enregistrez le fichier.
5. Avec l'explorateur de fichier, allez dans le dossier `NSI/Web` et double-cliquez sur le fichier `pagesimple.html`. Que se passe-t-il ?
6. Quel est le rôle des balises présentes dans ce document ?

## 3. Structure d'un document HTML et balises les plus utiles.

Une balise `HTML` classique est de la forme suivante :

```
<p class="Toto"> ... </p>
```

où :

- `p` est l'identificateur de balise, ici celui d'un *paragraphe* ;
- `class` est un **attribut** de la balise `<p>` ;
- `"Toto"` est la **valeur de l'attribut** `class`. Une valeur est **toujours entre guillemets**.
- `</p>` est la balise fermante. Entre les deux balises se trouve le corps du paragraphe.

#### Cas particulier : la balise `<img>`

La balise `<img>` est la balise permettant d'insérer des images dans le document HTML. Elle possède une particularité : elle est **auto-fermante** (car elle n'encadre aucun élément, elle contient la référence vers l'image).

Elle possède entre-autres les attributs suivants :

- `src` : l'adresse de l'image, absolue ou relative, ou sous la forme d'une URL (Attention aux URL google image, qui ne sont que temporaires !);
- `alt` : le texte de remplacement si l'image est manquante (voir remarque ci-dessus) ;
- `title` : le texte de l'info-bulle de l'image
- `width` : la largeur de l'image, en pixels ou en pourcentage ;
- `height` : la hauteur de l'image, en pixels ou en pourcentage ;
- ...

Les balises **HTML** sont des éléments précisant la **nature** du texte qu'elles encadrent. En aucun cas elles ne donnent d'indications de placement, d'alignement, bref de mise en page.

Il s'agit donc de préciser un **découpage logique** de la page.

#### Les balises

Pour chacune des balises suivantes, donner leur rôle et leurs principaux attributs, en s'aidant de la page [développeur Mozilla](#).

1. `<html>`
2. `<head>`
3. `<meta>`
4. `<title>`
5. `<body>`
6. `<header>` et `<footer>`
7. `<h1>`, `<h2>`, ..., `<h6>`
8. `<ol>`, `<ul>` et `<li>`
9. `<div>`
10. `<a>`
11. `<br>`

#### Les balises et attributs obsolètes

Attention, vous trouverez dans de nombreux sites web des conseils forts mauvais pour l'utilisation des balises. certains en effet continuent d'utiliser des **balises dépréciées ou obsolètes**, comme `<center>` par exemple.

Vous pouvez trouver [ici](#) quelques-unes de ces balises et attributs

### ? A vous de jouer

Créez une page web sur le sujet que vous souhaitez, et qui respecte au moins les conditions suivantes :

1. Un titre
2. Deux parties de 3 paragraphes, avec au moins une liste (numérotée ou non).
3. Un lien interne à la page
4. Un lien externe à la page
5. Une image locale (sur l'ordinateur)
6. une image distante (par URL)
7. Une incrustation vidéo (voir directement depuis YT pour obtenir le lien)

## 4. Vérifier son code : le vérificateur du W3C

Le **W3C** (*World Wide Web Consortium*) est l'organisme à buts non lucratifs, chargé de standardiser et promouvoir les technologies du web. Sa gestion est assurée conjointement par le MIT aux États-Unis, l'ERCIM en Europe (dont l'INRIA en France), l'université Keio au Japon et l'université Beihang en Chine. À côté d'industriels et d'éditeurs informatiques, en particulier les éditeurs des navigateurs (Mozilla Fondation, Microsoft, Apple, Opera ou Google), ses membres se composent aussi des centres de recherches (Inria, National Research Council Canada, etc.), des opérateurs de réseaux, ainsi que des entreprises investies dans le Web comme Braillenet, la Bibliothèque du Congrès ou la BBC.

Le W3C offre un service de **validation des pages web**, vérifiant que leur structure est conforme aux normes en vigueur. On le trouve à l'adresse <https://validator.w3.org/>.

On peut y tester :

- soit une URI (*Validate by URI*) ;
- soit un fichier (*Validate by File Upload*);
- soit en copiant-collant le code HTML (*Validate by Direct Input*).

1. [https://fr.wikipedia.org/wiki/Uniform\\_Resource\\_Locator](https://fr.wikipedia.org/wiki/Uniform_Resource_Locator) ←

2. Les traitements de textes classiques comme LibreOffice, Word, Google Doc, sont appelés des traitements de textes **WYSIWIG** (*What you see is what you get*). Ils sont extrêmement puissants, mais parfois difficiles à manipuler. De nombreux utilisateurs ne les utilisent que comme un substitut à la machine à écrire, sans avoir conscience des puissantes capacités d'édition de ces logiciels. Ces logiciels ne fabriquent en réalité pas un simple document texte, mais en réalité une archive (comme un type `zip`, `7z` ou `rar`), qui contient l'ensemble des informations textuelles, de formatage, et hypertextuelles du document. Il faut noter aussi qu'il existe des logiciels de traitements de textes non WYSIWYG, tels que ceux utilisant *L<sup>A</sup>T<sub>E</sub>X*, extrêmement puissants puisqu'ils permettent à un **utilisateur averti** de formater des documents exactement comme il le souhaite. Ces traitements de textes sont utilisés dans le domaine de la recherche, et savoir utiliser *L<sup>A</sup>T<sub>E</sub>X* est un véritable atout pour un étudiant en sciences. ←