

# 3IN017 - TECHNOLOGIES DU WEB

HTML – *HyperText Markup Language*

14 janvier 2025

Gilles Chagnon

# Introduction

# Origine

Vers 1990, par Sir Tim Berners-Lee au CERN

Invention du World Wide Web = convergence de trois technologies :

**URL** (*Universal Resource Locator*) : protocole d'adressage de ressources ;

**HTTP** (*HyperText Transfer Protocole*) : protocole de communication client/serveur ;

**HTML** (*HyperText Markup Language*)

C'est un langage de balisage issu de SGML (dont est issu en parallèle XML), avec un arbre d'éléments et d'attributs.

# Exemple de document HTML

```
<!DOCTYPE html>
<html lang="fr">
<head>
  <meta charset="utf-8" />
  <title>Nouveau document HTML</title>
</head>
<body>
  <h1>Un grand titre</h1>
  <!-- Un commentaire, caché à l'écran -->
  <p>Paragraphe avec <a href="http://site.com">un lien</a>&thinsp;!</p>
  <ul>
    <li>un item</li>
    <li>un autre</li>
  </ul>
</body>
</html>
```

# Principes fondateurs

- Format texte ouvert
- Facile à lire et à écrire (pour les machines et les humains)
- Sémantique, accessible (est consultable par les personnes handicapées si bien fait)
- Robuste, résilient (pas de « plantage »)

... des principes cependant parfois encore remis en cause (versions propriétaires, monopoles, etc.)

# Historique

- 1994 W2C créé par T. Berners-Lee pour protéger et faire évoluer les standards
- dans les années 90, différentes versions (2.0, 3.2, 4.01...)
- depuis 2004, une seule version, un *living standard*, HTML 5, développé conjointement par le W3C et le WHATWG (consortium d'éditeurs de navigateurs)

Quelques liens utiles :

- les spécifications : <https://html.spec.whatwg.org/multipage/>
- Mozilla Developer Network : une référence incontournable pour le développement : <https://developer.mozilla.org/fr/>

# Éléments HTML : structure globale

# Structure de base

```
<!DOCTYPE html>
<html lang="fr">
  <head>
    <!-- Entête -->
  </head>
  <body>
    <!-- Corps -->
  </body>
</html>
```

- `<!DOCTYPE html>` : déclaration de type de document, unique et toujours en toute première ligne
- `html` : élément racine, unique, qui contient dans cet ordre :
  - 1 head : entête unique
  - 2 body : corps de la page, unique



# L'entête head

## Contient

- les méta-informations de la page :
  - le titre
  - l'encodage de caractères
- les appels aux ressources annexes :
  - icône du site
  - feuilles de style CSS
  - fichiers JavaScript
  - flux RSS...

```
<head>  
  <meta charset="utf-8" />  
  <title>Mon site perso</title>  
  
  <meta name="description" content="Mes centres d'intérêt">  
  <meta name="viewport" content="width=device-width, initial-scale=1">  
  <meta name="author" content="Victor Hugo">  
  
  <link rel="stylesheet" media="screen" href="styles/main.css">  
  <link rel="alternate" type="application/rss+xml" href="feed.xml">  
  <link rel="icon" href="icone.png">  
  
  <script src="scripts/main.js"></script>  
  
</head>
```

# Structure macro du document

De grands blocs structurants : main, main, nav, header, footer, content, article, section, aside

