

# WEB

## 1) Du navigateur au serveur web

### 1) Différence entre l'internet et le web, HTTP, HTTPS

L'internet est le réseau mondial où les ordinateurs communiquent entre eux par le protocole TCP / IP.

L'information qui peut être partagée est multiple : musique, réseaux sociaux, vidéo, texte, etc ... Le World Wide Web correspond à l'ensemble des pages internet liées entre elles par le protocole HTTP : *hypertext transfer protocol*

**Exemple :** <https://bonne-terre.fr> est une page web, on utilise le protocole HTTPS pour se connecter. Le protocole HTTPS est la version sécurisée du HTTP, les pages HTTP disparaissent de l'internet car les pirates peuvent récupérer des informations durant la fréquentation de la page. [https://lycee.bonne-terre.fr/?page\\_id=248](https://lycee.bonne-terre.fr/?page_id=248) est une page du même site internet, <https://scolinfo.net> est une page d'un autre site internet, dans les deux cas on passe de l'une à l'autre par un lien hypertexte.

## 2) Historique

13 mars 1989 : Tim Berners-Lee lance l'idée du Web

1990 : le Web est baptisé, première page internet : <http://info.cern.ch/hypertext/WWW/TheProject.html>

1992 : premier site français

1993 : 10.000 sites internet

1994 : premier achat sur le web (un CD)

1995 : premiers moteurs de recherches dont Lycos <https://www.youtube.com/watch?v=NahiFFd4bwk>

1996 : netscape et internet explorer voient le jour. Netscape est l'ancêtre de Firefox, Internet Explorer celui de Edge

1998: création de Google

2018 : plus de connexions mobiles que sur ordinateur

## 3) URL, nom de domaine

Une **URL** est une adresse internet précise qui correspond à un document sur l'internet. Par exemple [https://lycee.bonne-terre.fr/?page\\_id=9689](https://lycee.bonne-terre.fr/?page_id=9689) et [https://lycee.bonne-terre.fr/?page\\_id=248](https://lycee.bonne-terre.fr/?page_id=248) sont deux urls différentes, elles appartiennent au même **nom de domaine**.

Le nom de domaine bonne-terre.fr n'est pas la propriété de l'établissement, il est « loué » chaque année à un prestataire de service agréé par l'ICANN. Un nom de domaine coûte de 2 € à 30 € par an à sa création. Certains noms de domaines peuvent prendre de la valeur et sont revendus plusieurs millions d'euros. La commande **whois** permet d'avoir des informations sur un nom de domaine.

L'extension .fr signifie qu'il s'agit d'un site français. Ce n'est pas une obligation, chacun est libre de poser l'extension qu'il veut, sachant que certaines extensions coûtent plus cher que d'autres.

On parle de « cybersquatting » lorsque quelqu'un dépose un nom de domaine d'une marque dont il n'est pas propriétaire ou d'une personne avec qui il n'a pas de lien. C'est ainsi que les grands groupes vont déposer tous les noms de domaine en lien avec leur marque et parfois même leurs noms

## WHOIS LOOKUP



**bonne-terre.fr is already registered\***

```
%%  
%% This is the AFNIC Whois server.  
%%  
%% complete date format : YYYY-MM-DDThh:mm:ssZ  
%% short date format : DD/MM  
%% version : FRNIC-2.5  
%%  
%% Rights restricted by copyright.  
%% See https://www.afnic.fr/en/products-and-services/services/whois/whois-special-notice/  
%%  
%% Use '-h' option to obtain more information about this service.  
%%  
%% [165.254.21.161 REQUEST] >> bonne-terre.fr  
%%  
%% RL Net [#####] - RL IP [#####.]  
%%
```

```
domain: bonne-terre.fr  
status: ACTIVE  
hold: NO  
holder-c: SARE22-FRNIC  
admin-c: SBT262-FRNIC  
tech-c: GR283-FRNIC  
zone-c: NFC1-FRNIC  
nsl-id: NSL29045-FRNIC  
registrar: GANDI  
Expiry Date: 2020-10-11T22:07:11Z  
created: 2006-06-06T14:46:45Z  
last-update: 2019-10-02T06:42:08Z  
source: FRNIC
```

dérivés ! Coca-cola.com, cocacola.com, coke.com sont la propriété de la compagnie Coca-cola, ainsi que toutes les extensions disponibles.

