

HTML : Historique, technique et nouveaux usages

1^{er} octobre 2014

Table des matières

- 1 Introduction
- 2 Evolution du langage, son histoire
- 3 Techniques
- 4 Nouveaux usages

Sommaire

1 Introduction

- Présentation
- Un web sémantique
- Un outil puissant

2 Evolution du langage, son histoire

3 Techniques

4 Nouveaux usages

HTML pour HyperText Markup Language

Format de données conçu pour représenter les pages web et permettant d'écrire de l'hypertexte (systèmes contenant des noeuds liés entre eux par des hyperliens permettant de passer d'un noeud à un autre, e.g. : passage d'une page web à une autre dans le World Wide Web).

Dérivé du Standard Generalized Markup Language (SGML, langage normalisé de balisage généralisé).

Langage de balisage

- Utilise des **balises** pour structurer un document et son contenu qui est ensuite compréhensible par un programme informatique (permet un traitement automatique du contenu).
- A terme, cette évolution permettra aux machines de réaliser toutes les tâches qui amèneront à un Web plus intelligent et rendra la recherche d'informations plus efficace (but du **web sémantique**).

HTML permet aussi :

- D'inclure des ressources multimédia (images, vidéos, audio, canvas (pour le dessin 2D/3D dans le navigateur))
- De créer des formulaires de saisie
- De créer des tableaux
- De créer des programmes informatiques (intranet (restreint à la partie site web interne))

- HTML est **interopérable** (e.g. : site web desktop/mobile, différents navigateurs et versions, personnes malvoyantes/malentendantes, etc.)
- Utilisé conjointement avec d'autres technologies (e.g. : Javascript, CSS (pour la présentation), PHP/Python/C/C++, etc.)
- Extension : .html (parfois .htm)

HTML (et plus généralement les technologies du World Wide Web) est promu par le World Wide Web Consortium (W3C), qui regroupe en février 2013, 383 entreprises partenaires (dont Mozilla, Microsoft, Apple)

Remarques

- Quand on parle de HTML5, on parle souvent d'un ensemble de technologies comprenant HTML5, CSS3, Javascript, permettant ainsi de développer des applications.
- HTML5 fait de la concurrence aux plugins tiers tel que Flash Player d'Adobe ou Silverlight de Microsoft.

Sommaire

- 1 Introduction
- 2 Evolution du langage, son histoire
 - Chronologie
- 3 Techniques
- 4 Nouveaux usages

Origines (1989 - 1992)

Fait parti des 3 inventions du World Wide Web (W3), avec HTTP et les adresses web. Conçu pour écrire des documents hypertextuels (appelés aujourd'hui « page web »)

1ers éléments du HTML :

- Titre du document
- Hyperliens
- Structuration en titre, sous titre, texte brut
- Listes

Pas de spécifications (les navigateurs gèrent le HTML comme ils le souhaitent)

Apports de NCSA Mosaic (1993)

- Tentative de normalisation du langage jusqu'à la fin des années 1990, mais infructueuse.
- Apparition de la balise `` (pour les images GIF et XBM) et des formulaires grâce au navigateur NCSA Mosaic => 1ers sites web de commerce électronique.

Apports de Netscape (1994)

- Apparitions de nouveaux éléments de présentation : attributs de texte, clignotement, centrage, etc.
- Apparition des tableaux pour améliorer la présentation des documents.
- Divergence de HTML : les développeurs de navigateurs cherchent à maximiser l'impact visuel
- Extension des capacités de description sémantiques (logos, notes de bas de pages, etc.) et les domaines d'application (mathématiques) de HTML par les concepteurs.
- Apparition des premières feuilles de style en cascade (CSS)

HTML 2.0 (1995 - 1996)

- Création du W3C (World Wide Web Consortium) (mars 1995)
- Abandon de HTML+ (brouillon d'html 3) ajoutant notamment : les tableaux, les figures et les expressions mathématiques (septembre 1995)
- Fin 1995 : Finalisation de la RFC (Request For Comment) décrivant HTML 2.0 (ne prend pas en compte les améliorations de Netscape).

HTML 3.2 et HTML 4.0 (1997)

■ Janvier 1997 : HTML 3.2

Prend en compte les ajouts de Netscape et Internet Explorer :

- Standardisation des tableaux.
- Ajout d'éléments de présentation et de balises destinées à recevoir du css et du javascript.

