

# 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

Le **World Wide Web** ( `www` ou *3W*) est un système *hypertexte* fonctionnant sur *Internet*. Il a été inventé par *Tim Berners-Lee* et *Robert Cailleau* au CERN de Genève, en 1989.

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

✓ ▼ Done

**Internet** est le réseau mondial d'ordinateurs accessibles au public. Nous verrons un historique rapide de la création d'Internet dans la partie *Réseau* du programme.

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

✓ ▼ Done

Un **hypertexte** est un document ou un ensemble de documents contenant des unités d'information liées entre elles par des **hyperliens**. Ce système permet à l'utilisateur d'aller directement à l'unité qui l'intéresse, à son gré, d'une façon non linéaire. Le terme a été créé en 1965 par *Ted Nelson*, sociologue américain.

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

✓ ▼ Done

Une URL (sigle de l'anglais : *Uniform Resource Locator*, littéralement « *localisateur uniforme de ressource* »), couramment appelée adresse web, est une chaîne de caractères uniforme qui permet d'identifier une ressource du World Wide Web par son emplacement et de préciser le protocole internet pour la récupérer (par exemple `http` ou `https`).<sup>1</sup>

5. Que signifie `http` et `https` ?

✓ ▼ Done

`http` signifie *HyperText Transfer Protocol*, soit littéralement \* « Protocole de transfert Hypertexte ». C'est un protocole de communication client-serveur développé pour le World Wide Web. Il existe aussi en version **sécurisée** (protocole `https`).

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

✓ ▼ Done

**Standard Generalized Markup Language** (« langage de balisage généralisé normalisé &raquo; - **SGML**) est un langage de description à balises, de norme ISO (ISO 8879:1986).

Ce langage, publié en 1986 par Charles Goldfarb, mais dont les prémisses remontent à 1969, est un langage qui a eu pour vocation de s'affranchir des capacités techniques spécifiques des imprimantes pour décrire un document texte. Ce standard, vite adopté par la communauté internationale, repose sur la séparation complète entre la *structure logique* d'un document (titres, chapitres, paragraphes, illustrations,...), qui est identifié par des **balises** insérées dans le document lui-même, de sa *mise en page*, qui dépend du support de présentation (livre, journal, présentation, écran, ...).

**HTML** (*HyperText Markup Language*) est une application de SGML, créée par Tim Berners-Lee en 1989.

## 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 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 commentaire 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 :
    <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>
    <p>
    <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>
  </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 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>` et `<span>`
10. `<a>`

## ? 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
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

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

Le **W3C** (*World Wide Web Consortium*) est l'organisme à but non lucratif 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 Foundation, 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 Brailenet, 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/>.

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