



# Le Web

Formation continue – Année scolaire 2019-2020



Jean-Damien Leroy – Philippe Morin



# Contenus et capacités attendues

BO spécial N° 1 du 22 janvier 2019

Contenus	Capacités attendues
Repères historiques	Définir 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 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. Sécuriser sa navigation en ligne et analyser les pages et fichiers.

# Exemples d'activités

BO spécial N° 1 du 22 janvier 2019

- Construire une page Web simple contenant des liens hypertextes, la mettre en ligne.
- Modifier une page Web existante, changer la mise en forme d'une page en modifiant son CSS. Insérer un lien dans une page Web.
- Comparer les paramétrages de différents navigateurs.
- Utiliser plusieurs moteurs de recherche, comparer les résultats et s'interroger sur la pertinence des classements.
- Réaliser à la main l'indexation de quelques textes sur quelques mots puis choisir les textes correspondant à une requête.
- Calculer la popularité d'une page à l'aide d'un graphe simple puis programmer l'algorithme.
- Paramétrer un navigateur de manière qu'il interdise l'exécution d'un programme sur le client.
- Comparer les politiques des moteurs de recherche quant à la conservation des informations sur les utilisateurs.
- Effacer l'historique du navigateur, consulter les cookies, paramétrer le navigateur afin qu'il ne garde pas de traces.
- Utiliser un outil de visualisation tel que Cookieviz pour mesurer l'impact des cookies et des traqueurs lors d'une navigation.
- Régler les paramètres de confidentialité dans son navigateur ou dans un service en ligne.

# Repères historiques

## 1965

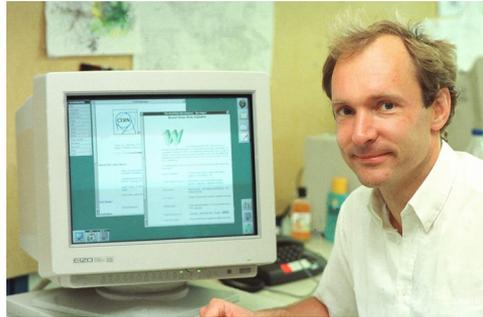
Concept d'hypertexte



Ted Nelson est un sociologue américain. Il publie en 1965 le concept d'hypertexte numérique : un texte numérique contenant des liens vers d'autres textes numériques ainsi interconnectés.

## 1989

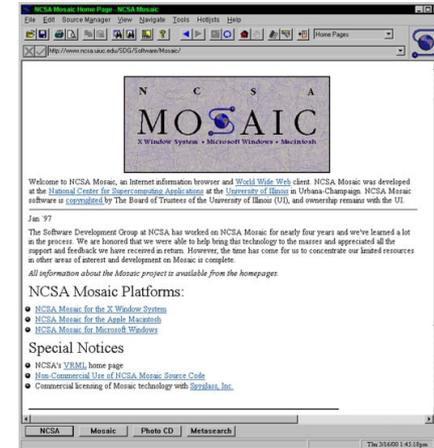
Invention du Web



Tim Berners-Lee, chercheur britannique, a inventé le Web au CERN en 1989. À l'origine, le projet, baptisé « World Wide Web », a été conçu et développé pour que des scientifiques travaillant dans des universités et instituts du monde entier puissent s'échanger des informations instantanément.

## 1993

Le 1<sup>er</sup> navigateur

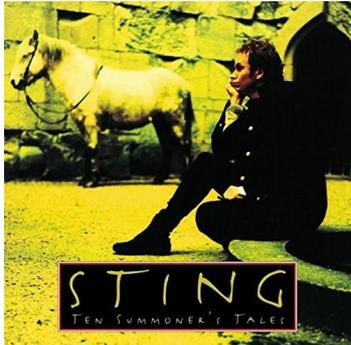


Le premier navigateur web graphique par une équipe de chercheurs du National Center for Supercomputing Applications de l'université de Illinois aux Etats Unis. Il est rebaptisé Mosaic Netscape en 1994.