■ Décembre 1997 : spécification de HTML 4.0

- Standardise les styles, scripts, frames, objets (inclusion généralisée de contenu).
- Séparation plus explicite entre la forme et le fond d'un document.

XHTML (2000 - 2006)

- XHTML (Extensible HyperText Markup Language) est un langage de balisage fondé sur la syntaxe du XML (Extensible Markup Language) et conçu à l'origine comme successeur de HTML car sa syntaxe est plus simple à utiliser.
- Aucune fonctionnalité n'a été rajoutée ou retirée entre XHTML et HTML 4. Seule la syntaxe change.

HTML5 (2007 - 2014+)

- Dernière révision majeure de HTML. Début du travail en 2007, développement à partir de 2013.
- La syntaxe de HTML5 est complètement revue pour la rendre le plus simple et intuitif possible à l'utilisation.
- De plus, de nombreuses nouvelles balises sont introduites, notamment pour donner un sens au document (pour un web plus sémantique).
- Le langage comprend également une couche application avec de nombreuses API, ainsi qu'un algorithme afin de pouvoir traiter les documents à la syntaxe non conforme.
- Le W3C encourage les développeurs Web à utiliser HTML5 dès maintenant.

Sommaire

1 Introduction

2 Evolution du langage, son histoire

3 Techniques

- Les balises
- Les attributs
- Quelques balises
- HTML & CSS

4 Nouveaux usages

Les balises

La structure d'une page HTML est construite à partir de balises qui ont pour rôle l'organisation du contenu.

Il existe plusieurs sorte de balise :

- **Les balises en paires** s'ouvrent et se referment plus loin (le contenu est au milieu)
- **Les balises orpheline** qui servent à insérer un élément à un endroit précis.

Exemple de balises en paires

```
<block>  
    Il y a du texte Il y a du texte  
    Il y a du <inline>texte</inline>  
    Il y a du texte Il y a du texte  
</block>
```

Remarques

Une balise doit donner un sens au contenu qu'elle englobe.
Il est possible d'imbriquer les balises les une dans les autres (principe d'encapsulation).

Les balises doivent être fermées dans l'ordre de leurs ouvertures :
<xxx><yyyy>.....</xxx></yyyy> n'est pas correct !

Exemple de balise orpheline

```

```

Remarque

Le / (slash) final n'est pas obligatoire. Il sert juste à différencier les balises orphelines des balises en paires.

Les attributs

L'attribut sert à ajouter des informations à une balise (e.g. : l'encodage d'un document, le chemin vers une ressource, etc.).

Remarques

Un attribut s'écrit toujours sous la forme :

```
attribut="valeur"
```

Cependant, la valeur de l'attribut est parfois optionnelle et prend dans ce cas une valeur par défaut (e.g. : pour rendre un champ de formulaire obligatoire, on peut rajouter `required` dans la balise, sans lui donner de valeur)

Remarques

- Les attributs se placent dans les balises ouvrantes
- La valeur doit être écrite entre guillemets (") (mais ce n'est pas une obligation tant que la valeur ne fait pas plus d'un mot)
- Les attributs sont caractéristiques du type de la balises, on ne peut pas mettre n'importe quel attribut dans une balise.

L'attribut id

L'attribut **id** permet d'attribuer un identifiant unique à un élément. Il ne peut y avoir plusieurs éléments avec le même identifiant. Cet identifiant peut servir à structurer le document grâce à CSS, mais il peut aussi servir dans une encre.

L'attribut class

L'attribut **class**, contrairement à l'attribut à l'id, n'est pas unique, c'est-à-dire que l'on peut ré-utiliser la classe dans une autre balise. Il permet de styliser une partie du document avec l'aide de CSS.

Présentation rapide de quelques balises I

Balises multimédia

- Balise `<video></video>` : permet l'intégration d'une vidéo dans un document
- Balise `<audio></audio>` : permet l'intégration d'un son dans un document
- Balise `` : permet l'intégration d'images dans un document

Présentation rapide de quelques balises II

Remarque

Pour afficher des boutons de contrôle dans ces balises, il faut ajouter l'attribut **controls**.

```
<audio src="my_song.mp3" controls></audio>
```

Présentation rapide de quelques balises III

Balises de structure

- Balise `<header></haeder>` : entête d'une section déjà existante
- Balise `<article></article>` : section de contenu indépendante, pouvant être extraite individuellement du document ou syndiquée (flux RSS ou équivalent), sans pénaliser sa compréhension
- Balise `<section></section>` : section de contenu plus générale que `<article>` (mais doit être placé dans un article, par exemple).

