

Chapitre 02 : Le Web

Les élèves doivent savoir

- Définir les étapes du développement du Web.
- Maîtriser les renvois d'un texte à différents contenus.
- Distinguer ce qui relève du contenu d'une page et de son style de présentation.
- Étudier et modifier une page HTML simple.
- Décomposer l'URL d'une page.
- Reconnaître les pages sécurisées.
- Décomposer le contenu d'une requête HTTP et identifier les paramètres passés.
- Inspecter le code d'une page hébergée par un serveur et distinguer ce qui est exécuté par le client et par le serveur.
- Mener une analyse critique des résultats fournis par un moteur de recherche.
- Comprendre que toute requête laisse des traces.
- Maîtriser les réglages les plus importants concernant la gestion des cookies, la sécurité et la confidentialité d'un navigateur.

Chapitre 02 : Le Web

1 Introduction :

Le web est un ensemble de ressources numériques reliées et accessibles par des liens cliquables, sur internet.

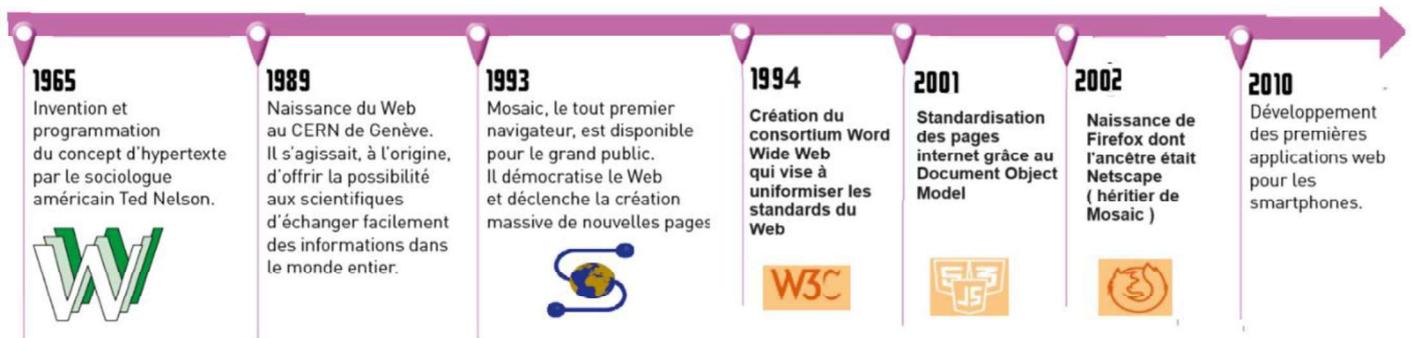
A partir des années 2000, le web « social » avec l'apparition de réseaux sociaux comme Facebook , Twitter, Youtube, ...etc, permet le partage en tout genre.

Importance économique :

- Le premier paiement en ligne a été effectué en 1994 (il s'agissait d'une transaction d'une dizaine de dollars pour acheter une chanson sur un CD audio).
- Plus de 700 milliards de paiements électroniques ont été effectués sur le Web en 2019.
- On estime à 1,4 milliards de sites Web ont été créés dans le monde depuis le début du Web.

2 Evolution du web

- 30 ans de développement ont été nécessaires au Web pour devenir un vaste réseau de pages web reliées entre elles par des liens hypertextes et consultables avec un navigateur web (voir frise chronologique ci-dessous) :



En 1989 **Tim Berners-Lee**, chercheur au CERN (centre européen de recherche nucléaire) à Genève, commence à travailler sur l'association entre hypertexte et internet en combinant hypertexte, TCP et DNS.

En 1990 il crée l'expression **World Wide Web** (www) et développe, avec Robert Cailliau* les adresses web (url), l'**HyperTexte Transfer Protocol** (HTTP) et l'hypertexte markup langage (html). Ils créent également un navigateur web, un éditeur et un serveur HTTP. L'ensemble est lancé en 1991.

3 Le fonctionnement du web

Le World Wide Web (Web ou "Toile à échelle mondiale ") est l'une des principales applications d'internet. L'internaute accède aux ressources en naviguant sur les pages Web les hébergeant.

