

Rappel du référentiel Web :

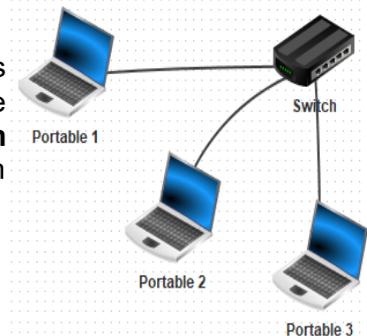
Contenus	Capacités attendues
Repères historiques	Définir les étapes du développement du <i>Web</i> .
Hypertexte	Maîtriser les renvois d'un texte à différents contenus.
Langages HTML et CSS	Distinguer ce qui relève du contenu d'une page et de son style de présentation. Étudier et modifier une page HTML 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 que toute requête laisse des traces.
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.

1 – Internet en deux pages

Avant de parler du **Web**, il faut peut-être clarifier la différence avec **Internet**.

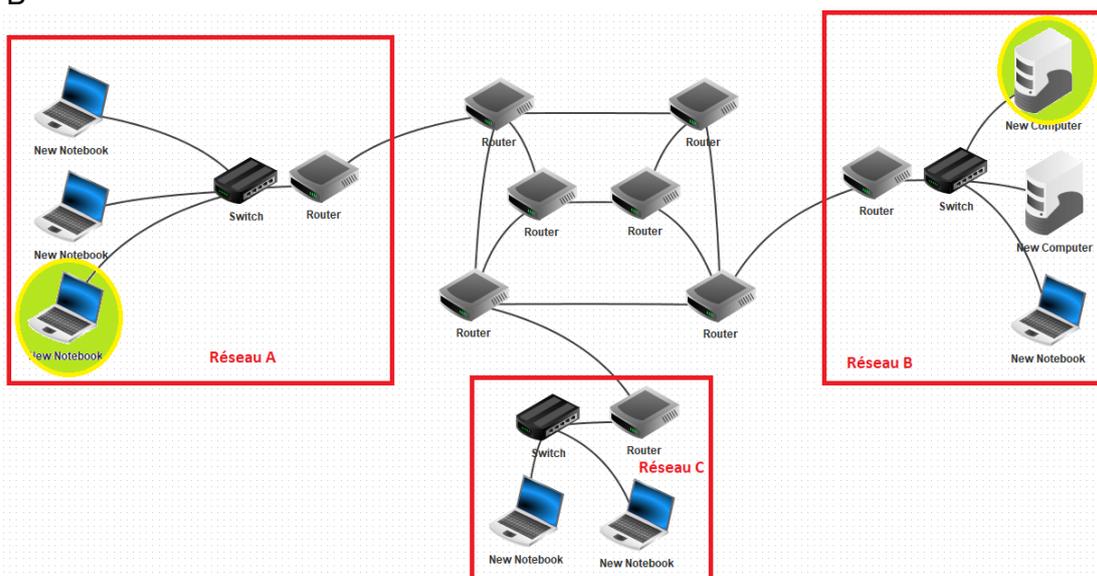
Qu'est-ce qu'un réseau informatique ?

- Un **réseau informatique** est un ensemble d'équipements informatiques (ordinateurs, serveurs...) reliés entre eux de façon à pouvoir échanger des informations. Le **switch** (commutateur) permet de relier entre eux les appareils d'un même réseau. Il agit comme un aiguillage.



Qu'est-ce qu'Internet ?

- **Internet** est lui le réseau informatique mondial accessible au public. Il permet donc de faire communiquer entre eux les réseaux informatiques qui y sont connectés. Internet est donc un **réseau de réseaux informatiques**. Ces réseaux sont reliés entre eux par l'intermédiaire de **routeurs** qui permettent le transport des messages d'un réseau à l'autre. Internet permet donc de faire communiquer deux machines qui ne sont pas sur le même réseau.
- Si l'un des ordinateurs du réseau A veut consulter une page Web accessible via le serveur situé sur le réseau B, une communication va devoir s'établir entre le réseau A et le réseau B