Présentation rapide de quelques balises IV

Balises de structure

- Balise `<aside></aside>` : section dont le contenu est un complément par rapport à ce qui l'entoure, qui n'est pas forcément en lien direct avec le contenu mais qui peut apporter des informations supplémentaires.
- Balise `<nav></nav>` : Section possédant des liens de navigation principaux (au sein du document ou vers d'autres pages)
- Balise `<footer></footer>` : Section de conclusion d'une section ou d'un article, voire du document entier (pied de page).

Présentation rapide de quelques balises V

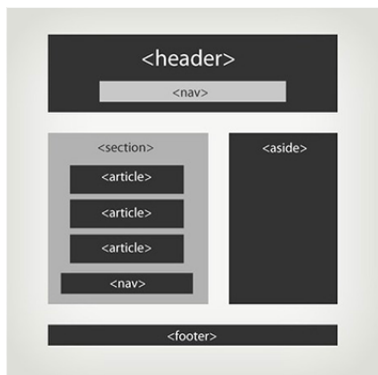


Figure: Exemple de structuration de page HTML - Copyright alsacreations.com

Présentation rapide de quelques balises VI

- Balises `<form>`, `<input>`, `<textarea>`, `<select>`, etc. : permettent la création de formulaires
- Balise `<table>`, `<tr>`, `<td>`, etc. : permettent la création de tableaux
- Balise `<p>` : indique qu'il s'agit d'un paragraphe
- Balise ``, `` : indiquent qu'il s'agit d'un contenu plus ou moins important
- Balise `<h1>` à `<h6>` : indiquent les niveaux de titres du plus important (h1) au moins important (h6)
- Balise `<a>` : pour naviguer d'une ressource à une autre (peut-être un lien interne ou externe ou une encre)

Les balises `div` et `span`

Les balises `<div>` et `` servaient à l'époque à structurer le document HTML : `<div>` pour agencer le contenu et `` pour associer un style à un élément. Grâce à HTML5, ce n'est plus nécessaire.

Sémantiquement ces balises n'ont pas de sens particulier.

Structure de base d'une page HTML5

```
<!doctype html>
<html lang="fr"> <!-- Permet aux liseuses de savoir
    quelle langue utiliser -->
    <head>
        <meta charset="utf-8" /> <!-- Encodage de
            la page -->
        <title>Yoo !</title>
        <!-- script et compagnie -->
    </head>

    <body>
        <!-- Le blabla de la page : img, p,
            progress, form, table, etc. -->
    </body>
</html>
```

Le CSS

- Le CSS (Cascading Style Sheets) s'occupe de la mise en forme d'un document HTML. Il s'utilise conjointement avec celui-ci.
- Pour inclure une feuille de style dans un document html, il faut rajouter dans l'entête :

```
<link rel="stylesheet" href="style.css" />
```

