SNT – Sciences numériques et technologie

Thème 1 : le Web

Cours complet

Seconde – Durée indicative : environ 2 heures

Objectifs : à la fin de ce thème, je sais...

- citer les grandes dates de l'histoire du Web ;
- décomposer l'URL d'une page et expliquer le rôle du protocole HTTP ;
- distinguer le rôle du HTML (contenu) et du CSS (mise en forme), reconnaître les principales balises ;
- expliquer comment un moteur de recherche trouve et classe les pages ;
- paramétrer son navigateur et reconnaître un site sécurisé ;
- respecter les droits d'auteur des contenus publiés sur le Web.

Ce cours contient tous les mots à placer dans le cours à compléter : ils sont mis en évidence. Les mots en gras sont à connaître.

1. Quelques repères historiques

Date	Ce qui se passe
1965	Ted Nelson invente le concept d'***hypertexte*** : un texte numérique qui contient des liens vers d'autres textes.
1989	Au CERN, Tim Berners-Lee propose d'utiliser l'hypertexte sur Internet : c'est la naissance du ***Web*** (*World Wide Web*, la « toile » mondiale).
1991	Le premier langage du Web, le ***HTML***, permet d'écrire des pages avec des titres et des liens : les pages sont statiques.
1993	Mosaic, l'un des premiers ***navigateurs***, affiche textes et images ensemble : le Web s'ouvre au grand public.
1994	Premier achat sécurisé sur le Web : c'est le début du commerce en ***ligne***.
1995	Les langages PHP et JavaScript rendent les pages ***dynamiques*** : elles réagissent aux actions de l'internaute.
2014	Le Web dépasse le ***milliard*** de sites.

À retenir

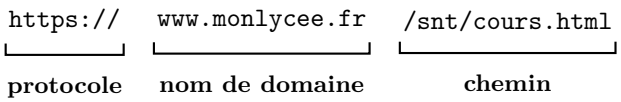
Le Web et Internet sont aussi différents qu'une peinture et la toile sur laquelle elle est peinte : **Internet** est le ***réseau*** qui relie les machines ; le **Web** est un ensemble de pages reliées entre elles par des liens ***hypertextes***, que l'on consulte grâce à Internet.

2. Le fonctionnement du Web

2.1 L'URL, l'adresse d'une page

Définition – URL

Une ***URL*** (*Uniform Resource Locator*) est l'adresse d'une page Web. Elle est composée de trois éléments :
— le ***protocole*** de communication (`http://` ou `https://`) ;
— le nom de ***domaine***, souvent celui d'une marque, d'une entreprise ou d'une association ;
— le ***chemin***, qui mène à une ressource ou à une page précise du site.



Les pages d'un site sont rangées en ***arborescence*** : une page peut se trouver dans un dossier, lui-même rangé dans un autre dossier (ici, la page `cours.html` est dans le dossier `snt`).

2.2 Le modèle client-serveur et le protocole HTTP

Pour afficher une page Web, le navigateur de ton ordinateur (le ***client***) envoie une ***requête*** à un ***serveur***, qui lui renvoie les ressources demandées (textes, images…) sous forme de code. Ces échanges utilisent le protocole ***HTTP*** (*HyperText Transfer Protocol*).

Ligne d'une requête HTTP	Ce qu'elle indique
`GET /snt/cours.html HTTP/1.1`	la ***méthode*** employée (GET : demander une ressource), le chemin et la version du protocole
`Host: www.monlycee.fr`	le nom de ***domaine*** du serveur
`User-Agent: Mozilla/5.0 ... Firefox`	le ***navigateur*** utilisé
`Accept-Language: fr`	la ***langue*** préférée de l'internaute

3. Les langages HTML et CSS

3.1 Le HTML : le contenu de la page

Une page Web est écrite en ***HTML*** (*HyperText Markup Language*), un langage de description. Il utilise des ***balises*** pour délimiter chaque élément : une balise ***ouvrante*** `<p>` au début, une balise ***fermante*** `</p>` à la fin. Les balises ne s'affichent pas à l'écran : elles indiquent au ***navigateur*** ce qu'il doit afficher.

Exemple – Le code d'une page Web très simple

```
<!DOCTYPE html>
<html>
   <head>
      <title>Mon lycée</title>
      <link rel="stylesheet" href="style.css">
   </head>
   <body>
      <h1>Bienvenue</h1>
      <p>Voici nos <a href="cours.html">cours de SNT</a>.</p>
   </body>
</html>
```

La partie `<head>` donne des informations au navigateur (titre de l'onglet, feuille de style) ; la partie `<body>` est la partie ***visible*** de la page.