**Activité** : aller sur le site gandi.net et essayez d'acheter le site coca-cola.com

#### 4) Navigateur web.

##### a) Les différents navigateurs web

Pour aller sur une page web il est nécessaire d'avoir un outil qui s'appelle le navigateur. C'est un logiciel qui permet d'afficher les pages web. Il existe bon nombre de navigateurs, voici les principaux.



Chrome, Opera, Brave, Edge Chromium sont construits sur le même « moteur », il s'agit du logiciel Chromium. Firefox et Safari possèdent leur propre moteur. Il faut comprendre que c'est un problème de fond car les programmeurs respectent des standards du web, de la même manière qu'on respecte les règles orthographiques. Comme l'ensemble des navigateurs sont construits sur Chromium, les développeurs de sites web ont tendance à coder pour Chromium et ne plus respecter les standards. C'est un peu comme si Peugeot parce qu'il vend beaucoup de voitures en France, décidait du code de la route. Il s'agit donc d'un enjeu économique et des libertés de l'internet car Chromium est principalement développé par Google.

- Chrome est le navigateur de Google
- Firefox est un logiciel libre, il appartient à la fondation Mozilla. Le code du programme est accessible à tout le monde.
- Safari est le navigateur de Apple
- Opera est le navigateur de la société Opera Software basée en Norvège
- Brave est un navigateur qui se veut libre et sécurisé, il appartient toutefois à une société commerciale, Brave Software.
- Edge Chromium est le navigateur de Microsoft. Microsoft a développé pendant des années le navigateur Internet Explorer aujourd'hui abandonné, il a continué à développer son propre moteur mais a fini par abandonner pour se baser sur Chromium.

Quel navigateur choisir ? Pas si simple. Un utilisateur des produits Google trouvera intéressant d'utiliser Chrome car l'ensemble des services qu'il utilise seront synchronisés les uns aux autres, tout comme un utilisateur d'Apple utilisera Safari. Le risque de croiser de nombreux services c'est de donner toutes ses informations à une société commerciale. Par exemple, si vous êtes utilisateurs d'un téléphone Android, Google sait qui vous appelez, où vous êtes, vos trajets, mais si vous utilisez aussi son navigateur, il connaît l'intégralité des sites que vous consultez. Firefox est un logiciel libre, il n'est pas en lien avec une société commerciale, il ne recueille donc pas de données. Brave devrait se positionner de la même manière, néanmoins il est apparu plusieurs fois qu'il avait fauté sur des liens publicitaires. Le choix d'un navigateur est avant tout personnel et n'est pas sans conséquence.

