

Les langages du World Wide Web



Le web est géré au niveau mondial par le W3C - World Wide Web Consortium. Cet organisme international indépendant définit les standards permettant aux ordinateurs connectés de pouvoir communiquer ensemble. Ces standards concernent les règles de communication, ainsi que les formats et langages à utiliser.

Plusieurs langages ont été standardisés par le W3C pour décrire le contenu des pages (HTML), leur forme (CSS) et pour programmer les interactions (JavaScript).

1. Le langage HTML : le langage de l'hypertexte

Le langage HTML (Hypertext Markup Language) a été conçu pour standardiser le format des pages envoyées par un serveur à un navigateur web. HTML est un langage de description des pages web permettant de structurer un texte en lui ajoutant des balises pour préciser comment le navigateur doit interpréter et mettre en forme le contenu.

Les balises

La page HTML contient le texte à afficher (contenu) et des balises entre les symboles `<>` donnant des indications sur la manière d'afficher le document. Les balises servent à délimiter une partie du texte situé entre une balise de début (par exemple `<title>`) et la balise de fin correspondante (`</title>`).

Ces indications sont à interpréter par le navigateur en fonction de la taille de l'écran et des préférences de l'utilisateur. La balise `<p>` signale les paragraphes, la balise `<h1>` les titres de niveau 1. Les hyperliens sont décrits par la balise `<a>` qui permet de donner l'URL d'une autre page et le contenu du texte à cliquer. La liste des balises usuelles est donnée dans l'aide mémoire HTML présent dans la boîte à outil sur Moodle.

Structure principale d'un fichier HTML

Un document HTML est composé de deux parties séparées par des balises :

- x son **entête** entre les balises `<head>` et `</head>`
- x son **corps** entre les balises `<body>` et `</body>`

```
1 <!doctype html>
2 <html lang="fr">
3 <head>
4   <meta charset="utf-8">
5   <title>Hello HTML</title>
6 </head>
7 <body>
8   <h1>Hello HTML</h1>
9   <p>Bye now.</p>
10 </body>
11 </html>
```

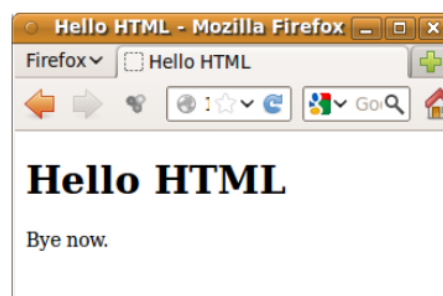


Figure 1: Fichier HTML minimal et son rendu

Dans l'entête on peut retrouver par exemple l'encodage, le titre de la page, le style de la page. Dans le corps, on retrouve l'ensemble des éléments publiés sur la page (texte, images, liens...)

Le document HTML peut être représenté par une arborescence de nœuds représentant les balises. Le premier parent est l'élément racine `<html>`, les "branches" enfants sont les éléments imbriqués et les "feuilles" sont le contenu des éléments (voir figure 2). Ce mode de représentation est appelé le DOM. Il constitue un modèle simple utilisable par divers programmes externes tels que Javascript (voir plus loin).

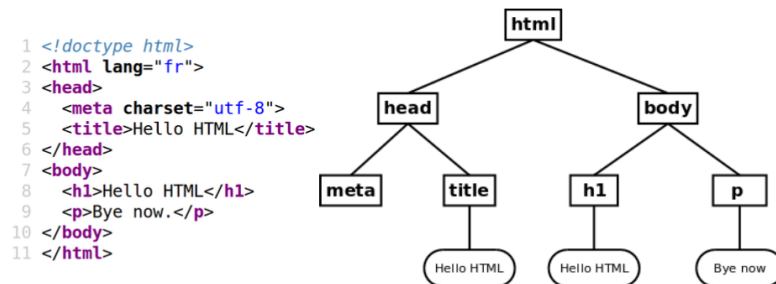


Figure 2: Un fichier HTML, sa représentation en arborescence (DOM) et son rendu

2. CSS : un langage de style.

Le style graphique d'une page (polices utilisées, tailles, couleurs...) peut être précisé par un fichier de style au format CSS.

Un fichier CSS (de anglais Cascading Stylesheet) contient des règles de mise en forme. Par exemple, tous les titres de niveau 1 (marqués par les balise `<h1>`) sont de telle couleur et de telle police. L'usage de fichiers CSS contribue à homogénéiser la présentation d'une page web et à séparer le fond, c'est-à-dire le contenu, de la forme.

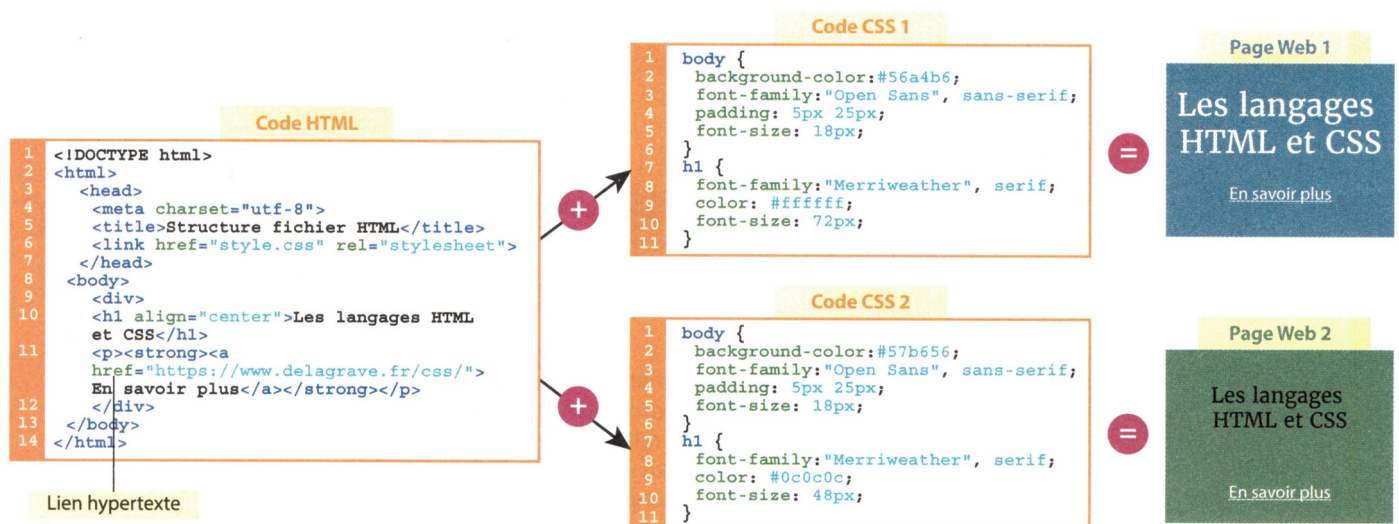


Figure 3: Application de styles CSS différents