# Repères historiques

## 1994

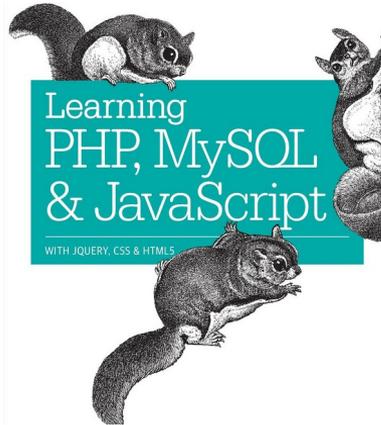
Commerce en ligne



Un Américain dénommé Dan Kohn avait lancé un site web appelé Net Market et sa première vente fut un CD de Sting. Plus tard en 1994, la compagnie Pizza Hut a également commencé à vendre des pizzas en ligne.

## 1995

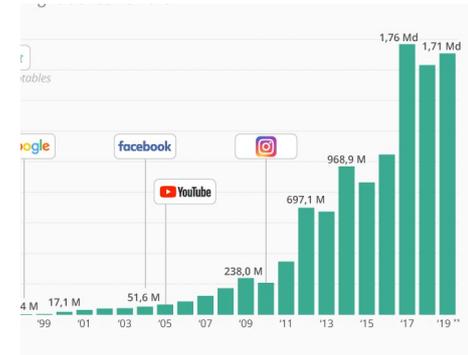
Web interactif et dynamique



En 1993, Mosaïc intègre des formulaires et des images. Les pages web s'enrichissent en 1995 en intégrant des langages comme PHP et JavaScript qui les rendent dynamiques

## 2014

1 milliard de sites



En septembre 2014, la barre du milliard de sites Web est franchie pour la première fois. En 2019 on compte près de 1,7 milliard de sites. Il y aurait plus de 4 milliards d'internautes.

# Notions juridiques

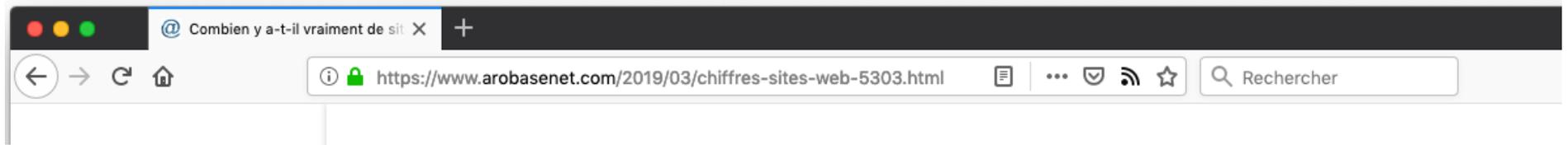
## Les droits sur internet

Tout ce qu'on trouve sur internet (textes, images, photos, etc) est la propriété de son auteur et ne peut être utilisé sans son autorisation.

Il existe les droits d'auteur en France, ou le *copyright* aux États-Unis. Mais la dimension mondiale du Web et l'impossibilité de contrôler le respect des différentes législations a donné naissance à de nouvelles pratiques, comme les licences *Creatives Commons*.

Licence Creative Commons	Bouton	Usage commercial permis?	Nouvelle version permise?
<b>Attribution</b>			
<b>Partage à l'identique</b> de la licence originale (Share Alike)			 Mais sous licence originale
<b>Aucune d'œuvre dérivée</b> (No Derivatives)			
<b>Non-commerciale</b> (Non-Commercial)			 La nouvelle œuvre et ses versions subséquentes devront être non-commerciale
<b>Non-commerciale</b> (Non-Commercial) + <b>Partage à l'identique</b> (Share alike)			 La nouvelle œuvre et ses versions subséquentes devront aussi être non-commerciale
<b>Non-commerciale</b> (Non-Commercial) + <b>Pas d'œuvre dérivée</b> (No Derivatives)			

# Le fonctionnement du Web : U.R.L.



## l'U.R.L.

Une URL (*Uniform Resource Locator*) est l'adresse d'une ressource d'un site Web. Elle indique où se trouve cette ressource dans l'arborescence du site.



# Le fonctionnement du Web : requêtes HTTP

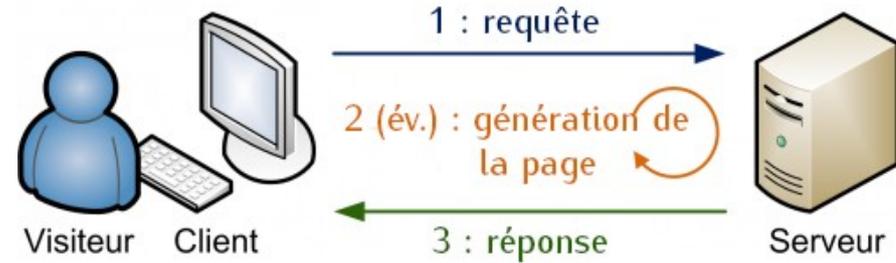
Le Web s'appuie sur un dialogue entre **clients** et **serveurs**.