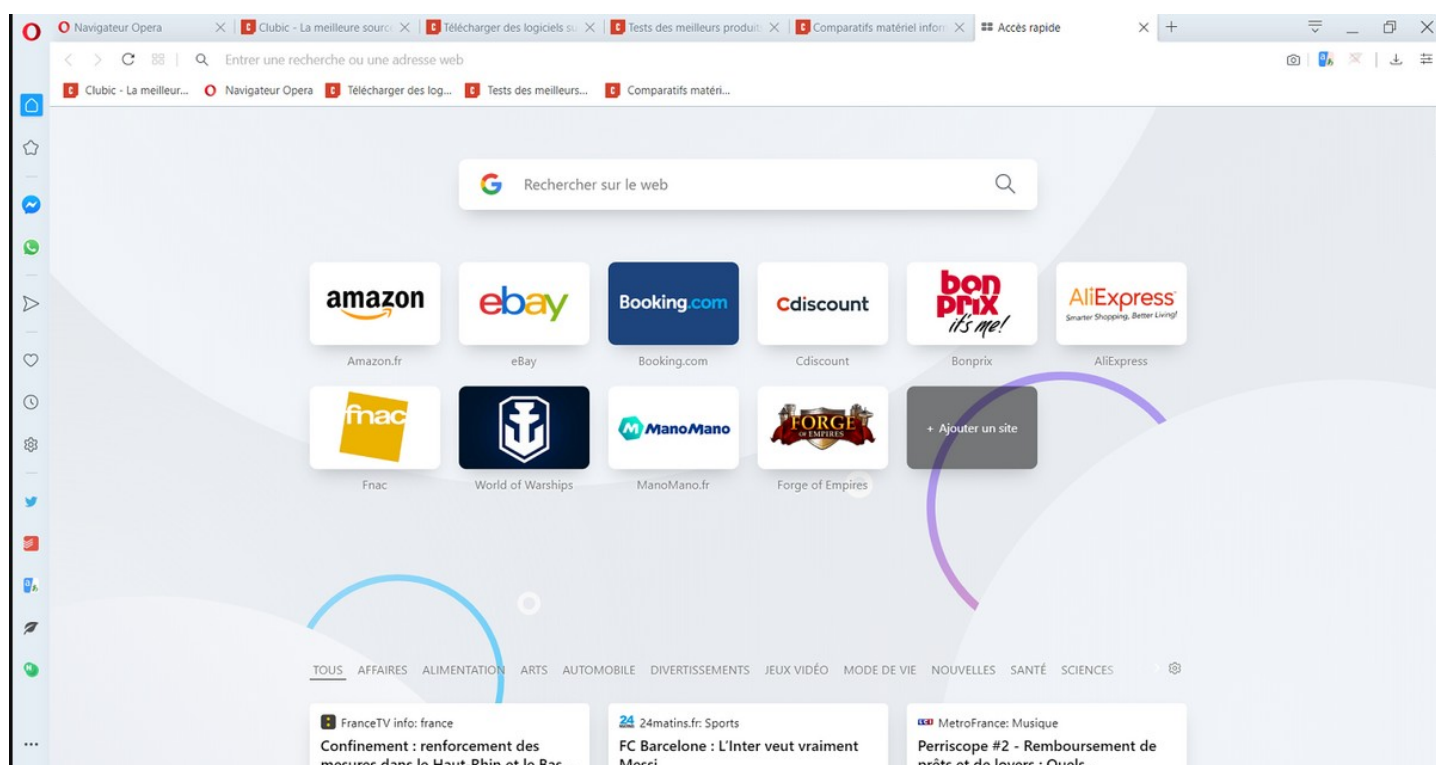
##### b) Modèle économique.

À part le navigateur Firefox qui n'appartient pas à une entreprise mais une fondation, tous les navigateurs sont en lien avec une entreprise. Une entreprise doit gagner de l'argent, un navigateur nécessite des centaines de développeurs pour corriger les problèmes, intégrer les nouveautés, il faut donc gagner de l'argent. La source principale d'argent des navigateurs c'est le contrat le liant aux moteurs de recherche.

Si vous ne connaissez pas l'URL, mais que vous cherchez une information sur l'internet, vous allez utiliser un moteur de recherche. Nous verrons plus loin que le moteur de recherche est une manière très juteuse de gagner de l'argent.

Par exemple, si Google a développé son propre moteur de recherche (Google), et qu'il paraît logique que Google Chrome utilise son propre navigateur, comme Edge Chromium va utiliser Bing, le moteur de recherche de Microsoft, Apple n'a pas développé son propre moteur de recherche. Google a passé un accord avec Apple de l'ordre de 9 milliards de dollars par an pour que Google soit le moteur de recherche par défaut dans Safari. Ainsi, si vous utilisez un iPad, un iPhone, un iMac, vous utilisez le moteur de recherche de Google, l'un des principaux concurrents de la marque à la pomme.

