

# Le Web

## SNT

Frédéric Junier

Lycée du Parc  
1 Boulevard Anatole France  
69006 Lyon

2 janvier 2022

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# Internet, un réseau informatique

## Définition

- Un réseau informatique est un ensemble de noeuds (pc, mobiles . . . ) qui peuvent communiquer entre eux par des liens de natures diverses (fils, cables sous-marins, ondes . . . ).
- **Internet** désigne une interconnexion mondiale de réseaux d'ordinateurs.

# Internet, un réseau informatique

## Définition

- Chaque noeud doit disposer d'une adresse unique sur le réseau pour communiquer avec les autres. Sur **Internet** il s'agit d'une **adresse IP**.
- Des règles universelles, appelées **protocoles** permettent aux noeuds de réseaux basés sur des technologies hétérogènes (Ethernet, Wifi, 4G . . . ) d'échanger sur **Internet**. Sur **Internet** il s'agit de la pile de protocoles d'une **TCP/IP**.

## Naissance d'Internet (1/2)

- Les premières expériences de connexions d'ordinateurs en réseaux ont lieu dans les années 60.
- Le 19 octobre 1969 a lieu par exemple la première connexion entre des ordinateurs d'universités distinctes : Stanford et UCLA.

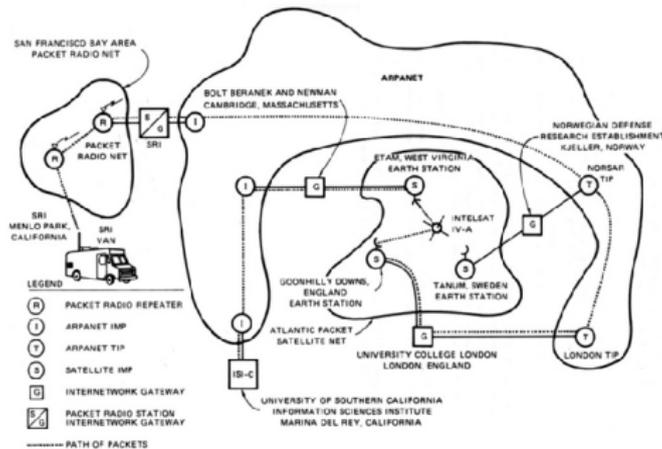
# Naissance d'Internet (1/2)

12900769	100	LOADED OP. PROGRAM FOR BEN BARKER BRV	CSK
22:30		Talked to SRI Host to Host	CSK
		Leftop. program running after sending a host dead message to imp.	CSK

## Naissance d'Internet (2/2)

- Dans les années 1970, **Vinton Cerf et Robert Kahn**, s'inspirant des travaux de **Louis Pouzin** développent la pile de protocole TCP/IP.
- À partir de 1983, TCP/IP permet d'interconnecter les réseaux ARPANET et CSNET donnant naissance à l'**Internet**.
- **Internet** reste principalement utilisé par des chercheurs et des acteurs institutionnels.

# Naissance d'Internet (2/2)



Une histoire d'Internet.

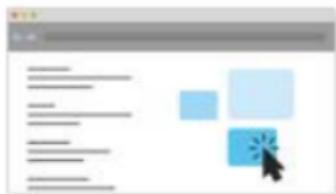
# Le Web, un réseau de documents

## Définition

- Le **Web** est un réseau de documents reliés par des liens hypertextes.
- Le Web est un service d'Internet basé sur une **architecture client / serveur** : un serveur fournit les ressources qui lui sont demandées par des clients. Un logiciel client est un **navigateur** : par exemple Firefox, Chrome, Edge, Safari ...

! Non sécurisé | http://exemplesite.fr

Client



Requête HTTP



Réponse HTTP

Serveur



source : « Le livre scolaire »

# Le Web, un réseau de documents

## Définition

- Clients et serveurs Web communiquent avec le protocole **HyperText Transfer Protocol (HTTP)**.
- Chaque ressource du Web a une adresse unique appelée **Uniform Ressource Locator (URL)**.

