

Contenus	Capacités attendues
Repères historiques	Connaître les étapes du développement du Web.
Notions juridiques	Connaître certaines notions juridiques (licence, droit d'auteur, droit d'usage, valeur d'un bien).
Hypertexte	Maîtriser les renvois d'un texte à différents contenus.
Langages HTML et CSS	Distinguer ce qui relève du contenu d'une page et de son style de présentation. Étudier et modifier une page HTML simple.
URL	Décomposer l'URL d'une page. Reconnaître les pages sécurisées.
Requête HTTP	Décomposer le contenu d'une requête HTTP et identifier les paramètres passés.
Modèle client/serveur	Inspecter le code d'une page hébergée par un serveur et distinguer ce qui est exécuté par le client et par le serveur.
Moteurs de recherche : principes et usages	Mener une analyse critique des résultats fournis par un moteur de recherche. Comprendre les enjeux de la publication d'informations.
Paramètres de sécurité d'un navigateur	Maîtriser les réglages les plus importants concernant la gestion des cookies, la sécurité et la confidentialité d'un navigateur. Sécuriser sa navigation en ligne et analyser les pages et fichiers.

Table des matières

1 L'architecture du Web	2
1.1 Qu'est-ce-que le Web ?	2
1.2 Le protocole de communication HTTP	2
1.3 Les URL	2
2 Les langages HTML et CSS	3
2.1 HTML	3
2.2 CSS	3
2.3 Les liens Hypertextes	3
3 Les moteurs de recherche	3
4 Le Web et les données personnelles	3
5 Le Web et le droit d'auteur	4
6 Repères historiques	4

1 L'architecture du Web

1.1 Qu'est-ce-que le Web ?

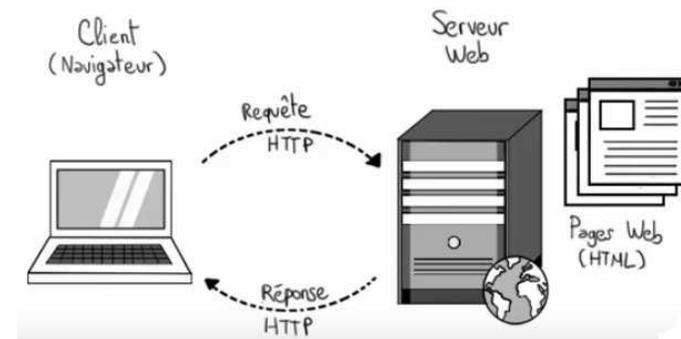
Le **Web** désigne un ensemble d'informations constitué par des milliards de documents reliés les uns aux autres et accessibles sur Internet, formant ainsi une gigantesque « toile d'araignée » mondiale : le World Wide Web.

Pour accéder au Web, on utilise un **navigateur** (comme par exemple Firefox ou Google Chrome).

Sur une page Web, les documents sont reliés entre-eux par des liens **hypertextes**. Ces liens sont soit des images, soit des morceaux de texte cliquables, que les concepteurs signalent, par convention, en les soulignant et en les bleuant.

1.2 Le protocole de communication HTTP

Pour consulter une page Web, un **client** envoie une **requête** à un **serveur**. Celui-ci retourne les ressources (textes, images,...) sous forme de code mis en forme par le navigateur.
Le protocole de communication entre un client et un serveur est le protocole HTTP (*Hypertext Transfer Protocol*).



Remarque. – Le protocole HTTPS est la version sécurisée du protocole HTTP. Il protège l'intégrité ainsi que la confidentialité des données lors du transfert d'informations entre le client et le serveur. Le protocole HTTPS est reconnaissable par un cadenas dans la barre d'adresse du navigateur.

1.3 Les URL

Les sites Web ont une structure en arborescence. Une ressource (page, photo...) peut se trouver dans un dossier, lui-même situé dans un autre dossier et ainsi de suite...

Une URL (*Uniform Resource Locator*) est l'adresse d'une ressource d'un site Web, elle indique où elle se trouve dans l'arborescence du site.