Le sponsoring est aussi un moyen de financer le navigateur, comme on peut le voir dans cette capture d'Opera.



Les sites mis en avant par défaut ont payés pour apparaître dans le navigateur. Pour Google Chrome et Edge Chromium, du fait d'utiliser leur propre moteur de recherche, la façon de gagner de l'argent est plus subtile : enfermement de l'utilisateur dans une suite de produit de la marque, récupération de données personnelles qui permettent de mieux cibler l'utilisateur.

### c) Moteurs de recherche

Avec plusieurs milliards de pages web existantes, le moteur de recherche est apparu comme une évidence pour **référencer** les pages internet. Seulement, comme sur un même sujet il existe parfois plusieurs millions de pages, il est nécessaire de trier les pages et de renvoyer un ordre d'apparition. Chaque jour, des robots (des programmes) indexent des millions de pages c'est-à-dire qu'ils vont trouver dans la page des mots clés qui vont revenir souvent afin de faciliter la recherche.

Si on fait une recherche sur « lycée agricole Hérault », il est normal de trouver le lycée Bonne Terre. On peut toutefois s'interroger quant au classement, pourquoi le lycée agricole Bonne Terre arrive avant le lycée agricole de Gigean par exemple. Un site web est noté en fonction de son **PageRank** c'est son classement selon un algorithme qui a été déterminé par Google. Comme il est impossible de juger de la qualité de milliards de sites, Google fait confiance à l'ensemble des sites entre-eux. Concrètement, plus un site a de liens vers lui, plus son PageRank augmente. Les sites les plus fréquentés sont sur le principe ceux qui sont le plus cités sur le web, sauf que ce n'est pas si simple.

Comme pour les navigateurs internet, entretenir un moteur de recherche coûte très cher, scanner des millions de pages par jour demande de gros serveurs, des ingénieurs, etc ... Le moteur de recherche se doit de gagner de l'argent.

Les premiers sites qui apparaissent dans un moteur de recherche ne sont pas les plus pertinents mais ceux qui payent.

The screenshot shows a Google search for "chaussures". At the top, there's a search bar with "chaussures" and a microphone icon. Below it, navigation tabs for "Tous", "Maps", "Shopping", "Images", "Actualités", "Plus", "Paramètres", and "Outils" are visible. A notification banner for Google's privacy policy is present. The search results are dominated by sponsored content:

- Announce - www.sarenza.com/**: "Chaussures sur Sarenza.com - Grand choix de marques". It offers categories for "Chaussures pour Femme", "Chaussures pour Enfants", "Chaussures pour Homme", and "Nouvelle Collection Femme".
- Announce - www.chaussures.fr/chaussures**: "Chaussures - Summer Deals jusqu'à -50% - chaussures.fr". It promotes "Collections à la mode Tendances estivales" and offers "Offre spéciale" for both "Femme" and "Homme".

On the right side, a "Résultats chaussures" section is highlighted with a red box. It features a grid of shoe products:

- PROMOTION**: Brooks Glycerin 17 425 Unisex Brooks, 119,00 € (470 €), 5 stars (580).
- Sandales Orthopédiques... Maxi Offres, 28,90 €, Livraison gratuite.
- Sandales orthopédiques... Corps & Vie, 29,99 €, Livraison gratuite.
- Brooks Adrenaline GTS 20 406... Brooks, 140,00 € (299), 5 stars (299).
- Tommy Hilfiger Baskets Harlow ... La Redoute, 59,90 €, 5 stars (76).
- Chaussure indestructible... Lockfeet, 49,99 €, Livraison gratuite.

At the bottom, a map of the region around Agde is shown, with a red box highlighting the location of "Chaussexpo" and "Besson Chaussures".

