



## Franck CELLIER

Formateur en Informatique

Technicien Qualité d'Organisme de Formation

<http://cellierfranck.alwaysdata.net>



# MS LANGAGE & DEVELOPPEMENT POUR LE WEB

## LE HTML

### QU'EST-CE QUE LE WORLD WIDE WEB ?

Le World Wide Web (Web) est un réseau de ressources et d'informations. Le Web repose sur trois mécanismes pour rendre ces données facilement disponibles le plus largement possible :

un système de nommage uniforme pour leur localisation sur le Web (par exemple, les URI) ;

des protocoles, pour accéder à des ressources nommées dont on possède l'adresse (par exemple, HTTP) ;

l'hypertexte, pour faciliter la navigation entre ces ressources (par exemple, HTML).

Les liens entre ces trois mécanismes apparaissent clairement tout au long de cette spécification.

### QU'EST-CE QU'UNE URL ?

L'URL, pour Uniform Resource Locator, désigne communément l'adresse web d'un site Internet. Elle se présente sous la forme d'une série de caractères permettant l'identification d'un site Internet et/ou d'une page web.

Sur un plan technique, l'URL indique aux navigateurs (Internet Explorer, Mozilla, Firefox, Safari, etc.) le chemin à emprunter pour accéder aux ressources contenues dans le World Wide Web.

Dans un navigateur Internet classique, l'URL est la formule que l'on renseigne dans la barre d'adresse (ou barre de navigation) et qui permet d'accéder directement à un site ou à une page internet. Voici sa forme la plus classique : <http://www.exemplemonsupsite.fr>

## Importance de l'URL pour le référencement naturel

Les moteurs de recherche sont très attentifs à plusieurs points relatifs aux URL. Pour le référencement d'une page, l'URL doit être claire, unique et pas trop longue. Il est conseillé d'y faire figurer un mot-clé important pour la page. Pour référencer une page dans les actualités, il est nécessaire qu'une suite de chiffres ne ressemblant pas à une date apparaisse également dans l'URL.

L'URL sert aussi, lorsqu'on crée un lien, à pointer une page précise. L'URL doit alors figurer proprement et clairement dans le code HTML de la page. De nombreuses règles régissent le netlinking et guident le travail des référenceurs.

Elle contient plusieurs éléments qui donnent des renseignements qui, mis tous ensemble, forment une adresse unique pour chaque page du web :

- Le protocole de communication utilisé pour accéder au contenu, qui se présente ainsi : "https://" ou "ftp://" (indique une URL dédiée au transfert de documents. Il est de plus en plus rare de voir des sites avec une URL : "http://" car depuis 2018, Google Chrome affiche ces sites comme insuffisamment sécurisés.
- Un nom de domaine : www.monsite.com Il s'agit du cœur de l'adresse, permettant d'identifier le site.
- Un identifiant, prenant la forme d'une suite de chiffre. Cela est notamment vrai pour les sites de presse, mais ne concerne pas tous les sites
- Le dossier, dans le cas d'une page ftp : il permet d'accéder, dans le répertoire, à l'emplacement où se trouve le document ciblé.
- Le nom de la page : ex : /cet-article-est-super/ Au sein d'un site internet, chaque page a un nom composé soit de mots clés, soit des mots du titre de la page. La page d'accueil d'un site n'a pas cette extension.
- 

## QU'EST-CE QUE HTML ?

Publier de l'information en vue d'une distribution mondiale nécessite l'usage d'un langage universellement compréhensible, d'une sorte de « langue maternelle » acceptée par tous les ordinateurs. Le langage de publication utilisé sur le World Wide Web est le langage de balisage hypertexte HTML.

HTML donne aux auteurs les moyens de :

- Publier des documents en ligne possédant des titres, du texte, des tables, des listes, des photos, etc.
- Ramener des informations en ligne via des liens hypertextes avec un clic ;
- Concevoir des formulaires pour conduire des transactions avec des services distants, dans le but de rechercher des informations, de faire des réservations, de prendre des commandes, etc.
- Inclure des feuilles de calculs, des séquences vidéos ou sonores et d'autres applications, directement dans les documents.

## UNE BRÈVE HISTOIRE DE HTML

Le langage HTML a été développé à l'origine par Tim Berners-Lee, alors au CERN, et popularisé par le navigateur Mosaic, développé au NCSA. Pendant les années 1990, il s'est épanoui en profitant de la croissance explosive du Web. Durant cette période, HTML s'est agrandi de nombreuses manières différentes. Le Web dépend des auteurs de pages et des éditeurs de logiciels qui partagent les mêmes conventions pour HTML. Cela a motivé un travail commun sur les spécifications de HTML.

La spécification HTML 2.0 (novembre 1995, voir [RFC1866]) a été développée sous l'égide de l'Internet Engineering Task Force (IETF) pour codifier les pratiques courantes à la fin de 1994. Les spécifications HTML+ (1993) et HTML 3.0 (1995, voir [HTML30]) proposaient des versions HTML bien plus riches. Malgré l'absence de consensus autour d'un standard, ces avant-projets ont conduit à l'adoption d'un éventail de fonctions nouvelles. Les efforts du groupe de travail HTML du World Wide Web Consortium pour la codification des pratiques courantes en 1996 donnèrent naissance à la spécification HTML 3.2 (janvier 1997, voir [HTML32]). Les changements intervenus entre HTML 3.2 et HTML 4 sont résumés dans l'appendice A.

La plupart des gens sont d'accord sur le fait que les documents HTML devraient fonctionner correctement sur des plates-formes et des navigateurs différents. Parvenir à l'interopérabilité permet de baisser les coûts des fournisseurs de contenus parce qu'ils n'ont à développer qu'une seule version d'un document. Si cet effort n'est pas fait, le risque est grand de voir le Web régresser vers un monde propriétaire de formats incompatibles, en réduisant en fin de compte son potentiel commercial au détriment de tous ses participants.

Chaque version HTML a essayé de refléter le plus grand consensus entre les acteurs de l'industrie, de sorte que les investissements consentis par les fournisseurs de contenus ne soient pas gaspillés et que leurs documents ne deviennent en peu de temps illisibles.

Le langage HTML a été développé avec l'intuition que les appareils de toutes sortes devaient pouvoir utiliser les informations sur le Web : les ordinateurs personnels avec des écrans de résolution et de profondeur de couleurs variables, les téléphones cellulaires, les appareils portables, les appareils de synthèse et de reconnaissance de la parole, les ordinateurs avec une bande passante faible comme élevée, et ainsi de suite.