



## Chapitre 03 Le web

Manuel pages 68 à 71

### I. Introduction

Le **Web** (toile ou réseau) désigne un système donnant accès à un ensemble de données (page, image, son, vidéo) reliées par des liens hypertextes et accessibles sur le réseau Internet.

### II. Repères historiques *manuel p. 58-59*

**1965** Invention et programmation du concept d'**hypertexte** par Ted Nelson.

**1989** Naissance au CERN par Tim Berners-Lee.

**1993** Mise dans le domaine public, disponibilité du premier navigateur **Mosaic**.

**1995** Mise à disposition de technologies pour le développement de site *Web* interactif (langage JavaScript) et dynamique (langage PHP).

**2001** Standardisation des pages grâce au **DOM (Document Object Model)**.

**2010** Mise à disposition de technologies pour le développement d'applications sur mobiles.

### III. Normalisation de la présentation de l'information

Sur le Web, les textes, photos, vidéos, graphiques, sons, programmes sont exprimés et assemblés dans divers formats normalisés par un consortium mondial (W3C : **World Wide Web Consortium**), ce qui permet une circulation standardisée de ces informations. Les pages Web sont écrites dans le **langage de balises HTML (HyperText Markup Language)**. Leur style graphique est défini dans le **langage CSS (Cascading Style Sheets)**.



HTML d'un site Web sur le lycée

```

1 <!DOCTYPE html>
2 <html>
3 <head>
4   <meta charset="utf-8">
5   <title>Le lycée</title>
6   <link rel="stylesheet"
7     href="style.css">
8 </head>
9 <body>
10  <h1>Les Classes de Seconde</h1>
11  <ul>
12    <li>La seconde</li>
13    <ul>
14      <li>Générale</li>
15      <li>Professionnelle</li>
16    </ul>
17  </ul>
18 </body>
19 </html>
  
```

Indique au navigateur que la page est en HTML  
Balise ouvrante

En-tête (*head*) : informations générales sur le document qui ne s'affichent pas dans le navigateur

Corps (*body*) : contenu de la page (textes, images, liens)

Balise fermante

Le langage **CSS (Cascading Style Sheets)** permet de définir des feuilles de style, c'est-à-dire le style graphique d'une page Web (arrière-plan, type et taille des polices de caractères, bordures, etc.).

```

1 body {
2   background-color: white;
3   font-family: "Open Sans", sans-serif;
4   padding: 5px 25px;
5   font-size: 18px;
6   margin: 0;
7   color: #C8C8C8;
8 }
9
10 h1 {
11   font-family: "Merriweather", serif;
12   font-size: 32px;
13 }
  
```

Indique que le style ci-dessous correspond au texte courant  
Couleur du fond : blanc  
Nom de la police de caractère : OpenSans  
Taille de la police : 18  
Couleur de la police : gris (code #C8C8C8)

Indique que le style ci-dessous correspond aux titres de niveau 1

Une feuille de style CSS

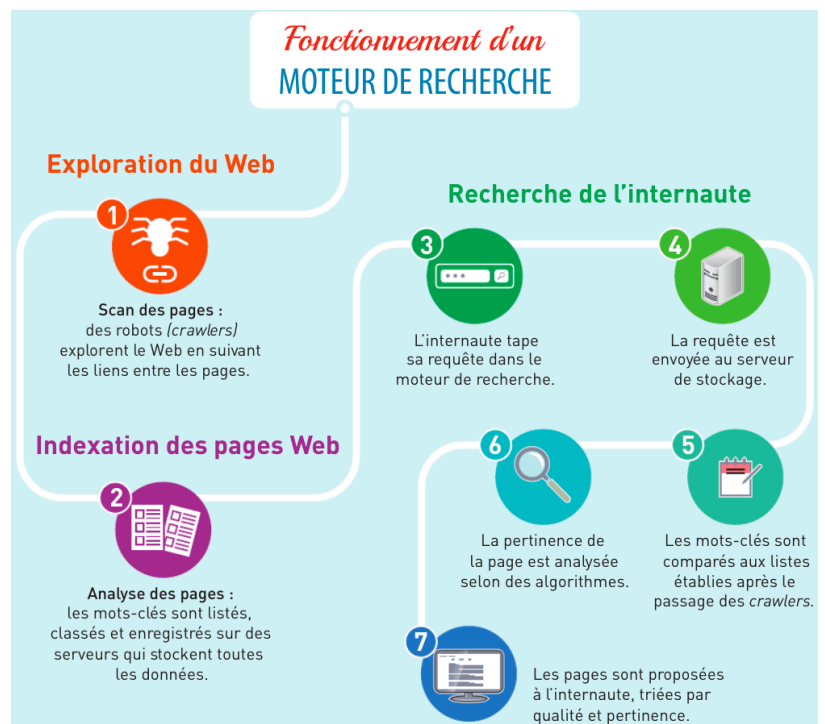
Les pages ont une adresse unique, nommée **URL (Uniform Resource Locator)**. Elles sont accessibles *via* Internet en utilisant le protocole **HTTP (HyperText Transfer Protocol)** ou sa version sécurisée **HTTPS** qui crypte les échanges. L'affichage des pages est réalisé chez l'utilisateur par un programme appelé navigateur. Un hypertexte est un texte augmenté de renvois automatiques à des textes, des images ou des sons. Initialement, un hypertexte se restreignait à la mémoire d'un seul ordinateur. Dans une page Web, ce renvoi se fait sur n'importe quelle machine du réseau Internet, par le truchement de l'adresse de la page Web du texte (URL) auquel il fait référence. La toile d'araignée construite par les liens peut être représentée sous la forme d'un graphe qui matérialise la structure du Web.