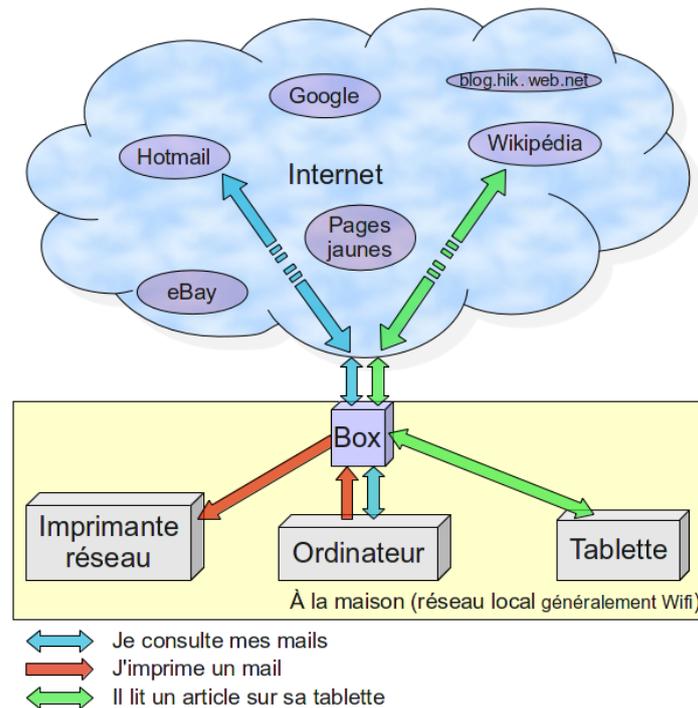
- Pour permettre cette communication, plusieurs règles de fonctionnement ont été édictées : les **protocoles**. Cette communication devra passer par des équipements particuliers qu'on nomme les **routeurs** : ce sont des ordinateurs dont la tâche est de faire transiter les communications.
- Comme on le voit sur le schéma, il n'existe pas de chemin unique pour passer du réseau A au réseau B.
- Une Box fait office à la fois de switch et de routeur.

Remarque : en réalité, votre box n'est pas directement reliée à Internet. Avant le routeur de la box se trouve le modem qui va permettre de transmettre les informations au point d'accès de votre fournisseur d'accès à Internet, le FAI.

Deux exemples de protocoles utilisés pour Internet (les plus connus)

1. Le **protocole TCP** (*Transmission Control Protocol*) est une façon de superviser la transmission d'un message entre une machine X et une machine Y. Ce protocole s'occupe de la découpe du message à envoyer, en petits bouts numérotés, d'informations (nommés segments) et vérifie qu'ils soient bien arrivés à destination.
2. Le **protocole IP** (*Internet Protocol*) est utilisé pour réaliser l'acheminement de données. Pour passer de machines en machines, on utilise donc notamment ce protocole.

On voit parfois Internet symbolisé par une sorte de nuage, donnant l'impression d'une immatérialité du réseau. Mais l'infrastructure matérielle est bien présente et indispensable. L'idée derrière le nuage est plutôt qu'il existe un grand nombre de chemins possibles et que ces chemins sont variables en fonction du moment.



(Bruno Voisin – Wikipedia - [CC BY-SA 3.0](https://creativecommons.org/licenses/by-sa/3.0/))

- Il existe une multitude d'applications qui utilisent Internet pour fonctionner. On trouve ainsi :
 - le FTP (*File transfert protocol*) qui permet le téléversement et le téléchargement sur un serveur distant (1971)
 - les courriers électroniques / emails (1971)
 - Usenet pour envoyer et lire des messages sur les forums de news (1979)
 - Internet Relay Chat (IRC) (1988)
 - **le Web qui permet de lire des documents hypertextes (1991)**
 - le pair-à-pair / peer-to-peer / P2P (1999)
 -
 - ...

2 – Le Web, un ensemble de documents hypertextes

- Le **Web** est **l'une** des applications qui utilisent Internet pour transmettre leurs données d'une machine à une autre.
- De façon fondamentale, le **Web** est constitué de l'ensemble des documents **hypertextes** auxquels il donne accès : on relie un document à un autre à l'aide d'un **lien** qu'on peut utiliser pour basculer vers un nouveau contenu hypertexte sans avoir à parcourir l'ensemble du document initial.

01 - Ouvrir un navigateur et aller sur Qwant (<https://www.qwant.com/?l=fr>) et faire une recherche sur « web ».

Et voilà : vous avez une nouvelle page texte qui comporte une multitude de liens vers d'autres documents. C'est bien un document **hypertexte** : un document texte amélioré qui permet d'atteindre un autre document hypertexte à l'aide d'un lien.

02 – Ouvrir la page Wikipédia traitant du Web :

https://fr.wikipedia.org/wiki/World_Wide_Web

03 – Partir en bas de page jusqu'à **Liens externes**. Cliquer sur le lien proposé (info.cern.ch) qui vous propose de vous montrer la première page du Web.