Les clients sont des applications qui se connectent au Web, comme les navigateurs.



Ils envoient des **requêtes** aux serveurs où sont stockées les données.

Ces requêtes utilisent le protocole HTTP (*HyperText Transfert Protocol*), permettant aux ordinateurs de communiquer entre eux.



## Exemple de requête

```
GET /HTTP/1.1
Host : www.arobase.net
User-Agent : Mozilla/5.0 (Windows NT 10.0 ; WIN64 ; X64 ; rv:65.0) GECKO/200100101
Firefox/65.0
Accept : text/html
Accept-Language : fr-FR,fr,q=0.8,en-US;q=0.6,en;q=0.4
Referer : https://www.qwant.fr
```

Version du protocole

Nom de domaine du serveur

Identification du navigateur

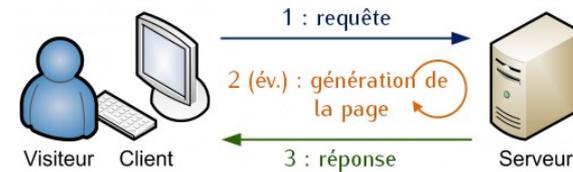
Indique que le client va recevoir du HTML

Indique la page d'où vient le navigateur

Langage accepté

# Le fonctionnement du Web : client / serveur

Une fois la requête reçue, le serveur renvoie une réponse (3).



## Exemple de réponse du serveur

```
HTTP/1.1 200 OK
Date: Thu, 15 feb 2019 12:02:32 GMT
Server: Apache/2.0.54 (Debian GNU/Linux) DAV/2 SVN/1.1.4
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html; charset=ISO-8859-1
<!doctype html>
<html lang="fr">
  <head>
    <meta charset="utf-8">
    <title>Voici mon site</title>
  </head>
  <body>
    <h1>Hello World! Ceci est un titre</h1>
    <p>Ceci est un <strong>paragraphe</strong>
    bien compris ? </p>
  </body>
</html>
```

En-tête de la réponse



# Les moteurs de recherche

Les moteurs de recherche sont indispensables au bon fonctionnement du Web. Ils permettent de trouver des informations dans des pages dont on ne connaît pas l'adresse, voir dont on ignore l'existence.



## Le cas de Qwant

Qwant est un moteur de recherche créé en France et lancé en version bêta le 16 février 2013.

Contrairement à Google, lorsque vous effectuez une recherche sur Qwant, aucun cookie de traçage n'est déposé sur votre navigateur. La recherche n'est pas stockée sur le serveur, ni associée à votre profil pour être réutilisée plus tard.

# COMMENT FONCTIONNE LA RECHERCHE GOOGLE



## ROBOT D'EXPLORATION

Des logiciels d'exploration, appelés "crawlers" ou "spiders", explorent régulièrement le web en parcourant les réseaux de liens. Ils scannent les pages qu'ils trouvent.

## ROBOT D'INDEXATION

Le robot d'indexation extrait les données des pages scannées (URL, titre, mots-clés sous forme de chaînes de caractères), puis les classe pour former l'index de la base de données de Google

## MOTEUR D'INTERROGATION

Interface entre l'internaute et l'index, le moteur d'interrogation récupère la requête tapée dans le formulaire, la compare aux données de l'index, puis recueille les résultats qu'il classe avant de les afficher. (voir l'algorithme de recherche Google)

# Les langages du web : le HTML

Contrairement à une page réalisée avec un logiciel de traitement de texte, on peut accéder au *code source* d'une page web. Ce code décrit le contenu de ce document.

**HTML** Les pages web sont écrites en HTML (HyperText Markup Language). C'est un langage de *balisage* : il permet d'insérer dans le document des balises qui définissent ce que signifient les différentes parties du texte, et de créer des liens.

Ce langage n'est pas **compilé** : il est **interprété** par le navigateur web.

