

Le WEB :

la Toile(=web), est un **système hypertexte** public fonctionnant sur Internet. Le Web permet de consulter, avec un navigateur, des pages accessibles sur des sites. L'image de la toile d'araignée vient des **hyperliens qui lient les pages web entre elles**.

Le Web **n'est qu'une des applications d'Internet**, distincte d'autres applications comme le courrier électronique, la messagerie instantanée et le partage de fichiers en pair à pair.

Histo :

1989 : Tim Berners Lee invente dle World Wide Web(système d'information basé sur l'hypertexte)

1990 : définition de HTTP et HTML

1993 : premier navigateur

1994 : fondation par Tim Berners Lee du wwwconsortium (W3C) pour homogénéiser les protocoles.

1995 et 98 : fondation d'Amazon et de Google : les deux premiers GAFAM

2010 : les sites deviennent responsive(implique un minimum de redimensionnement (zoom), de recadrage, et de défilements multidirectionnels de pages : s'adapte à tous les appareil de lecture et de navigation.)

Chiffres : 4,3 milliards d'internautes, 1,4 milliards de sites internet, 92 % des entreprises ont un site internet, 85 % des internautes ont déjà fait un achat sur internet.

1- principe du WEB

1-A- Architecture client /serveur

Le WEB(ou toile) fonctionne grâce à un échange de données entre clients et serveurs : les clients demandent des données, les serveurs les mettent à disposition. L'utilisation de cette architecture est permise par les applications/ logiciels : les navigateurs internet.

Serveur : ensemble d'ordinateurs interconnectés qui abritent des données réparties à travers le monde. Ils mettent à disposition ces données.

Client : un navigateur web, une interface graphique, permet d'accéder aux données/ informations des serveurs.

Le navigateur est en charge de :

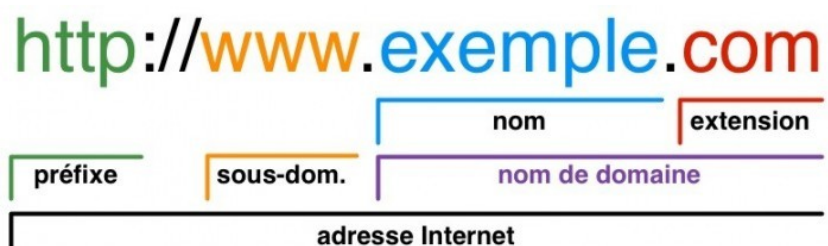
- la demande de la page souhaitée par le client, associée à un URL
- de l'affichage de la page conformément à sa création, aux dimensions de l'écran.

1-B- L'URL

Structure d'un URL (Uniform Resource Locator) :

quatre parties significatives, présentées ici de la droite vers la gauche :

1. L'**extension** désigne soit un pays sous la forme d'un code à deux lettres, soit un domaine d'activité générique : com, org, net, edu, info, mobi, etc. Appelé aussi Top Level Domain(TLD)



2. Le **nom**, souvent appelé nom de domaine par simplification excessive, est choisi pour identifier et comprendre le contenu du site web associé. La rédaction du nom répond à des règles particulières.
3. Le **sous-domaine** sert à dédier certaines zones du domaine à des usages spécifiques.
4. Le **préfixe** indique la nature de la communication(protocol) établie par l'internaute avec le domaine, par exemple http est le préfixe qui correspond au web, ftp au transfert de fichiers, etc.

1, 2 et 3 donne le nom de domaine entièrement qualifié :

www.google.fr → Nom de domaine entièrement qualifié(associé à un IP...)

exemple :

<https://www.cairn.info>

→ ou <https://185.34.32.93>

Dans ce domaine, on peut accéder à une arborescence de dossiers

/exemple/ → arborescence du dossier dans lequel se trouvent les fichiers ressources de la page.

<https://www.cairn.info/revue-etudes-de-communication-2018-2.htm>

1-C- Le protocole HTTP(Hypertext Transfer Protocol)

Une adresse Internet est toujours constituée de la façon suivante :

protocole://adresse-du-serveur:port/chemin/ressource

Rq : par défaut, le port est rarement affiché, il est de valeur 80...