Liens externes [modifier | modifier le code]

- Notices d'autorité ✎ : Bibliothèque nationale de France✎ (données✎ Gemeinsame Normdatei✎ · Bibliothèque nationale d'Espagne✎
- "30 ans du web : l'âge de la trahison ?", La Méthose Scientifique, Fra
- (en) *A Little History of the World Wide Web*✎ [archive]
- (en) (fr) info.cern.ch✎ [archive], rétrospective historique du CERN
- (en) *Architecture of the World Wide Web, Volume One*✎ [archive], rec

En allant sur ce site, avec le premier lien proposé, vous pourrez voir le premier site Web de l'histoire (le site présentait le projet World Wide Web lui-même), visualisé avec votre navigateur.

Le deuxième lien vous permet même d'y naviguer en simulant l'affichage de l'époque ! Sans souris. Tout au clavier.

Le troisième lien explique (en anglais) que le Web a vu le jour en **1989** au CERN en Suisse sous l'impulsion de **Tim Berners-Lee**, un chercheur Britannique.

3 – Historique du Web

A l'époque, **en 1989**, **Tim Berners-Lee** travaille au CERN, il est ennuyé par la difficulté que représente le partage d'informations. Internet existait déjà et on pouvait partager des documents sur des serveurs FTP mais la compatibilité et la facilité de consultation posaient problème.

Il propose alors à ses supérieurs de travailler sur un projet de partage d'informations basé sur la consultation de documents hypertextes sur le réseau.

Tim Berners-Lee travaille alors avec **Robert Cailliau**, ingénieur informatique belge. Ensemble, ils mettent au point le premier serveur HTTP et le premier client HTTP de l'histoire **en 1990**.

En **mai 1991**, le projet Web est déployé pour les scientifiques du CERN. En **août 1991**, Tim Berners-Lee publie le projet sur **Usenet** et donne ainsi au monde entier (enfin, à ceux qui peuvent se connecter sur Usenet) la documentation complète du projet.

Lors de l'année **1992**, de nombreuses universités s'emparent du projet et créent leur propres serveurs également. On dénombre un trentaine de serveurs, dont un à Lyon.

C'est à partir de **1993** que le Web va réellement décoller ;

- De plus en plus de serveurs se créent : plus de 200 serveurs en fin d'année.
- Le **30 avril 1993** le CERN met le WWW dans le **domaine public** : les entreprises privés pourront donc à partir de cette date utiliser le Web.
- En **février 1993**, l'Université de l'Illinois publie le premier navigateur graphique **Mosaic** (porté sur Macintosh, Windows et Commodore Amiga).
- Fin **1993**, les grands groupes de presse commencent à publier sur le Web.

En **1994**, une première convention mondiale sur le Web permettra de poser les bases de l'institution dédiée à veiller à l'avenir du Web, le **World Wide Web Consortium (W3C)**.

En **fin 1994**, il y aura plus de **1000 sites** et une nouvelle avancée : l'apparition de **Netscape**, le premier navigateur graphique pour le grand public. Les premiers **moteurs de recherche** (Yahoo!) apparaîtront et les premières publicités également.

En **1995**, Microsoft lance son propre navigateur **Internet Explorer**, qui sera préinstallé sur tous les pc vendus avec Windows 95. La guerre des navigateurs commencent. Pendant de nombreuses années, Internet Explorer sera en situation de quasi-monopole et tentera d'imposer ses propres fonctionnements plutôt que de suivre les normes. Mise à disposition du **javascript** pour rendre la page interactive localement et du **php** pour permettre de rendre un site dynamique (l'internaute peut laisser des messages ou passer des commandes qui sont traitées automatiquement, sans passer par des emails lus par un service clientèle). Création d'**Amazon** et début du « commerce numérique » pour les sites Web.

En **1997**, on atteint un million de sites Web.

En **1998** sort un nouveau moteur de recherche beaucoup plus performant : **Google**. Il est basé sur un algorithme, le PageRank, classant la pertinence de la recherche par autre chose que la simple étude des mots présents sur le site : il intègre le nombre de liens externes visant le site étudié. Il dépasse rapidement ses concurrents.

En **2001**, lancement de **Wikipedia**, l'encyclopédie collaborative Wiki en ligne.

En **2004**, apparition de **Firefox 1.0**, un navigateur web libre et gratuit et de **Facebook**. Framasoft est réorganisé et voit la création de l'**Association Framasoft** et son site dédié, visant à valoriser les logiciels libres.