## Histoire du Web (1/2)

- En 1989, **Tim Berners-Lee**, qui s'occupe de la gestion des documents scientifiques au **CERN**, propose avec le projet *mesh* (filet) de faire traverser le réseau internet à **l'hypertexte** imaginé par **Ted Nelson** en 1965.
- En 1993, les dirigeants du CERN annoncent officiellement que la technologie du Web sera gratuite et libre de droits. Microsoft essaie de développer des formats propriétaires concurrents mais en vain (la première version de Windows 95 était livrée sans navigateur Web !)
- À partir de cette date, le Web se développe dans le grand public pour le partage de connaissances mais aussi pour le commerce et la publicité.

## World Wide Web

The WorldWideWeb (W3) is a wide-area [hypermedia](#) information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an [executive summary](#) of the project, [Mailing lists](#), [Policy](#), November

### [What's out there?](#)

Pointers to the world's online information, [subjects](#), [W3 servers](#), etc.

### [Help](#)

on the browser you are using

### [Software Products](#)

A list of W3 project components and their current state. (e.g. [Line Mode](#), [X11 Viola](#), [NeXTStep](#), [Servers](#), [Tools](#), [Mail robot](#), [Library](#))

### [Technical](#)

Details of protocols, formats, program internals etc

### [Bibliography](#)

Paper documentation on W3 and references.

### [People](#)

A list of some people involved in the project.

### [History](#)

A summary of the history of the project.

### [How can I help?](#)

If you would like to support the web..

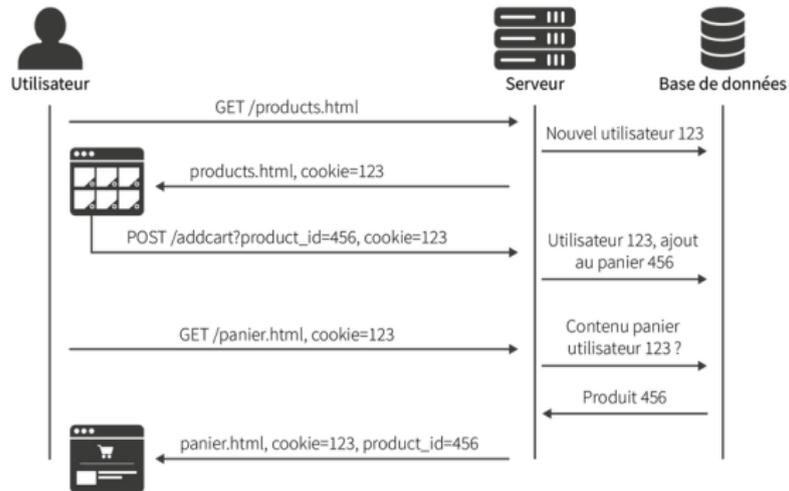
### [Getting code](#)

Getting the code by [anonymous FTP](#), etc.

Lien hypertexte vers la première page Web !

## Histoire du Web (2/2)

- En 1994, le système des *cookies* petits fichiers textes échangés entre client et serveur, est intégré dans le navigateur Netscape afin de conserver des informations sur le client (panier) lors de sa navigation sur un site Web.
- En 2000, le chiffre d'affaire de la publicité sur le Web est de 8 milliards d'euros.
- Les années 2000 sont celles du Web 2.0 ou Web social avec l'apparition des blogs, des réseaux sociaux et l'extension du Web à la mobilité avec le succès des ordiphones.



## Une histoire des cookies

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# QCMS

- Faire le **Quizz 1**
- Faire le **Quizz 2**

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web**
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# URL

Les ressources (ou pages) sur le Web sont adressées par une **Uniform Ressource Locator(URL)**.

Une URL comporte plusieurs informations, par exemple décomposons

<https://www.gnu.org/gnu/linux-and-gnu.fr.html>

- le protocole est `https` ;
- le nom de domaine sur Internet de la machine est `www.gnu.org`, un nom de domaine est un alias humainement lisible pour l'adresse IP du serveur (service d'annuaire **DNS** d'Internet) ;
- `gnu/linux-and-gnu.fr.html` est le chemin vers la ressource sur le serveur : le fichier `linux-and-gnu.fr.html` qui se trouve dans le dossier `gnu`.