Pour que 2 personnes se comprennent, il faut qu'elle parle la même langue. Et bien en informatique, c'est pareil : il faut que le navigateur sache dans quelle langue il doit parler au serveur. Cette langue est appelée protocole.

L'acronyme HTTP signifie "Hypertext Transfert Protocol". Il a été inventé par Tim-Berner Lee dans les années 1990. C'est le protocole le plus utilisé sur Internet. Le but du protocole HTTP est de permettre un transfert de fichiers (essentiellement au format HTML) localisés grâce à une chaîne de caractères appelée URL entre un navigateur (le client) et un serveur Web. Ce protocole est un protocole de communication client-serveur et fonctionne sur le principe "requête-réponse"

C'est le protocole de base du web, il est chargé des requêtes d'affichage des pages web.

Ce protocole utilise un protocole TCP pour faire transiter les données des pages web entre le client (navigateur) et le serveur.

Déroulé du protocole en bref:

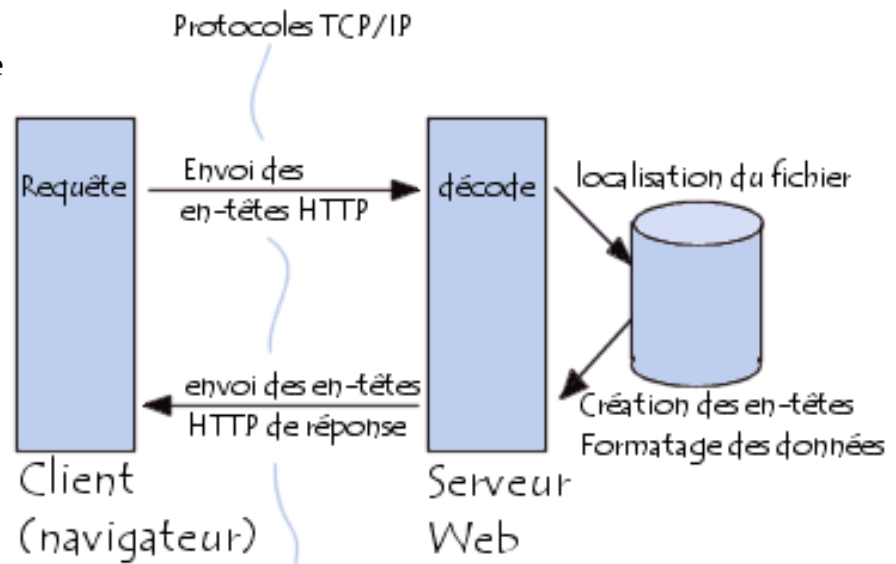
0- Le client « demande » éventuellement au préalable d'adresse IP auprès d'un serveur DNS(Aller retour d'infos via le serveur FAI)

1- le client envoie une requête au serveur web via le réseau internet.

2- Le serveur web envoie la confirmation de l'existence de la page web demandée et établit une connexion TCP