Elle se compose de 3 grandes parties : le **protocole de communication client-serveur**, le **nom de domaine** et le **chemin vers la ressource**.

<https://www.facebook.com/pages/creation/>

2 Les langages HTML et CSS

2.1 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.

2.2 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.

2.3 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.

Un lien est composé principalement d'une URL cible et d'un libellé (le texte cliquable souvent de couleur, ou une image).

Le code HTML d'un lien hypertexte est :

```
<a href="URL">libellé</a>
```

3 Les moteurs de recherche

Le Web contient des milliards de pages et il est constitué de millions de sites.

Pour pouvoir trouver de l'information dans cette masse de données, les **moteurs de recherche** parcourent le Web pour créer un **index** (base de donnée/répertoire) à partir des contenus.

L'index permet de répondre aux requêtes des utilisateurs. Par exemple, si l'internaute recherche des informations sur « Informatique au lycée », le moteur recherchera dans son index toutes les pages qui contiennent ces mots.

Le moteur de recherche classera ensuite les réponses par ordre de pertinence, en combinant des critères tels que l'importance des mots dans chaque page, la popularité des pages, le profil de l'utilisateur.

Exemple. – Le moteur de recherche le plus connu est Google. Son algorithme de classement est appelé **PageRank**. Il mesure quantitativement la popularité d'une page Web (de 0 à 10).

4 Le Web et les données personnelles

L'accès à la plupart des sites Web est gratuit. Pour financer les coûts d'hébergement dans des serveurs, mais aussi de production des contenus, beaucoup de sites font appel à la publicité. Afin que les annonceurs puissent cibler ces publicités en fonction du profil de l'utilisateur, les sites Web sont amenés à collecter des données personnelles.

Certaines des données personnelles sont fournies volontairement par l'utilisateur (par exemple, lorsqu'il crée un compte sur un site Web), mais la plupart sont collectées à son insu, lorsqu'il navigue d'une page Web à l'autre en choisissant de cliquer sur un lien plutôt qu'un autre.

La collecte de ces données utilise plusieurs mécanismes. Le plus courant est le **cookie**, petit fichier texte qui se stocke dans l'ordinateur lors de la navigation, qui permet de conserver des données de l'utilisateur afin de faciliter la navigation et de permettre certaines fonctionnalités.

Grâce au **RGPD** (Règlement Général sur la Protection des Données) et aux options des navigateurs, il est possible de réduire ce traçage en réglant les paramètres de sécurité, mais il reste très difficile de l'empêcher complètement.

5 Le Web et le droit d'auteur

Tout le monde peut mettre des contenus sur le Web, soit en créant son propre site, soit en les téléversant sur des sites existants (comme Youtube ou Dailymotion pour les vidéos).

Si l'on souhaite utiliser un contenu dont on n'est pas l'auteur, il faut respecter les droits de celui qui l'a créé. En général cela nécessite son autorisation, à moins que le contenu soit dans le **domaine public**.

Si l'on veut autoriser la publication d'un contenu dont on est l'auteur, on peut utiliser une licence. La licence *Creative Commons* est souvent utilisée sur le Web. Elle permet de spécifier, par exemple, que toute publication doit inclure la mention explicite de l'auteur.

6 Repères historiques

- **1965.** – Naissance du concept d'hypertexte numérique par Ted Nelson : un texte numérique contenant des liens vers d'autres textes numériques ainsi interconnectés.
- **1989.** – L'invention du Web par Tim Berners-Lee : échange d'informations dans le réseau informatique du CERN (Conseil Européen pour la Recherche Nucléaire).
- **1993.** – « Mosaic », le premier navigateur à avoir popularisé le Web, développé aux États-Unis.
- **1994.** – Première transaction sécurisée sur le Web pour l'achat d'un album du chanteur Sting.
- **1995.** – Les pages Web deviennent dynamiques grâce au langage PHP et JavaScript. Elles étaient jusqu'à présent statiques avec le langage HTML.
- **2014.** – Plus d'un milliard de sites Web sur Internet.