L'année **2005** est parfois vue comme celle du début du Web 2.0, centré sur les internautes eux-mêmes plutôt que sur la documentation disponible. Les internautes créent du contenu, interagissent avec le Web et deviennent ciblés par le marketing. Explosion des blogs, des réseaux sociaux et des wikis. Apparition de **youtube** qui sera racheté en 2006 par Google. Google lance également Google Analytics qui permet aux webmasters de suivre facilement la fréquentation de leurs sites et à Google de suivre les utilisateurs.

En **2007** : commercialisation de l'iphone qui marque le début d'un nouveau type de terminal pour atteindre le Web et surtout apparition d'**AppStore** qui permet d'utiliser **des applications** plutôt que de surfer sur le Web. Vers la fin du principe de l'hypertexte ?

En **2008**, apparition de **Chrome** : un navigateur web propriétaire développé par Google (mais basé sur le projet libre **Chromium** qui est la base de beaucoup d'autres navigateurs et qui est tout à fait fonctionnel). Il séduit un grand nombre d'utilisateurs et devient rapidement le leader du marché. Création de la **Journée Mondiale contre la Censure sur Internet** par Reporters sans frontières. La principale cible de cette manifestation était à l'origine la Chine et ses lois concernant le contrôle d'Internet et du Web.

En **2012** : fermeture du site **MegaUpload** par la justice américaine. Fin officielle du Minitel.

En **2013** : apparition de **Qwant**, un moteur de recherche Français qui annonce ne pas tracer ses utilisateurs. Il n'est néanmoins ni un logiciel libre, ni un moteur ouvert.

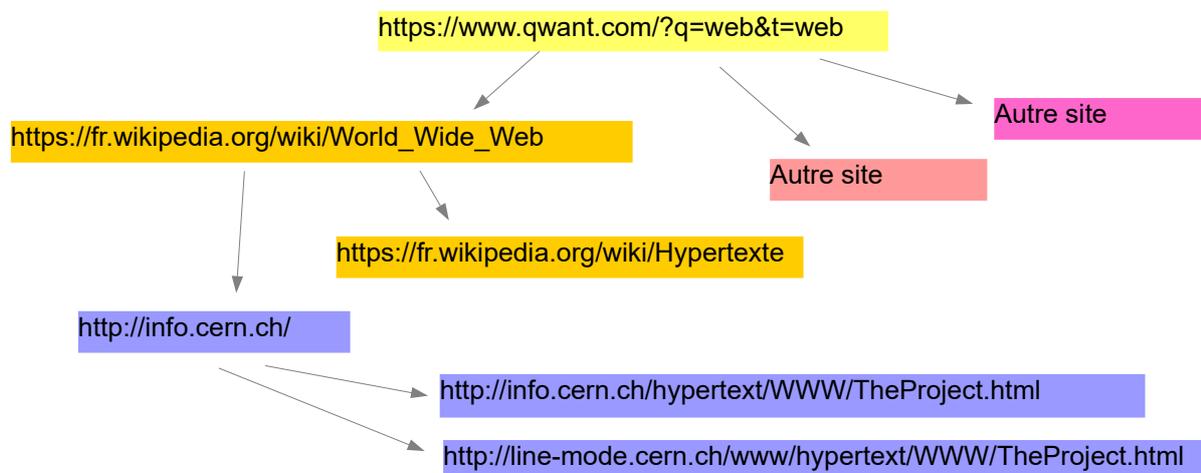
En **2014**, le W3C décide d'intégrer un **système de protection** des contenus (notamment vidéos) au sein du HTML5 même si cela va à l'encontre de l'ouverture du Web. La fondation Mozilla cède sous peine de voir le Web se fractionner.

En **2018**, Microsoft modifie son moteur de rendu pour Edge en le remplaçant par Chromium. On retrouve un important monopole dans le domaine (cette fois opéré par Google) et la fondation Mozilla craint que les dérives d'Internet Explorer ne se répètent.

En **2019** : Plus de 9 milliards d'utilisateurs mais assez mal répartis sur la planète.

3 – Visualiser le Web ?

Réaliser une carte du Web dans son ensemble est loin d'être aisé : il faudrait relier chaque page aux pages qu'elle permet de joindre. Tout en sachant que le Web évolue constamment : des liens se créent, des pages disparaissent...