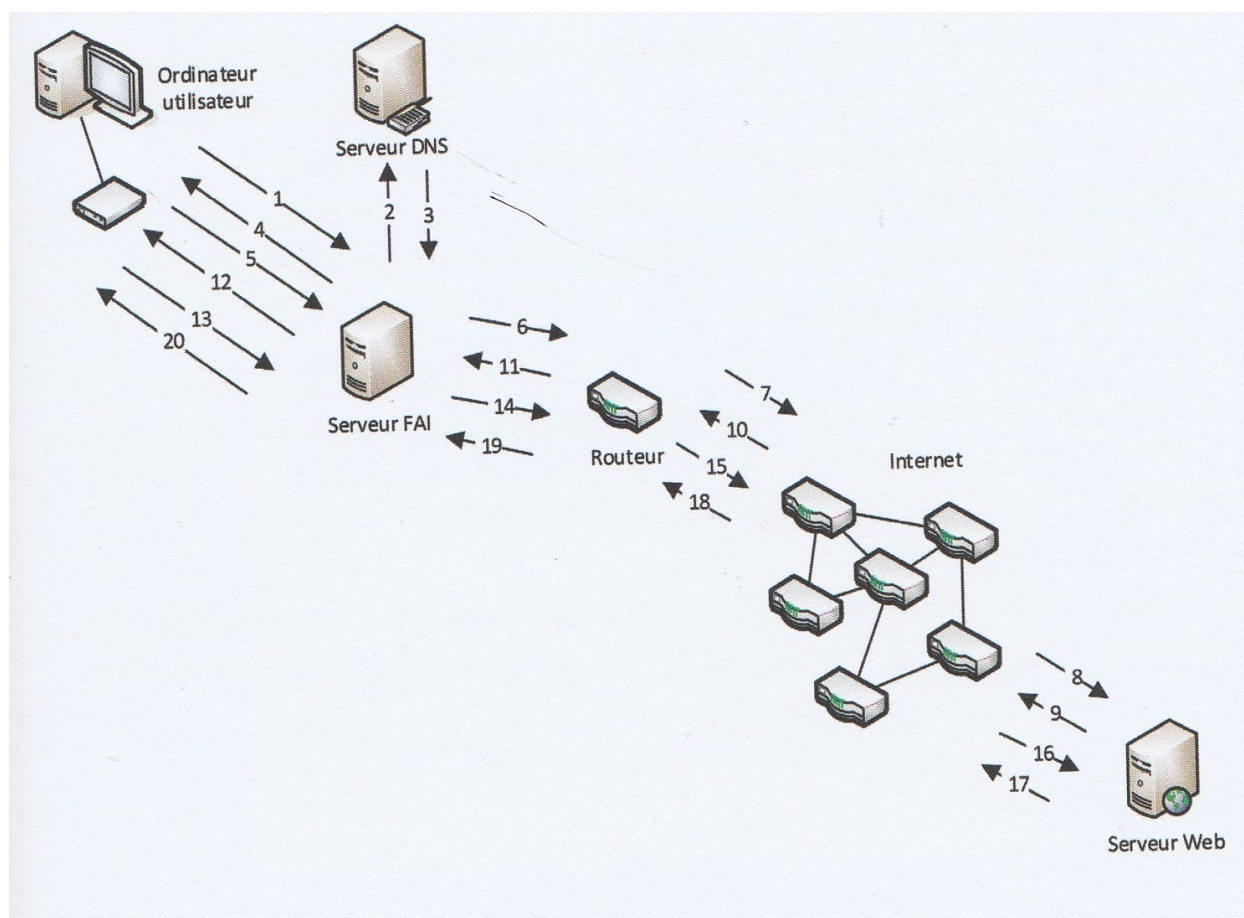
3- échange de demandes par le client sous forme de paquets TCP et d'envois de paquets TCP contenant des infos par le serveur web afin d'envoyer tous les éléments de la page demandée.

4- reconstruction de la page web et affichage par le navigateur du client.



Activités :

1- Légender le document suivant : Vous pouvez regrouper les légendes en groupes de légende de même fonction...



FAI= fournisseur accès internet.

1 à 4 : demande de l'adresse ip du serveur web abritant le site désigné par l'URL

