

WEB - BILAN (ce que dit le programme)

• 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.

• Repères historiques

- 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.

• 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** (*HypertextMarkupLanguage*). Leur style graphique est défini dans le **langage CSS** (*Cascading Style Sheets*).

Les pages ont une adresse unique, nommée **URL** (*Uniform Ressource 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*.

• 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 inversé 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 site *Web* peuvent améliorer le référencement de leurs pages en choisissant bien les mots et en les plaçant à des endroits stratégiques dans les pages.

- **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 leur résultat : 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.

- **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.

- **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.

Notion de la partie Internet, mais vu ici :

La neutralité du Net, présente dès l'origine du réseau, exprime l'idée que les routeurs doivent transmettre les paquets indépendamment du type de leur contenu : texte, vidéo, etc. Mais elle est constamment remise en cause par certains lobbies industriels.

| Contenus | Capacités attendues |
|--|---|
| Repères historiques | Connaître les étapes du développement du <i>Web</i> . |
| 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. |