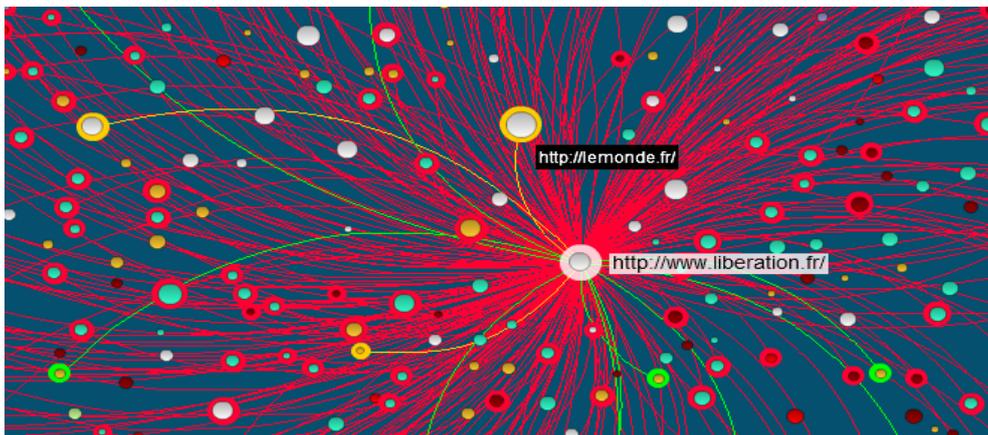


En plus, certains liens sont des liens internes au site (on reste sur le même site), certains liens sont des liens internes à la page (on reste sur la même page) et d'autres sont des liens externes (on quitte le site pour aller sur un autre site) . . .

La visualisation du Web peut se faire sous forme de **graphe** : chaque site est représenté par un cercle et on relie les sites/cercles entre eux en fonction des liens existants. Ce type de représentation donne une multitude d'informations.

Lors de la **campagne Européenne de 2009**, une équipe a ainsi réalisé un graphe représentant les interactions entre les différents sites parlant de la campagne. On ne montre pas l'intégralité des pages d'un site mais uniquement les liens d'un site à l'autre.

<https://www.touteurope.eu/actualite/cartographie-de-la-toile-europeenne.html>



J'ai choisi de visualiser les liens entrants ou sortants de liberation.fr
 Les liens rouges sont les liens entrants (les sites qui créent un lien vers liberation.fr)
 Les liens jaunes sont les liens sortants (les sites dont liberation.fr donnent le lien).
 Les liens verts sont les liens réciproques.

04 – En quoi un tel graphe peut-il être intéressant pour un sociologue ou un politologue ?

On trouve également assez facilement une autre carte sur le Web : celle réalisée en 2012 par un programmeur russe.

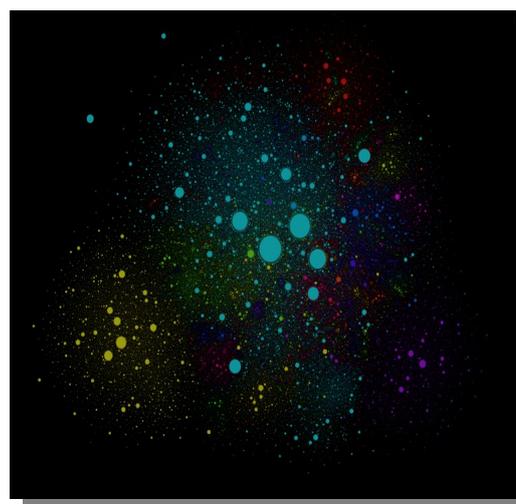
<http://internet-map.net/>

Sur cette représentation, on ne peut pas voir les liens. Néanmoins, chaque pays est représenté par une couleur.

Chaque site possède une taille lié à son trafic et l'algorithme générant la carte semble placer les sites en fonction des hyperliens entre eux et également des centres d'intérêt.

Les couleurs sont les suivantes :

- en bas à gauche (en jaune) : la Chine
- en bas à droite (en violet) : le Japon
- en bleu ciel : les USA
- en haut à droite (en rouge) : la Russie
- au milieu à droite (différentes couleurs, violet pour France) : les pays Européens
- au milieu à gauche (en vert) : l'Inde



Si vous avez encore un peu de temps :

05 – Quels sont les grands états de Web en 2012 ?

06 – Quels étaient les grands sites français ?

07 – Une telle représentation est-elle nécessairement objective ? L'algorithme choisi ou les données sélectionnées ont-elles un impact fort sur la carte obtenue ?



Merci à Hervé Owsinski pour son travail (très peu modifié par JNB)