*http://*<sup>Protocole</sup>*www.exemple.com/*<sub>Nom de domaine</sub>*chemin/vers/une – ressource.html*<sup>Chemin</sup>

#### IV. Moteurs de recherche

Les moteurs de recherche permettent de trouver des informations dans des pages dont on ne connaît pas l'adresse, voire dont on ignore l'existence. La méthode de recherche appelée référencement naturel se décompose en trois grandes activités, réalisées par les moteurs de recherche :

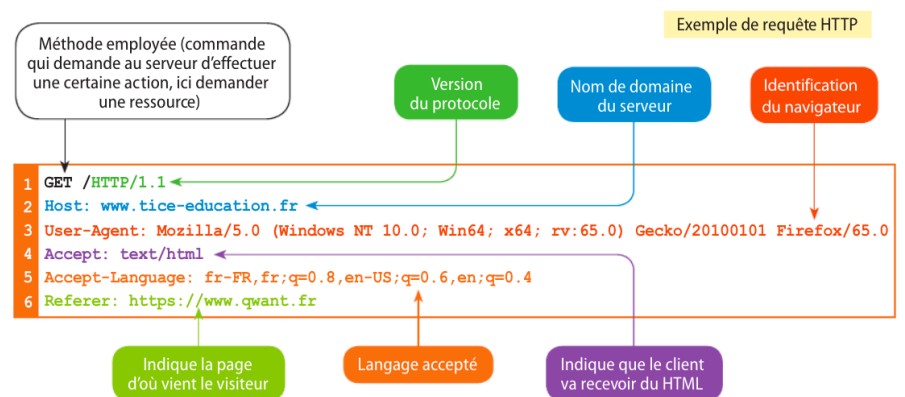
- ① **le parcours automatique du Web** pour collecter les pages visitées (aspiration des pages Web effectuée par des robots) ;
- ② **l'analyse du contenu des pages et leur indexation** sur les mots qu'elles contiennent (constitution d'un annuaire inverse qui associe à chaque terme les URL des pages où il apparaît) ;
- ③ la troisième activité, réalisée à chaque fois qu'un internaute fait une requête, **construit une liste ordonnée des pages (classement)** comportant les mots-clés de la requête. Leur ordre dépend notamment de leur popularité (principe des liens), de leur pertinence (aux mots de la requête), et de l'ordre des termes de la requête.



Les concepteurs de sites Web peuvent améliorer le référencement de leurs pages en choisissant bien les mots et en les plaçant à des endroits stratégiques dans les pages.

#### V. Interaction client/serveur

Le Web s'appuie sur le dialogue entre **clients** et **serveurs**. L'interaction est à l'initiative des clients (les applications qui se connectent au Web, dont les navigateurs), qui envoient des **requêtes HTTP** aux serveurs. Ces derniers renvoient leurs résultats : des pages qu'ils ont stockées ou qu'ils créent dynamiquement en fonction de la requête formulée. Les pages reçues par les clients peuvent contenir des codes exécutables (souvent en JavaScript) qui permettent aux clients d'effectuer des traitements en accédant aux ressources de son ordinateur



et en interagissant avec les serveurs. Les applications peuvent être paramétrées pour autoriser ou interdire l'accès à des ressources locales aux programmes téléchargés par les pages.

## **VI. Sécurité et confidentialité**

En formulant des requêtes sur des sites Web dynamiques et en laissant des programmes s'exécuter sur sa machine, l'utilisateur prend des risques : il peut communiquer des informations personnelles à son insu à des serveurs qui en gardent une trace, à distance ou localement par des *cookies*, ou encore charger des pages contenant des programmes malveillants, par exemple permettant d'espionner en continu les actions de l'utilisateur. Par ailleurs, un navigateur peut garder un historique de toutes les interactions, et le laisser accessible aux sites connectés. L'utilisateur peut utiliser des services qui s'engagent à ne pas garder de traces de ses interactions, par exemple certains moteurs de recherche. Il peut aussi **paramétrer son Navigateur** de façon à ce que celui-ci n'enregistre pas d'historique des interactions. De fausses pages peuvent encore être utilisées pour l'hameçonnage des utilisateurs. Un nom de lien pouvant cacher une adresse Web malveillante, il faut examiner cette adresse avant de l'activer par un clic.

## **VII. Impacts sur les pratiques humaines**

Dans l'histoire de la communication, le Web est une révolution : il a ouvert à tous la possibilité et le droit de publier ; il permet une coopération d'une nature nouvelle entre individus et entre organisations : commerce en ligne, création et distribution de logiciels libres multi-auteurs, création d'encyclopédies mises à jour en permanence, etc. ; il devient universel pour communiquer avec les objets connectés.

Le Web permet aussi de diffuser toutes sortes d'informations dont ni la qualité, ni la pertinence, ni la véracité ne sont garanties et dont la vérification des sources n'est pas toujours facile. Il conserve des informations, parfois personnelles, accessibles partout sur de longues durées sans qu'il soit facile de les effacer, ce qui pose la question du droit à l'oubli. Il permet une exploitation de ses données, dont les conséquences sociétales sont encore difficiles à estimer : recommandation à des fins commerciales, bulles informationnelles, etc. En particulier, des moteurs de recherche permettent à certains sites d'acquérir de la visibilité sur la première page des résultats de recherche en achetant de la publicité qui apparaîtra parmi les liens promotionnels.