

Le Web

I/ Le fonctionnement du Web

Lorsque l'on navigue sur le **Web**, notre poste informatique, tablette ou smartphone est identifié comme un **client** qui se connecte à un **serveur**. Ce serveur, qui héberge les pages Web ou les applications demandées, renvoie les contenus (textes, images, fichiers, etc.) au client. Chaque ressource fait l'objet d'une demande appelée « requête **HTTP** » par le client et d'une réponse par le serveur.

Pour identifier une page recherchée, le serveur utilise son **URL** qui se décompose en 3 parties : le protocole HTTP, le nom de domaine et le chemin vers la ressource.

a) URL

Une **URL** (Uniform Resource Locator) est l'adresse d'une page Web.

Elle est composée de trois éléments :

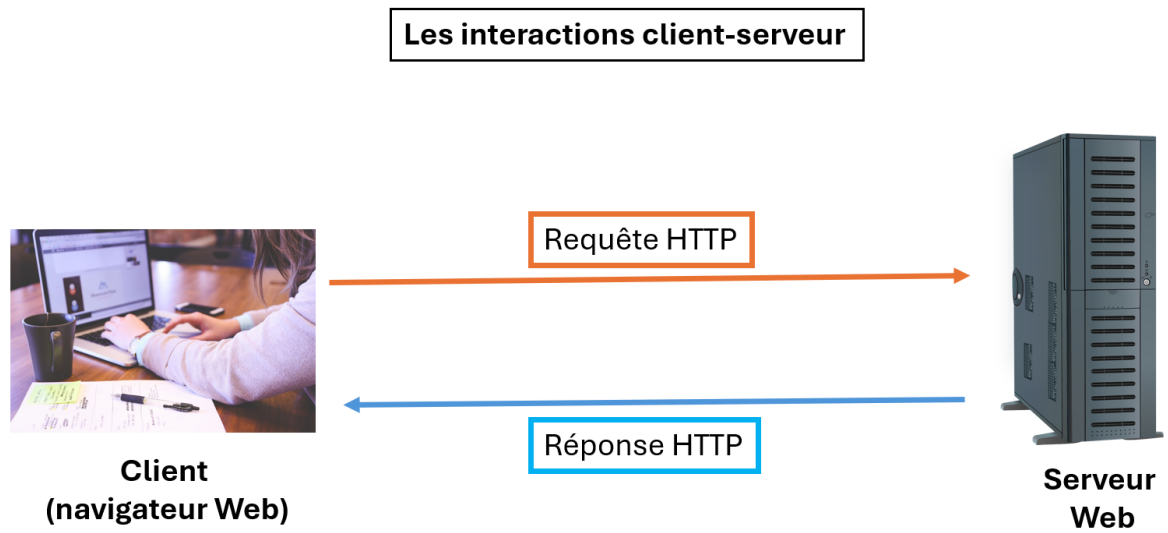
- `http://` ou `https://` qui correspond au protocole de communication client-serveur développé pour le Web ;
- Un nom de domaine, souvent une marque, une entreprise, une association, etc...
- Un chemin qui pointe vers une ressource ou page précise.

Décomposition d'une URL



b) Le modèle client-serveur et les requêtes HTTP

Pour consulter une page Web, un poste **client** envoie une requête à un **serveur**. Celui-ci lui retourne les ressources (textes, images, etc.) sous forme de code mis en forme par le navigateur. Les requêtes sont basées sur le protocole **HTTP** (HyperText Transfer Protocol) qui permet la communication entre les postes clients et les postes serveurs.



II/ Les langages HTML et CSS

Une page Web est constituée de deux langages : le **HTML** pour le contenu et une ou plusieurs feuilles de style **CSS** pour la mise en forme (arrière-plan, type et taille des polices de caractères, couleurs, etc.). Ces deux langages sont basés sur l'utilisation de balises identifient les différentes instructions.

Les **liens hypertextes** peuvent pointer dans la même page HTML, vers d'autres pages du site ou vers d'autres sites Web.

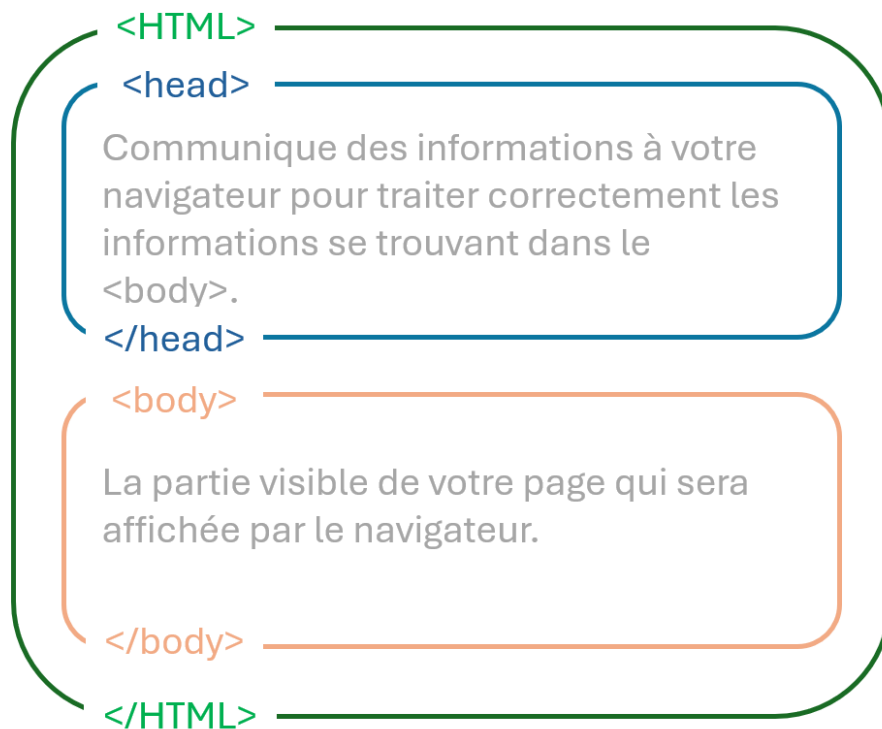
a) HTML

La programmation d'une page Web repose sur l'utilisation d'un langage de description appelé « **HTML** » (*HyperText Markup Language*). Le langage HTML utilise des balises pour délimiter le début et la fin de chaque instruction lue par le navigateur. Chaque instruction HTML est délimitée par une balise ouvrante `<élément>` et une balise fermante `</élément>`. Ces balises sont invisibles à l'écran mais elles permettent au navigateur d'interpréter ce qu'il doit afficher.