Le principe consiste à encadrer un texte par deux balises complémentaires :

- avant le texte : **<balise>**
- après le texte : **</balise>**

Exemples :

- décrire un paragraphe :

```
<p>Contrairement à une opinion  
répandue, le Lorem Ipsum n'est pas...</p>
```

- décrire un titre de niveau 1 :

```
<h1>Qu'est ce que le Lorem Ipsum ?</h1>
```

- décrire un lien vers un site :

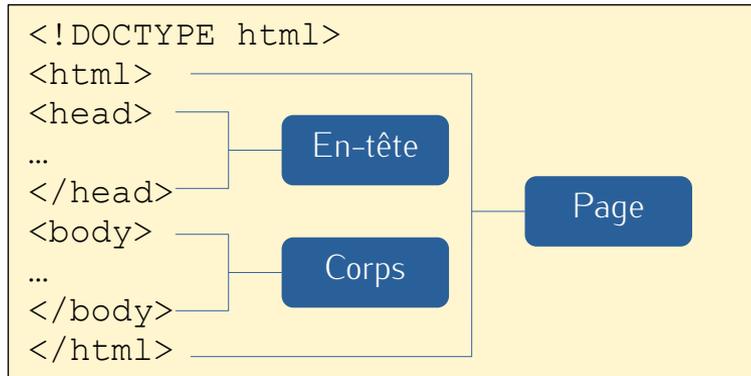
```
<a href="https://fr.lipsum.com">Le site  
Lorem ipsum</a>
```

Ces balises peuvent s'imbriquer :

```
<p>Un lien vers <a href="https://fr.lipsum.com">le site  
Lorem ipsum</a> sur lequel vous pouvez cliquer.</p>
```

# Les langages du web : le HTML

La forme globale d'un document HTML est la suivante :



## L'en-tête

Il regroupe des informations qui ne sont pas affichées par le navigateur dans la page :

- le titre de la page,
- des métadonnées (auteur, mots-clés,...),
- Les références aux feuilles de style,...

```

<head>
  <meta charset="utf-8">
  <!--On peut insérer des commentaires-->
  <title>Le lycée</title>
  <link rel="stylesheet" href="style.css">
</head>

```

# Les langages du web : le HTML

## La partie <body>

Elle contient l'ensemble de la page web qui sera affichée : le contenu et les balises.

Quelques balises :

 	Saut de ligne
<h1>	1 <sup>er</sup> niveau de titre, à utiliser une seule fois, pour le titre principal.
<h2> à <h6>	Titres de niveau 2 à 6.
<a>	Lien hypertexte
<img>	Image
<strong>	Renforcement

## Un premier code

```
<head>
  <meta charset="utf-8">
  <title>Une première page</title>
</head>
<body>
  <h1>Le titre principal</h1>
  <h2>Une liste non numérotée</h2>
  <ul>
    <li>Choix 1</li>
    <li>Choix 2</li>
    <li>Choix 3</li>
  </ul>
  <h2>Un premier texte</h2>
  <p>Ceci est un premier paragraphe
contenant du texte.<br>Ici un lien vers un
<a href="https://fr.lipsum.com"> site de
faux-texte.</a> pour créer des
paragraphe.</p>
  <h2>Une image</h2>
  
</body>
```

# Les langages du web : le HTML

## Avec quel logiciel saisir ce code ?

Avec n'importe quel *éditeur de texte* ! À ne pas confondre avec un *traitement de texte*..

Sous Windows, vous pouvez télécharger la version portable de Notepad++ :



Notepad++

- 1) Se rendre sur le site <https://notepad-plus-plus.org>
- 2) Cliquer sur Downloads puis sur la dernière « release ».
- 3) Choisir la version « zip package »
- 4) Dans Windows, décompresser le fichier et copier le dossier obtenu dans sa clé USB.

## Un premier code

```
<head>
  <meta charset="utf-8">
  <title>Une première page</title>
</head>
<body>
  <h1>Le titre principal</h1>
  <h2>Une liste non numérotée</h2>
  <ul>
    <li>Choix 1</li>
    <li>Choix 2</li>
    <li>Choix 3</li>
  </ul>
  <h2>Un premier texte</h2>
  <p>Ceci est un premier paragraphe
contenant du texte.<br>Ici un lien vers un
<a href="https://fr.lipsum.com"> site de
faux-texte.</a> pour créer des
paragraphe.</p>
  <h2>Une image</h2>
  
</body>
```