Comme on peut le voir sur une simple recherche dans Google, les premiers liens qui apparaissent sont des annonces, cela signifie que ces sites ont payé pour arriver en premier. Cela ne veut pas dire qu'ils sont meilleurs, mais simplement qu'ils ont donné de l'argent sans que le moteur de recherche vérifie ou non la pertinence du site. Il n'est pas rare que soient mises en avant des escroqueries sans aucune réaction du moteur qui se contente de prendre l'argent.

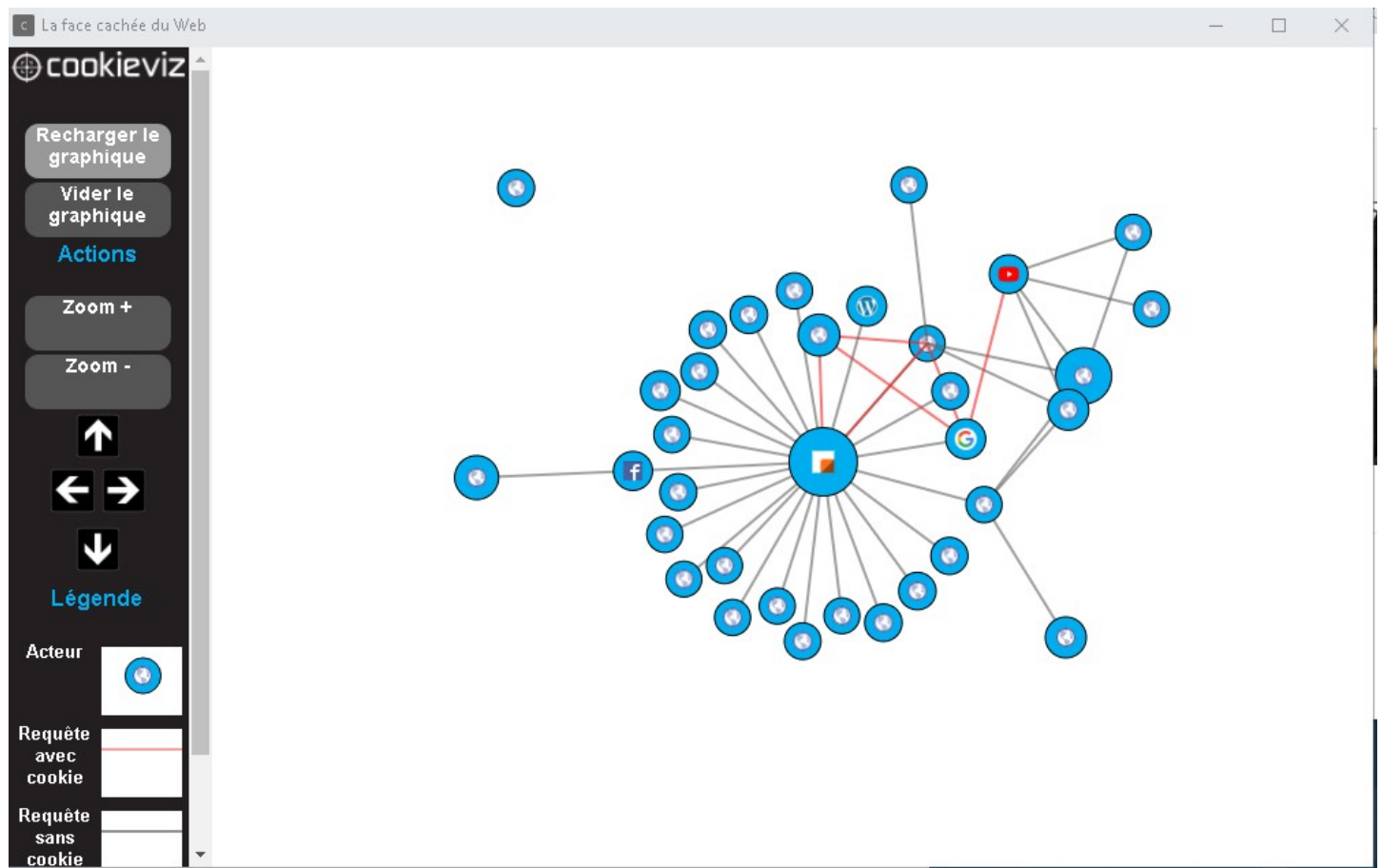
#### d) Sécurité, éthique

Se connecter à une page web avec le navigateur c'est réaliser une relation client serveur. Le site web est sur un ordinateur distant, il envoie l'information au navigateur mais la communication se fait dans les deux sens. À chaque visite sur un site internet, le navigateur va écrire un fichier **cookie** qui enregistre une partie de votre activité sur le web : les sites visités, votre adresse IP donc votre géolocalisation, vos recherches sur le site. Lorsque vous visitez une nouvelle page web, cette dernière va lire vos fichiers cookies et adapter ses contenus par rapport à vous.

Vous ne serez pas surpris si vous avez visité des sites de chaussures d'avoir sur un site commercial des propositions de chaussures à acheter. De la même manière vous ne serez pas étonné de trouver des publicités avec le nom de votre ville. Il est désormais obligatoire pour les sites d'afficher un message vous informant qu'il y a des cookies, tout comme il est possible de désactiver les cookies depuis le navigateur. Néanmoins comme

c'est une information particulièrement importante pour les sites, vous pouvez vous retrouver bloqué dans l'utilisation du site.

**Activité** : utiliser cookieviz 2.0 pour vérifier l'impact du traçage par les cookies lorsque vous vous déplacez d'un site à l'autre



Google représente 90 % des recherches au monde, il écrase l'intégralité de la concurrence. Bing le moteur de recherche de Microsoft ne fait que quelques pourcents, tout comme Qwant le moteur de recherche promu par la France.