Composition d'une page web

`<!DOCTYPE>`

Renseigne le navigateur sur le langage utilisé par votre page Web



Quelques balises HTML

`<p></p>` Paragraphe

`<h1></h1>` Titre de niveau 1

`<h2></h2>` Titre de niveau 2

`` Image

`Ce qui est affiché à l'écran ` Lien hypertextes

Exemple

`https://www.facebook.com/chemin/vers/la/ressource/`

b) CSS

Le langage **CSS** (Cascading Style Sheets) permet de définir le style graphique d'une page Web (arrière-plan, type et taille des polices de caractères, etc.). La modification du style d'un élément dans une feuille de style se répercute automatiquement à toutes les pages HTML qui y font appel.

c) Les liens hypertextes

Le lien **hypertexte** est sans doute la balise la plus emblématique : elle permet un renvoi à un autre endroit d'une page Web, à une autre page du site ou à un autre site. Elle est identifiée par la balise `<a = href>` et se termine par ``. Un lien est composé principalement d'une URL cible et d'un libellé (le texte cliquable souvent de couleur, ou une image).

III/ Les moteurs de recherche

Les **moteurs de recherche** parcourent le Web, et indexent ses pages au moyen de robots **d'indexation** (crawler) et les classent au moyen d'algorithmes.

Le référencement naturel ou SEO (*Search Engine Optimization*) permet d'améliorer le classement d'un site Web dans les résultats de recherche grâce notamment à des mots-clés.

a) Les critères de classement des résultats

Les **moteurs de recherche** parcourent le Web, indexent les pages au moyen de robots (crawler) et classent les pages grâce à des algorithmes qui analysent les textes, mots-clés, balises, popularité de la page, liens entrants, etc.

Le fonctionnement des moteurs de recherche :

1. Exploration du Web (Crawling)
2. Indexation
3. Traitement des requêtes
4. Classement des résultats (Ranking)
5. Affichage des résultats

b) Le référencement naturel

Le **référencement naturel** ou SEO (*Search Engine Optimization*) permet d'améliorer le classement d'un site Web dans les résultats de recherche grâce notamment à des mots-clés placés dans le contenu rédactionnel. Le référencement naturel est à distinguer du référencement payant des annonces et les liens sponsorisés.

IV/ Sécurité et confidentialité sur le web

Pour éviter de laisser des traces, il est possible de régler les paramètres de contrôle de son **navigateur** (cookies, historique, sécurité, confidentialité).

Les sites basés sur le protocole **HTTPS** sont sécurisés. Les ressources publiées sur le Web (textes, photos, images, vidéo) ne peuvent pas être reprises sans l'autorisation de leur auteur.

a) Paramétrer son navigateur

Pour surfer sur le Web en toute sécurité et confidentialité, il faut éliminer ses traces en paramétrant son **navigateur**. Il est possible d'effacer son historique à partir des options du navigateur ou par la combinaison des touches CTRL+SHIFT+SUPPR.

b) Les sites sécurisés

HTTPS (HyperText Transfert Protocol Secure ou protocole de transfert hypertexte sécurisé) est un protocole de communication qui protège l'intégralité ainsi que la confidentialité des données lors du transfert d'informations entre l'ordinateur de l'internaute et le serveur. Le protocole HTTPS est reconnaissable par un cadenas dans la barre d'adresse du navigateur.

c) Notions juridiques

Les ressources publiées sur le Web (textes, photos, images, vidéos) sont la propriété de leur auteur et ne sont pas libres de droits. Leur usage est réglementé par différentes législations en fonction des pays ou encore par des régimes spécifiques dont peuvent se réclamer les auteurs, comme les licences *Creative Commons*.

Vocabulaire

Client : logiciel envoyant des demandes à un serveur.

CSS (*Cascading Style Sheets*) : feuille de style décrivant la présentation des documents HTML.

HTML : (*HyperText Markup Language*) : langage de balisage utilisé pour la création de pages web

HTTP ou HTTPS : protocole de transmission permettant d'accéder à des pages web via un navigateur. Le HTTPS est sécurisé.

Indexation : analyse du contenu des pages Web pour détecter les mots-clés.

Lien hypertexte : élément d'une page Web (texte, image, etc.) qui, lorsque l'on clique dessus, renvoie vers une autre zone de la même page, une autre page ou un autre site Web.

Moteur de recherche : application informatique permettant de rechercher une ressource (pages Web, images, vidéos, fichiers, etc.) à partir d'une requête sous forme de mots.

Navigateur : logiciel permettant d'afficher des sites Internet (Chrome, Firefox, Edge, Safari, etc.).

Serveur : ordinateur exécutant des opérations suivant les requêtes effectuées par un ordinateur appelé « client ».

URL (*Uniform Resource Locator*) : adresse d'une page d'un site.

Web (*World Wide Web*) : système hypertexte, utilisant le protocole HTTP, permettant de visiter des pages sur Internet.