# Les langages du web : le HTML

## Un premier code

```
<head>
  <meta charset="utf-8">
  <title>Une première page</title>
</head>
<body>
  <h1>Le titre principal</h1>
  <h2>Une liste non numérotée</h2>
  <ul>
    <li>Choix 1</li>
    <li>Choix 2</li>
    <li>Choix 3</li>
  </ul>
  <h2>Un premier texte</h2>
  <p>Ceci est un premier paragraphe
contenant du texte.<br>Ici un lien vers un
<a href="https://fr.lipsum.com"> site de
faux-texte.</a> pour créer des
paragraphe.</p>
  <h2>Une image</h2>
  
</body>
```

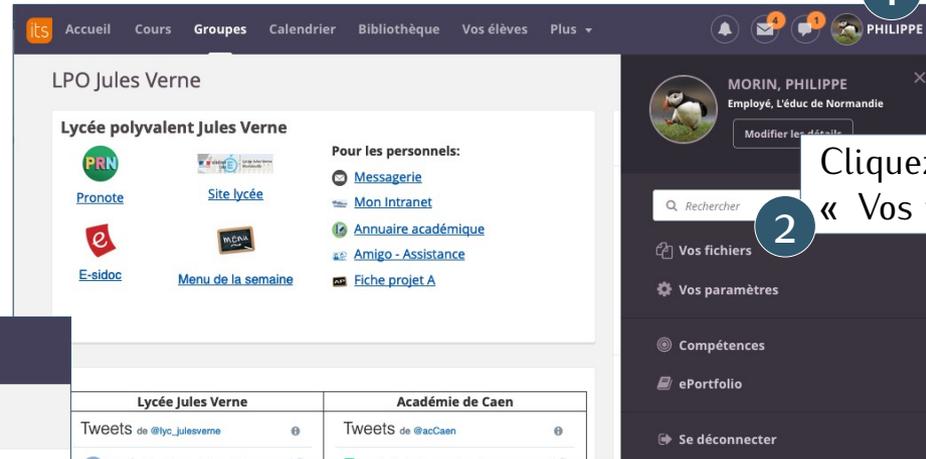
Saisir ce code.  
Enregistrer le fichier `essai.html`.  
Ouvrir ce fichier avec Firefox.



# Les langages du web : l'hébergement

Le fichier `essai.html` n'est accessible que par...celui qui possède le fichier !  
Il faut donc que celui-ci soit copié dans un dossier accessible sur internet.  
L'ENT académique le permet.

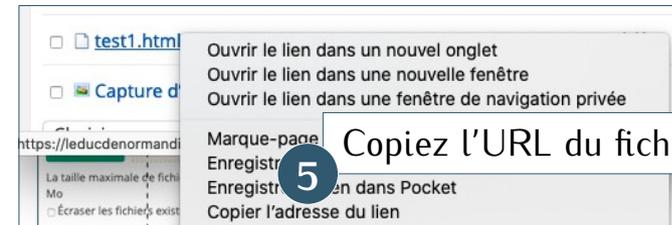
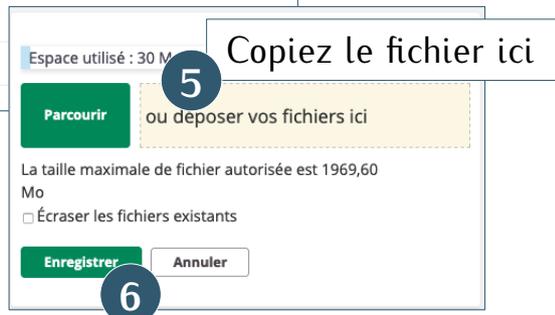
1 Cliquez sur votre prénom



2 Cliquez sur « Vos fichiers »



4 + Charger le fichier



# Les langages du web : le style CSS

Le langage *CSS* (*Cascading Style Sheets*, feuilles de style en cascade) permet de définir la manière dont sont affichées les éléments de la page en leur appliquant des styles : arrière-plan, police, marges, bordures,...



« Le fond »

+



« La forme »

## style.css

```
body {
  background-color:black ;
  font-family:calibri ;
}

h1 {
  background-color:#2e6ed6 ;
  color:#ele8f2 ;
  text-align:center ;
}

h2 {
  color:orange ;
}

p {
  color:white ;
}
```

À ajouter dans la partie « head » du code HTML :

```
<link rel="stylesheet" href="style.css">
```