

# Chapitre 2

## Le Web

★ Vidéo introductive d'Allô la Hotline.

### 2.1 Le Web et son fonctionnement

Le Web est un réseau d'informations (textes, images, sons...) supportées par Internet ; il est donc à ne pas confondre avec ce dernier : les deux sont aussi différents qu'un dessin et la feuille sur laquelle il est fait. Ces informations ou documents qui les contiennent sont reliées entre elles par le principe de l'hypertexte : texte contenant des liens vers d'autres documents. Ces documents sont hébergés sur des serveurs informatiques.

Un site Web est l'un des formats possibles de ces informations mais on peut directement accéder à un document (par exemple un pdf) sur le Web si on en connaît l'adresse URL (Uniform Resource Locator).

Pour consulter une page Web, un ordinateur client envoie une requête à un serveur qui lui renvoie les informations demandées, celles-ci sont alors mise en forme par le navigateur. Les échanges sont effectués selon le protocole HTTP(S) (Hypertext Transfer Protocol (Secure)) qui permet la communication entre client et serveur.

★ Vidéo de Monsieur Bidouille sur le HTTPS.

### 2.2 Les langages HTML et CSS

Les pages Web sont programmées en langages HTML (Hypertext Markup Language) et CSS (Cascading Style Sheets). Le HTML sert à programmer le contenu de la page, c'est donc le fichier .html qui contient les informations de la page (le texte, les menus, les adresses des fichiers qu'il doit appeler : styles, images, sons...). Le CSS sert à programmer l'apparence de la page (couleurs, polices, styles...), le fond est donc séparé de la forme !

Les liens hypertextes ou hyperliens envoyant vers d'autres pages ou documents apparaissent généralement soulignés ou colorés sur les pages Web. Le langage HTML est un langage « balise » où les instructions commencent avec une balise ouvrante `<commande>` et finissent avec une balise fermante `</commande>` (par exemple `<b>` pour du texte en gras `</b>`). La balise permettant de coder les hyperliens est

```
<a href="adresse url de la cible">nom du lien</a>
```

La lecture des pages Web – donc des fichiers HTML et CSS – se fait par le navigateur ; c'est la principale fonction de ces logiciels qui sont à ne pas confondre avec les moteurs de recherche (un navigateur peut afficher un site hébergé sur son ordinateur sans connexion internet). En France, les principaux navigateurs sont Firefox, Safari et Chrome.

## 2.3 Les moteurs de recherches

Les moteurs de recherche parcourent le Web, notamment grâce à des robots nommés *crawler*, l'indexent et classent les pages Web. Ces classements se font par des algorithmes analysant les textes, mots-clés, nombres de liens entrant (popularité) et sortant, etc. Il est toutefois possible de payer pour améliorer le référencement ou classement de son site.

Lorsque l'on effectue une recherche sur un moteur de recherche, celui-ci regarde donc dans la base de pages web qui a parcouru quelles sont celles qui correspondent aux mots-clés que l'on a entré dans la barre de recherche et les classe en fonction de ses algorithmes. Les résultats, pour une même recherche, diffèrent donc d'un moteur à l'autre.

Des exemples de moteurs de recherche : Google, Bing, Qwant, Duckduckgo, Écosia, Lilo.

★ Vidéo de Monsieur Bidouille sur Tor.

## 2.4 Sécurité et confidentialité sur le Web

Naviguer sur le Web ne se fait pas sans partager nombre de données personnelles (adresse IP, modèle de l'appareil que l'on utilise, version du navigateur, recherches, historiques des pages que l'on a visité...) avec les différents moteurs de recherche ou sites sur lesquels on va, et même des sites sur lesquels on ne va pas !

Il est possible de paramétrer les navigateurs afin de limiter, de choisir ou de masquer les données que l'on partage. Un autre gage de sécurité et de confidentialité sur une page est le protocole HTTPS symbolisé par un cadenas dans la barre d'adresse ; il ne concerne cependant que le transfert d'informations entre le client et le serveur et ne peut rien pour le contenu ou les informations partagées.

Il est aussi possible d'utiliser des bloqueurs de publicité (comme uBlock par exemple) mais cela ne va pas pour autant dire que la régie publicitaire proposant cette publicité ne recueille pas d'informations sur l'utilisateur. Les régies publicitaires sont des entreprises permettant aux annonceurs de diffuser les publicités de leurs produits à des clients ciblés en fonction des informations qu'elles ont récolté sur eux sur le Web.

On peut enfin utiliser un VPN, il s'agit d'un service (proposé par une entreprise ou une association) visant à masquer son identité et le contenu de sa navigation à son fournisseur d'accès internet et aux sites Web sur lesquels on navigue en utilisant un serveur intermédiaire qui va masquer l'adresse IP de l'utilisateur et chiffrer sa navigation. On peut aussi utiliser les VPN pour accéder à certains sites accessibles seulement à l'étranger ou des contenus spécifiques à certains pays (le catalogue de Netflix n'est pas le même aux États-Unis qu'en France). Toutefois, les fournisseurs de VPN sont soumis aux lois de leurs pays et peuvent enregistrer les informations que l'on souhaite masquer à l'insu de l'utilisateur ; il faut donc être encore vigilant sur le service que l'on utilise.

### 2.4.1 Fake news et recherches Web

★ Vidéo de Défakator sur les recherches Web

Quelques questions pour alimenter la réflexion et la discussion.

- Comment se prémunir des fake news ? Comment les détecter ?
- Pourquoi le profilage des utilisateurs est-il profitable aux sites, annonceurs et régies publicitaires ? Peut-on envisager des modèles économiques différents ?
- Quelles sont les informations qui peuvent être collectées sur les utilisateurs et comment ?
- Le profilage des utilisateurs par les moteurs de recherches et sites web est-il bénéfique aux utilisateurs ? Doit-il être encadré ?
- Pourquoi les informations sensationnelles (quitte à être fausses) sont mises en avant sur les sites ?
- Les bulles filtres liées aux profils des sites influencent-elles nos recherches ? Nos opinions ?

### 2.4.2 Profilage et marketing politique

★ Vidéo de Datagueule sur le marketing politique

Quelques questions pour alimenter la réflexion et la discussion.

- Comment le profilage d'utilisateurs permet de cibler des individus à des fins politiques ? Ce ciblage peut-il aller au delà du simple marketing politique ?
- Notre navigation sur le web influence-t-elle nos opinions ?
- Quelles sont les données que l'on récolte sur les utilisateurs ? Sont-elles toutes pertinentes ? Cette récolte est-elle souhaitable ?
- Des puissances étrangères peuvent-elles utiliser le marketing politique afin d'influencer la vie politique d'un autre pays ?
- Le profilage des utilisateurs doit-il être régulé ?
- Le marketing politique sculpte-t-il notre vision du monde ?