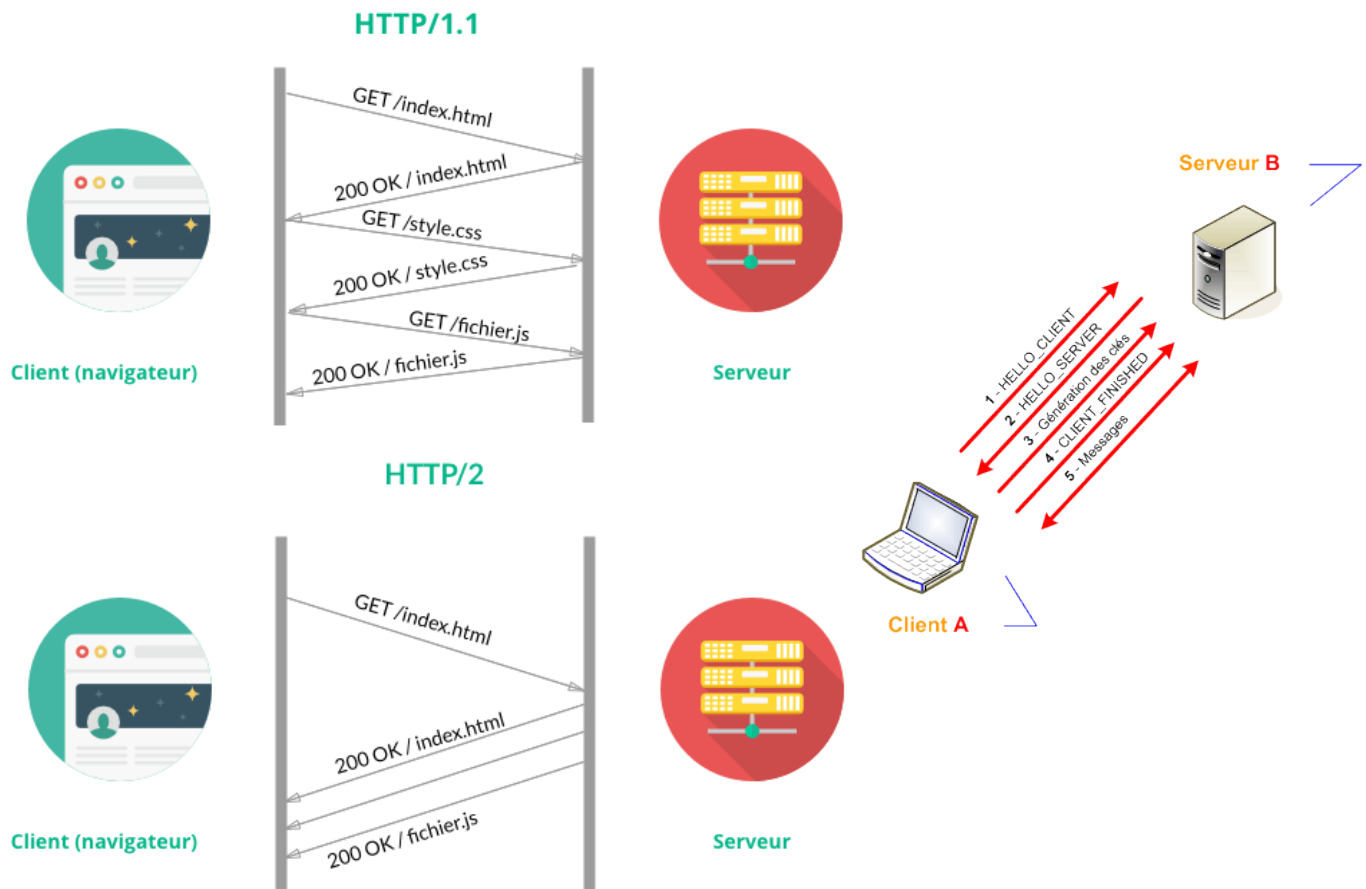
→ le client sait que le site correspondant à l'adresse désignée par l'URL existe, a obtenu l'adresse ip du serveur. Il faut à présent créer la connexion avec lui.

500 et 503 : erreur serveur ; 504 : le serveur n'a pas répondu.

HTTP peut être associé à un protocole de sécurisation SSL pour donner la version https qui garantit la sécurité des données renseignées sur la page web(indispensable lors des paiements....).

protocole(s) HTTP :

et en HTTPS :



1-D- Choix du navigateur et du moteur de recherche :

De nombreux navigateurs permettent l’affichage des pages web mais ils tous leurs spécificités :

1- Google Chrome :

Les plus

- + Performances
- + Simple et agréable
- + Bien sécurisé

Les moins

- Il reste l’un des outils de récolte de données de Google

Dix ans après sa création, le logiciel est toujours à la pointe de la technologie grâce notamment au rythme effréné des mises à jour / À force d’enrichir ses fonctionnalités, **Google Chrome** n’est

Le meilleur en 2019

Le plus accessible



Google Chrome

- + Stable et fiable
- + Versatile
- + Intégration de l’écosystème Google



Mozilla Firefox

- + Fonctionnel et efficace
- + Respect de la vie privée...
- + ...et des standards du web

plus aussi léger et véloce qu'à ses débuts cependant **Chrome coiffe en effet tous ses challengers sur le poteau en termes de gestion des nouveaux standards du web. Capable de gérer sans sourciller un très nombre d'onglets ouverts simultanément, Chrome est sans aucun doute le navigateur le plus stable du marché.**

2. Mozilla Firefox

Les plus

- + Un des plus innovants du marché
- + Riche en fonctionnalités pour optimiser l'UX
- + Gère les derniers standards vidéo AV1