3.1 Hypertexte :

3.1.1 Définition :

L'hypertexte est une liaison entre les textes par l'utilisation de mots clefs (comme les mots mis en évidence dans une page Wikipédia qui vous amènent vers une autre définition).

Le mot hypertexte a été créé en 1965 par Ted Nelson et au départ il servait à relier des documents entre eux sur la même machine.

C'est sur ce concept que s'est basé Tim Berners-Lee en l'étendant au réseau Internet naissant.

Exemple celui-ci est un [lien hypertext qui pointe sur la frise](#)

3.2 Le protocole de communication HTTP :

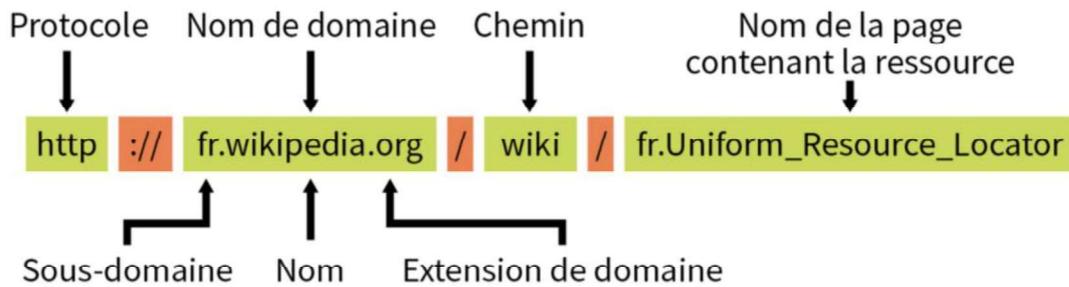
Pour s'afficher dans le navigateur, un document doit avoir fait l'objet d'une demande de la part de l'utilisateur (**une requête**) : il s'agit d'une demande auprès du serveur qui héberge ce document de le lui envoyer.

C'est cette relation client-serveur que désigne le **protocole http** (HyperText Transfert Protocol, "protocole de transfert hypertexte ").

Pour que cela soit possible, il faut que le document en question puisse être localisé sur le Web par une adresse unique appelée **URL** ou "**Uniform Ressource Locator** ", expression anglaise pouvant se traduire par " localisateur universelle de ressources ",

- L'URL est donc une adresse d'une ressource sur le Web comme une page Web, une image, une vidéo ou un fichier.
- L'URL commence par **http** ou **https** (si la page est sécurisée) et se décompose en un domaine principal et plusieurs sous-domaines.

Exemples d'URL : https://fr.wikipedia.org/wiki/Uniform_Resource_Locator



3.3 Les noms de domaines et les extensions :

Les extensions de domaine ou noms de domaines de premier niveau se divisent en deux catégories

- Celles associées à un pays ou un territoire *.fr* , *.uk* , *.ma* ...etc
- Celles associées à des activités : "*.edu*" (pour éducation) , "*.com*" (pour le commerce), "*.org*" pour les association et organisation diverses

Un utilisateur particulier ou une entreprise est libre de choisir le nom de domaine associé à aux pages Web qu'il crée, à condition qu'il soit disponible.

Pour louer un nom de domaine, on doit passer par des distributeurs agréés par **PICANN (Internet Corporation for Assigned Names and Numbers)** : C'est l'autorité de régulation d'internet.

Pour qu'une adresse soit accessible via son nom de domaine, ce dernier doit être enregistré dans deux DNS au minimum. Le DNS se charge de faire le lien entre le Web et l'internet.

4 Moteurs de recherche

Le moteur de recherche est une invention qui a vraiment révolutionné le web : il permet de retrouver une information précise en peu de temps. Encore faut-il savoir ses limites et comment en tirer le meilleur parti.

4.1 Une énorme base de données

Un moteur de recherche est un serveur qui dispose de logiciels fonctionnant en simultanées, nommés «**robots d'indexation**» (les *crawlers* ou les *spiders* en anglais) dont le rôle est d'explorer tous les sites internet à intervalle de temps régulier et à stocker les mots clefs découverts dans une base de données.