# URL

<https://eduscol.education.fr/numerique>

**Protocole** (*HyperText Transfer Protocol Secure*) de communication

**Adresse symbolique** : serveur où est stockée la ressource

**Chemin d'accès à la ressource demandée** : exprimé depuis la racine du domaine

source : « *Le livre scolaire* »

# HTTP

Lors d'un échange Web, le **client** et le **serveur** dialoguent selon des règles fixées par un protocole standardisé, **HyperText Transfer Protocol**.

Si la page demandée n'est pas disponible, le serveur retourne le célèbre code d'erreur 404. D'autres codes d'erreurs qui ne manquent pas d'humour : 418 « *I'm a teapot* » ou 451 : « *unavailable for legal reasons (censorship)* » => référence à « *La vie de Brian* » des *Monty Python*.

# HTTP code 451

HTTP/1.1 451 Unavailable For Legal Reasons

Link: <https://sqr.example.org/legislation>; rel="blocked-by"

Content-Type: text/html

```
<html>
<head><title>Unavailable For Legal Reasons</title></head>
<body>
<h1>Unavailable For Legal Reasons</h1>
<p>This request may not be serviced in the Roman Province
of Judea due to the Lex Julia Majestatis, which disallows
access to resources hosted on servers deemed to be
operated by the People's Front of Judea.</p>
</body>
</html>
```

# HTTP : exercice

- 1 Accéder depuis un navigateur à la page Web <http://frederic-junier.org/NSI/sandbox/index.html>.  
Ouvrir la barre d'outils de développement en appuyant sur la touche de fonction F12 et sélectionner l'onglet Réseau.  
Analyser les entêtes de la requête et de la réponse HTTP.
- 2 Le site <https://httpie.org/> propose un client HTTP en ligne de commandes permettant de décomposer les requêtes HTTP en précisant la méthode et l'URL de la ressource demandée.
  - Saisir la commande `http -v GET https://frederic-junier.org/NSI/sandbox/index.html`
  - Décrire le fonctionnement de la méthode GET du protocole HTTP : distinguer la requête et la réponse.

# HTTP

```
- $ http -v GET https://frederic-junier.org/NSI/sandbox/index.html
GET /NSI/sandbox/index.html HTTP/1.1
Accept: */*
Accept-Encoding: gzip, deflate
Connection: keep-alive
Host: frederic-junier.org
User-Agent: HTTPie/2.3.0

HTTP/1.1 200 OK
Accept-Ranges: bytes
Connection: Upgrade, Keep-Alive
Content-Length: 555
Content-Type: text/html
Date: Tue, 20 Apr 2021 21:19:38 GMT
Keep-Alive: timeout=10, max=100
Last-Modified: Tue, 21 Apr 2020 13:26:14 GMT
Server: Apache
Upgrade: h2,h2c
Vary: Accept-Encoding

<!DOCTYPE html>

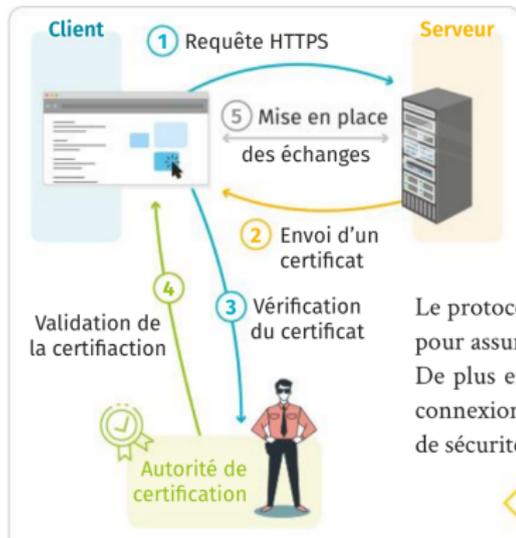
<html lang="fr">

<head>
  <title>Protocole HTTP </title>
  <meta charset="utf-8">
</head>

<body>

  <h1> HTTP le protocole du Web. </h1>
```

# HTTPS



## 2 Protocole HTTP et échange client/serveur

Le protocole HTTPS (*HyperText Transfer Protocol Secure*) a été mis au point pour assurer plus de sécurité dans les dialogues entre le client et le serveur. De plus en plus utilisé, il assure des garanties supplémentaires lors de la connexion. Ce protocole repose sur la vérification systématique de certificat de sécurité délivré par une entité appelée Autorité de certification.