Les moins

- Scores un peu décevants en benchmark
- Nécessite une configuration moyenne ou haut de gamme

Firefox fût un temps plus rapide que Chrome tout en consommant près d'un tiers de moins de mémoire grâce notamment à une nouvelle architecture multiprocesseur. **Mozilla est une fondation à but non lucratif qui n'exploite pas les données de ses utilisateurs. Trois niveaux de confidentialité** sont proposés - Standard, Strict, et Personnalisé - pour bloquer les traqueurs et les cookies. A noter qu'il **bloque désormais par défaut la lecture automatique des contenus audio/vidéo et les pop-ups.**

3. Microsoft Edge

Les plus

- + Performances correctes
- + Plus léger, rapide et moderne
- + Capables de tourner sur terminaux mobiles

Les moins

- Interface vieillissante
- Délaissée progressivement par Microsoft

4. Opera :

Les plus

- + Stable et performant
- + Débits rapides
- + Sécurité accrue avec outils de chiffrement efficaces

Les moins

- S'adresse aux utilisateurs avancés
- Mode turbo que sur les domaines http

5. Brave :

Les plus

- + Respecte la vie privée
- + Interface et ergonomie soignée
- + Rapide

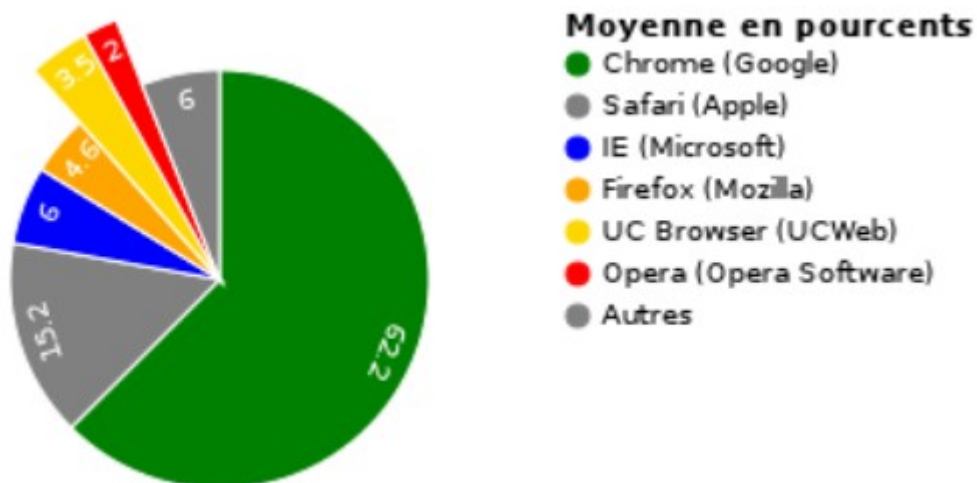
Les moins

- Fonctions de synchronisation perfectibles
- Encore quelques bugs pour devenir N°1

....

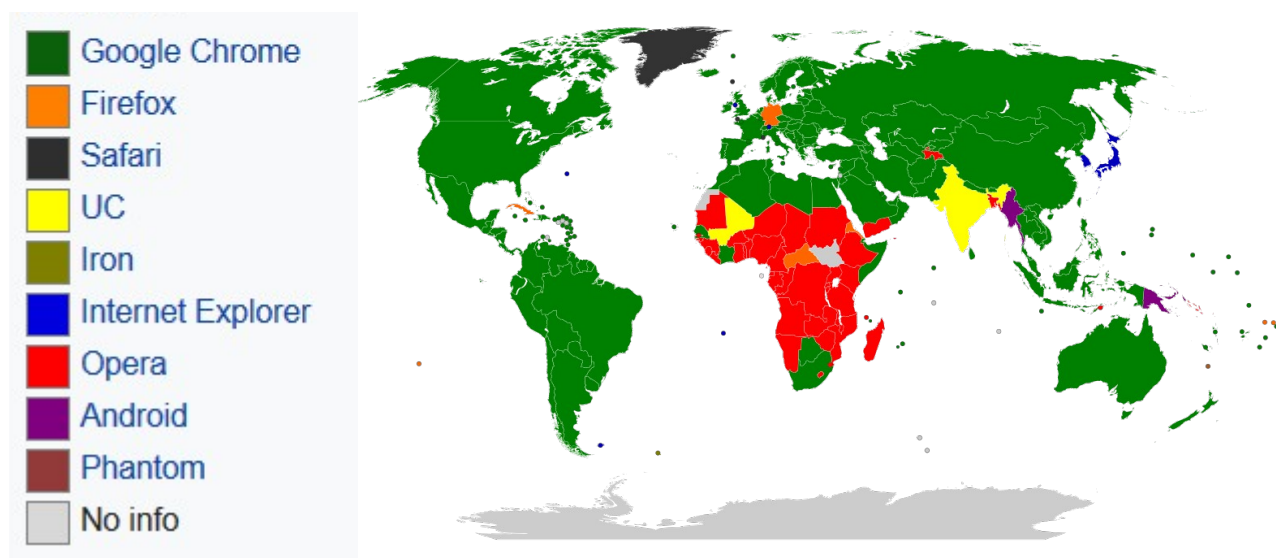
6. Safari : par défaut sur les OS apple.