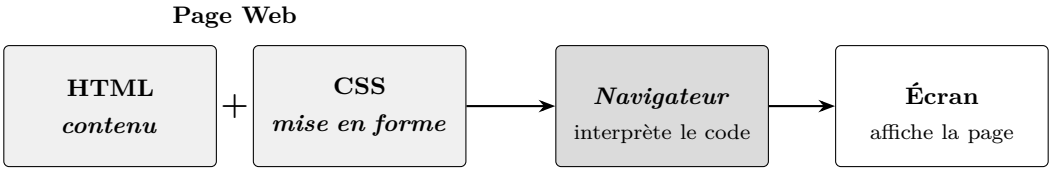
Balise	Rôle
`<h1>`, `<h2>`	***titre*** de niveau 1, de niveau 2
`<p>`	***paragraphe***
`<a>`	lien ***hypertexte***
`<img>`	***image***
`<ul>`, `<li>`	liste à puces, élément de la liste
`<table>`, `<tr>`	tableau, ligne du tableau

3.2 Le CSS : la mise en forme

Le langage ***CSS*** (*Cascading Style Sheets*, feuilles de style) définit la **mise en forme** de la page : couleurs, arrière-plan, type et taille des polices… Le HTML s'occupe du **contenu**, le CSS de la présentation. Une modification dans la feuille de style change automatiquement ***toutes*** les pages qui l'utilisent.

Exemple : avec la règle `h1 { color: red; }`, tous les titres de niveau 1 s'affichent en ***rouge***.

Complète le schéma.



3.3 Les liens hypertextes

Un lien ***hypertexte*** renvoie vers un autre endroit de la page, vers une autre page du site ou vers un autre site. Il s'écrit avec la balise `<a>` et l'attribut ***href***, qui donne l'adresse (l'URL) de la cible. Dans l'exemple ci-dessus, `<a href="cours.html">cours de SNT</a>` affiche les mots « cours de SNT » sur lesquels on peut cliquer.

4. Les moteurs de recherche

Les ***moteurs*** de recherche parcourent le Web à l'aide de ***robots*** (en anglais *crawlers*) qui ***indexent*** les pages : ils analysent leurs textes, leurs mots-clés, leurs balises… Des ***algorithmes*** classent ensuite les pages selon de nombreux critères : mots-clés, popularité, liens qui pointent vers elles, etc.

1. Les robots indexent des milliards de pages.
2. L'internaute tape sa ***requête*** (des mots-clés).
3. Le moteur extrait les pages qui correspondent à la requête.
4. Il trie et ***classe*** les résultats.
5. Il affiche les résultats.

Le référencement ***naturel*** (SEO, *Search Engine Optimization*) améliore la place d'un site dans les résultats grâce aux mots-clés de ses pages. Il ne faut pas le confondre avec le référencement ***payant*** : les annonces et les liens sponsorisés.

Tous les moteurs ne donnent pas les mêmes résultats, et certains collectent beaucoup plus de données personnelles que d'autres.

5. Sécurité et confidentialité sur le Web

5.1 Paramétrer son navigateur

Le ***navigateur*** (Firefox, Chrome, Edge, Safari…) garde des traces de ta navigation : l'***historique*** des pages visitées, les ***cookies*** et les formulaires remplis. On peut les effacer dans ses paramètres, ou avec les touches CTRL + MAJ + SUPPR.

— Un **cookie** est un petit fichier qu'un site enregistre dans le navigateur : il peut retenir un panier ou une connexion, mais aussi suivre l'internaute pour lui montrer de la publicité. Les sites doivent demander ton ***consentement*** avant de déposer des cookies publicitaires.

— La navigation ***privée*** n'enregistre ni l'historique ni les cookies sur l'ordinateur une fois la fenêtre fermée. Mais les sites visités et le fournisseur d'accès à Internet voient toujours la connexion.

5.2 Les sites sécurisés

Le protocole ***HTTPS*** (*HyperText Transfer Protocol Secure*) chiffre les échanges entre l'ordinateur et le serveur : il protège la ***confidentialité*** et l'intégrité des données. On le reconnaît à l'adresse qui commence par `https://` ; certains navigateurs affichent aussi un ***cadenas*** dans la barre d'adresse.

Attention : HTTPS protège la connexion, mais ne garantit pas que le site est honnête : un faux site peut, lui aussi, utiliser HTTPS.

5.3 Les droits d'auteur

Les textes, photos, images et vidéos publiés sur le Web sont la propriété de leur ***auteur*** : ils ne sont pas libres de droits. Pour les réutiliser, il faut l'autorisation de l'auteur, sauf si une licence le permet, comme les licences ***Creative*** Commons (CC).

6. Ce qu'il faut retenir

À retenir

— Le **Web** est un ensemble de pages reliées par des liens hypertextes, consultées grâce à Internet.

— Une **URL** se compose d'un protocole, d'un nom de domaine et d'un chemin. Le navigateur (**client**) envoie une requête **HTTP** au **serveur**, qui lui renvoie la page.

— Une page Web est écrite en **HTML** (le contenu, avec des balises) et mise en forme par le **CSS**.

— Les **moteurs de recherche** indexent les pages grâce à des robots et les classent avec des algorithmes.

— Trois règles d'or pour un Web plus sûr : naviguer sur des sites sécurisés (**HTTPS**), paramétrer son **navigateur**, respecter les **droits d'auteur**.

— Fin du thème —