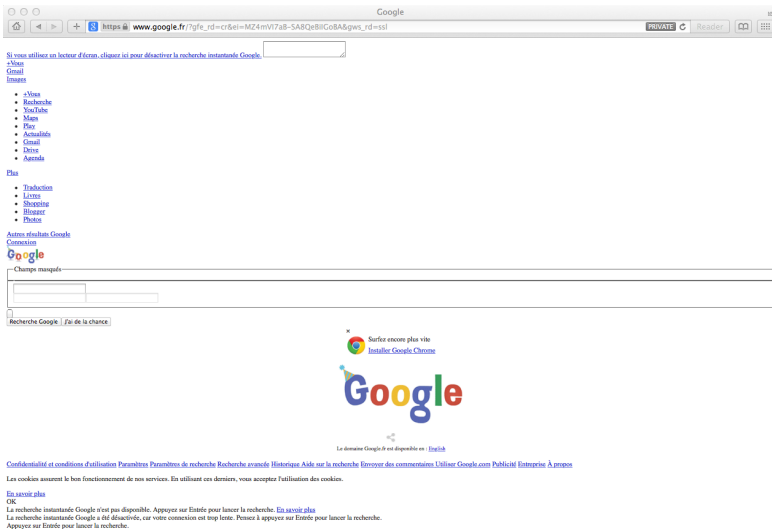



Figure: Site web google.fr sans CSS

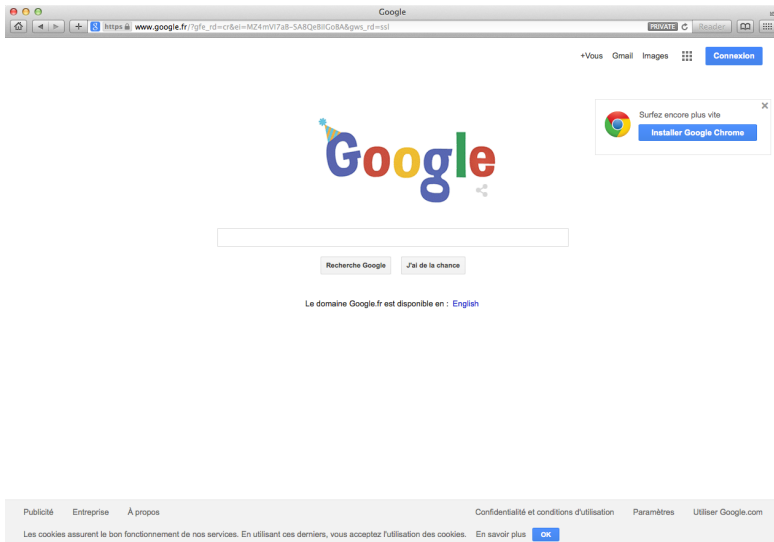


Figure: Site web google.fr avec CSS

Sommaire

- 1 Introduction
- 2 Evolution du langage, son histoire
- 3 Techniques
- 4 Nouveaux usages**
 - Domaine 3D
 - Echange et communication
 - Application internet avancée
 - Ergonomie & interactivité
 - La sémantique

WebGL

- Le **WebGL** est une spécification d'affichage 3D pour les navigateurs web qui permet d'utiliser le standard OpenGL ES 2.0 grâce à l'élément **Canvas** de HTML5 afin de créer un environnement 3D.
- Il a été introduit par le Khronos Group.

Websockets

Websockets est un protocole de communication bidirectionnel permanent entre un client et un serveur via un seul socket TCP (Transmission Control Protocol = permet de transférer des données).

Quelques exemples d'utilisation :

- Discussion instantanée (visioconférence, chat)
- Jeu multijoueur en ligne

Remarque

Un socket est une interface logicielle permettant d'exploiter facilement et de manière uniforme les services d'un protocole réseau.

Rich Internet Application (RIA)

A quoi servent les RIA ?

Les RIA sont des applications utilisable en ligne ou hors-ligne grâce à des API introduites par HTML5.

Exemples d'API :

- **WebStorage** : une alternative plus complète aux cookies
- **AppCache** : permettant aux applications d'avoir leur propre cache local

Autres usages de HTML5

- Le glisser/déposer
- Le couper/coller
- La géolocalisation
- L'affichage de barres de progression en pourcentage (progress, e.g. : upload)
- L'affiche de barres de progression en volume (meter, e.g. : espace disque utilisé)
- Contenu éditable

Certains système d'exploitation utilise une interface et un écosystème d'applications développées à l'aide de HTML5 (e.g. : Firefox OS ou Google Chrome OS).

Les microdatas I

Les **microdatas** sont des attributs intégrés par HTML5 ayant pour objectif de donner du sens aux éléments HTML. Cela permet d'éviter l'utilisation de l'attribut **class** pour inclure les **microformats**.

- **itemscope** : il permet de créer un item dont le type peut-être précisé grâce à l'attribut **itemtype**
- **itemtype** : couplé à **itemscope**, il accueille une URL vers la référence du schéma (e.g. : <http://schema.org/Review>) qui devra respecter les valeurs de ce schéma

Les microdatas II

- **itemid** : lorsque l'item possède une référence globale unique, comme par exemple un l'identifiant d'un livre (isbn). Il est attribué en même temps que l'attribut **itemscope**.
- **itemprop** : il est porteur d'une valeur permettant de préciser la nature d'une donnée au sein d'un schéma spécifié précédemment (e.g. : un code postal pour un schéma sur les adresses postal).
- **itemref** : il permet de faire un lien complémentaire entre deux données dispersées sur le document.

Les microdatas III

```
<p itemscope itemtype="http://schema.org/Restaurant  
  " itemref="restaurant-cruchon">
```

```
  Le restaurant <span itemprop="name">Le Cruchon  
    Repus</span> est clairement le meilleur  
    restaurant de Paris.
```

```
</p>
```