Rapide, stable et bonne protections des données.



Source :

<https://www.clubic.com/article-282232-1-comparatif-internet-explorer-firefox-opera-safari.html>
https://fr.wikipedia.org/wiki/Navigateur_web



Il existe aussi de nombreux moteurs de recherche. Ces moteurs possèdent des interfaces web où l'entrée de quelques mots clés permettent d'afficher de nombreux sites référencés. Afin de créer les relations nécessaires pour afficher des pages web pertinentes :

- des robots (petits programmes) (= crawlers ou spiders) parcourent et scannent tout le contenu du net en sautant d'hyperliens en hyperliens et de temps en temps en se plaçant sur le réseau de façon aléatoire pour éviter de se retrouver coincé dans un recoin... Ils permettent d'inventorier les mots clés de chaque page et d'actualiser l'index...

- un index en perpétuel changement répertorie et classe les URL en fonction de leur pertinence par rapport aux mots clés de la recherche.

Le résultat de recherche prendra la forme d'une liste hiérarchisée de page en lien avec la recherche.

Parmi les moteurs de recherche :

stocke vos recherches associées à votre IP : collecte les données personnelles

Google : Google scanne 20 milliards de sites internet par jour. **Et chaque jour, 500 millions de nouvelles requêtes** sa part de marché est de 94% spécialisé dans la recherche d'images avec son moteur spécialisé **Google image**. Il est même possible de faire des recherches par images. utilisez l'option « Recherche par images »

Bing : Souvent intégré directement au sein des systèmes informatiques, Bing tire son épingle du jeu en étant le deuxième moteur utilisé en France.

Yahoo : C'est plus de 8 millions d'utilisateurs uniques qui se sont connectés à ce moteur de recherche au cours du mois de Septembre, contre 5,1 millions l'année précédente.

NE stocke PAS vos recherches associées à votre IP : PAS de collecte les données personnelles

Qwant : Ce moteur de recherche croise les résultats des autres moteurs et des réseaux sociaux, sans stocker de cookies et sans vous « pister ».

Duck Duck GO : ses informations dans les API des moteurs Yahoo, Wikipédia, Wolfram, Bing et Blekko. Et bon point pour votre vie privée : il n'enregistre pas votre IP et les informations personnelles ! C'est le même positionnement que Qwant

→ **mais il y en a des dizaines d'autres, plus ou moins spécialisés dans des domaines précis....**

source : https://fr.wikipedia.org/wiki/Liste_de_moteurs_de_recherche

Activité : faire une recherche simultanée grâce à dukduckgo, à google et à Ecosia avec les mêmes mots clés (ex : vente chien paris).

Que remarquez vous ?

Rechercher quelle la particularité d'Ecosia.

→ **pas les mêmes résultats. Google affiche en premier des sites commerciaux.**

→ **Ecosia plante des arbres... Mais collecte les données personnelles.**

1-E- Impact sur les pratiques humaines.

Le web a pris une place prépondérante en 25 ans d'existence :

- source d'information n°1/ partage des savoirs.
- le **commerce** y est devenu très conséquent.
- Permet des **projets collaboratifs** internationaux.
- développement des **loisirs** (jeux, vidéos, musiques, arts créatifs...)
- outils de **communication** et base des **réseaux sociaux**....