**Activité** : Que pensez-vous de la situation de monopole de Google dans la recherche ?

**Activité** : dans bing faire une recherche sur Chrome ou Firefox. Que se passe-t-il ? Qu'en pensez-vous ?

**Activité** : réaliser une carte mentale dans laquelle vous ferez apparaître les grands groupes informatiques, les appareils qu'ils produisent, le système d'exploitation utilisé, le navigateur utilisé et le moteur de recherche utilisé



## II) Pages web, html et CSS

Un serveur web est un simple ordinateur sur lequel on a des pages écrites dans un langage web. Le navigateur interprète le code html pour afficher la page avec la mise en forme demandée.

**Activité :** se rendre sur le site <https://html5ed.it/editeur-html/>. Cliquer sur l'icône du professeur et le à



The screenshot shows the HTML5ed.it online editor. The top navigation bar includes 'Fichier', 'Editer', 'Voir', 'Insérer', 'Format', 'Outils', and 'Tableau'. The main editor area is split into two panes. The left pane (WYSIWYG) displays a preview of the HTML code with a yellow cartoon character and a 'NETTOYER' button. The right pane (Code source) shows the corresponding HTML code with line numbers and syntax highlighting. The code includes a header, a main content area with a table, and a footer with a link to GeekPrank.

Il apparaît que le langage html est réalisé sous forme de **balises ouvrantes et fermantes**. Par exemple le couple de balises `<h3> </h3>` permet de réaliser un titre de niveau 3. Si ici, il s'agit d'une page simple, il faut imaginer qu'un site internet est réalisé parfois par plusieurs centaines de pages. De façon à ce que les pages aient toutes la même présentation, on va rajouter un fichier CSS qui contient l'intégralité de la présentation des pages du site web.

**Activité :** déterminer l'utilité de chacune des balises.

**Activité :** rajouter une colonne et une ligne au tableau en utilisant les balises `<td> </td>` et `<tr> </tr>`. Changer l'image en remplaçant le lien vers une image de votre choix. Changer le lien, changer les couleurs.

**Version 1.0 :** le 10/07/2020 premier jet.