4.2 Effectuer une recherche

La recherche dans une base de données d'un moteur de recherche se fait sur un ensemble de mots clefs qui doivent être saisis dans un champ texte. Le moteur va ensuite donner des réponses pertinentes (ou pas) si les mots clefs sont bien saisis (sans faute) en fonction d'un algorithme interne, généralement secret, qui va mettre en avant certains sites..

4.3 Réponse d'un moteur de recherche

Il convient d'être très attentif aux réponses données par un moteur de recherche. Certains sites paye le moteur de recherche pour apparaître en premier dans les recherches et indique donc à coup sûr que celui-ci cherche à se rémunérer sur ses visiteurs.

5 Sécurité sur le Web :

Lorsque vous surfez sur le Web ou vous faites des recherches, les navigateurs enregistrent chronologiquement votre activité : Sites visités, cookies, préférences, ...etc.

Toutes ces informations sauvegardées dans un fichier, composent l'historique. L'intérêt premier est de simplifier la navigation de l'utilisateur : retrouver un site qui a été consulté, garder en mémoire les mots de passe des sites visités , fréquemment, préremplir les formulaires avec ses informations personnelles (coordonnées, numéro de carte bancaire...etc.,

- Les cookies contenant toutes sortes d'informations sont stockées et consultables sans que vous ne soyez au courant.*
- Les sites que vous visitez sont conservés dans un historique de navigation.*
- Vous pouvez stocker des mots de passe pour ne pas avoir à les retaper, parfois même des numéros de carte de crédit.*

Ces données restent cependant accessibles aux sites connectés, ils peuvent donc potentiellement être volées par un site web indélicat à la faveur d'une erreur de programmation du navigateur (faille de sécurité).

La navigation privée ne rend pas invisible

5.1 Utilisation des documents trouvés sur Internet

Une fois que vous avez trouvé une information sur Internet par le biais d'un moteur de recherche ou par un accès direct, respectez les droits d'utilisation de ces informations.

Vous ne pouvez réutiliser une information puisée sur Internet que si l'auteur vous en donne le droit. Il est généralement possible de cibler précisément les informations libres de droits sur votre requête de recherche.

Pour en savoir plus sur vos droits, vous pouvez consulter les documents de la Commission Nationale Informatique et Liberté, la CNIL (<https://www.cnil.fr/>) et ceux du règlement général sur la protection des données, RGPD (<https://www.cnil.fr/fr/rgpd-par-ou-commencer>).

6 Le langage et le développement des pages Web.

L'accès au Web est rendu possible par l'existence des liens hypertexte entre les pages du Web et au protocole de communication HTTP. Mais pour être échangées à travers le monde, les pages doivent aussi utiliser des langages standardisés.

6.1 HTML et CSS

6.1.1 4.1.1 La base du HTML

Définitions :

- Le HTML est un langage constitué de balises qui permettent la mise en page d'un document.
- HTML signifie "**HyperText Markup Language**", c'est un langage de description de textes liée avec **des balises**.
- Une page HTML est un code source qui sera interprété par le navigateur web. C'est un simple texte.

Exemples de codage html :

Il est structuré en arborescence avec deux blocs principaux : **en-tête et corps**

```
1  <!DOCTYPE html>
2  <html lang="fr">
3  <head>
4      <meta charset="UTF-8">
5      <title>Programme SNT</title>
6  </head>
7  <body>
8      <h1> Bien venu sur le site Web de la 2GT01_SNT </h1>
9  </body>
10 </html>
11
```

6.1.2 Les balises HTML

Une balise HTML respecte une syntaxe simple et stricte :

- Le symbole ouvrant : <
- Le nom de la balise : Titre
- Des attributs optionnels : un espace, suivi du nom de l'attribut, d'un signe {=} et d'une valeur entre guillemets {" "}
- Le symbole fermant : >

A chaque balise ouvrante : < **exemple** > doit correspondre une balise fermant : </ **exemple**>
Il existe néanmoins des balises auto-fermantes telles que : <**img** />