Il faut donc prendre du recul : L'info est elle fiable ? Objective ?/ Mon achat est il influencé par des publicités très ciblées ? La sécurité de paiement est elle garantie ? Comment sont gérées mes données personnelles ?

Quel réseaux social choisir ? Comment protéger mes informations et les laisser disponibles à certaines personnes ?

2- Langages pour créer des pages web.

2-A- Le langage HTML

T04-02 1/a/ le langage HTML : découverte(1heure)

Une page web est codée au minimum grâce à un fichier html.
Celui-ci est codée grâce à un **enchaînement de balises** : **<balise>**

<html> → pour ouvrir le doc html !

<head> → pour ouvrir l'entête d'information

<title> Titre de la page </title> → donne le titre de la page

<meta charset="utf-8"> → donne le codage des données de caractère

<h1> → ouvre une zone d'affichage de titre en format préétabli h1(les plus gros titres)

Titre de la page --> le titre de la page

</h1> → ferme une zone d'affichage

</head> → pour fermer l'entête d'information type h1

<body> → ouvre le corps de la page

<h2> → ouvre une zone d'affichage en format préétabli h2(les un peu moins gros titres)

Partie 1 : le savoir vivre → texte de titre format h2

</h2> → ferme une zone d'affichage

<p> → ouvre une zone de texte à afficher

texte, blablablala,</br> → envoie à la ligne **le savoir vivre c'est aussi ne pas parler de politique à table...**

</p> → ferme une zone de texte à afficher

</body> → ferme le corps de la page

</html> → pour fermer le doc html !

1, 2 voir 3 langages et plus..

HTML, CSS et Javascript...

Une page html peut contenir aussi des liens vers d'autres pages(sinon elle se trouve en bout de réseau et on ne peut pas naviguer à partir d'elle... elle est donc défectueuse!!), vers d'autres documents : **les hypertextes : la balise est**

texte affiché hyperlien

ex :

clique sur moi pour aller vers le moteur de recherche canardeux

Une page web contient aussi souvent **des images**, la balise d'insertion est **** l'adresse peut être :

- une **adresse d'une image stockée sur un serveur**(on peut copier le lien image de n'importe quel page web autre)

- l'**adresse sur un disque** de votre ordinateur avec pour racine le dossier dans lequel se trouve le fichier html.

Ex : ****

Du coup on peut mettre une image en tant que lien hypertext :

`<a href= « https://duckduckgo.com/»></a>`

Activé 04-02 1/b le langage HTML : action(1heure)

2-B- Le langage CSS :

On peut mettre les indications dans les balises : ici une balise `<p>` normale

`<pstyle="color: blue ; font-size: 10 px"> Texte bleu de taille 10px </p>`

Le style s'applique alors au paragraphe p

On peut mettre les indications de mise en forme dans la partie head du fichier html :

`<style>`

`pstyle=" color:blue ; font-size: 10 px"> Texte bleu de taille 10px`

`</style>`

Le style s'applique alors par défaut à la page entière

Pour un site entier, afin de ne pas se répéter pour toutes les pages, garder des styles uniformes et modifier les styles de toutes les pages en même temps : c'est plus simple de séparer la mise en forme afin d'attribuer des formes à plein de pages et de parties de pages en même temps...

Pour cela, on crée un fichier css(ou plusieurs) et on le référence dans l'en tête head de la page web. Pour chaque partie de la page, on indiquera le style voulu qui fait référence aux informations du fichier CSS.

Exemple :

Dans le fichier html:

`<head>`

`<link rel="stylesheet" href="style1.css " />`

`</head>`

ET

au sein du body :

`<divclass=" autreStyle ">`

`<p>`

texte d'un autre style que le reste du body(couleur rouge des caractères)

`</p>`

`</div>`

fichier style1.css

body {

padding-left: 11em;

font-family: Georgia, "Times New Roman" ;

Times, serif;

→ caractéristiques par défaut du body

color: purple;

background-color: #d8da3d ; }

autreStyle {

color : red ; } → mais on peut utiliser des attributs différents

```
a:link {  
  color: blue }
```

→ liens hypertextes en bleus

```
a:visited {  
  color: purple }
```

→ liens hypertextes en violet si consultés