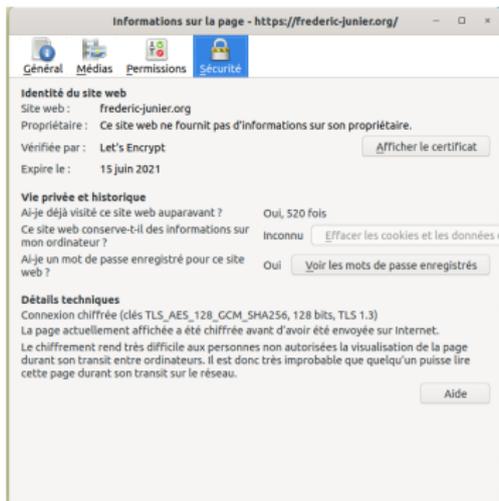
## 3 Protocole HTTP Sécurisé (HTTPS)

source : « *Le livre scolaire* »

# HTTPS : exercice

- 1 Depuis un navigateur, accéder à la page <http://frederic-junier.org/> : comment s'affiche le cadenas devant l'URL dans la barre d'adresse ? Quelle est sa signification.
- 2 Depuis un navigateur, accéder à la page <https://frederic-junier.org/> : comment s'affiche le cadenas devant l'URL dans la barre d'adresse ? Quelle est sa signification.
- 3 Corriger ses réponses à l'aide de la page <https://support.mozilla.org/fr/kb/comment-savoir-si-ma-connexion-est-securisee>. Que désignent les expressions « *blocage de contenu mixte* » et « *attack man of the middle* » ?

# HTTPS : infos du cadenas



# HTML

## Définition

Les **premières pages Web** sont des **pages statiques** : la page retournée par le serveur ne dépend pas du client. Le premier langage du Web est **HyperText Markup Language** qui est un langage de description de documents à l'aide de **balises**.

HTML permet de définir la structure d'un document en le découpant en blocs logiques délimités par une **<balise>** ouvrante et une **</balise>** fermante. En particulier il permet de créer des liens hypertextes entre deux blocs d'un même document ou de documents différents.

# HTML

Code source HTML

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
  </head>
  <body>
    <h1 class="titre">Langage HTML</h1>
    <h2 class="titre">Recommandation</h2>
    <p>Un <a id="monsite" href="https://frederic-junier.org/wp">Super site</a>
      à visiter absolument ! </p>
    <!-- Commentaire : sauts de ligne et indentation
      pas affichés mais
      facilitent la relecture -->
  </body>
</html>
```

Affichage dans le navigateur

## Langage HTML

### Recommandation

Un [Super site](https://frederic-junier.org/wp) à visiter absolument !

# Plan

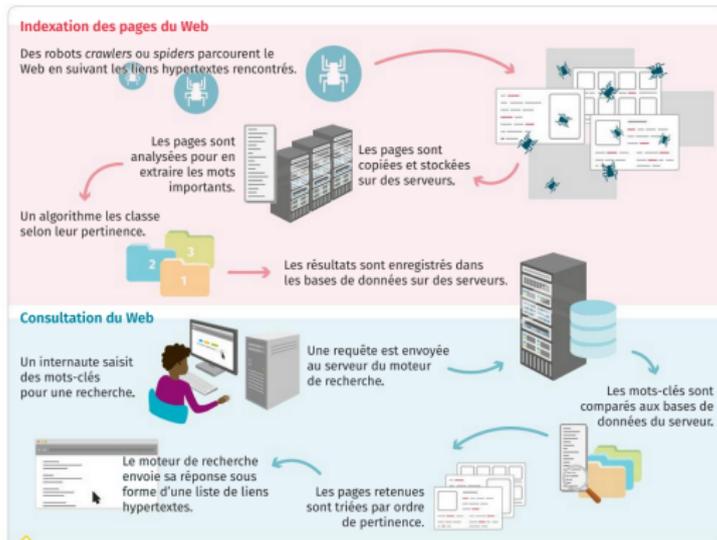
- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche**
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# Moteurs de recherche

Le Web est devenu un gigantesque hyertextet et pour y chercher une information, de puissants logiciels, les **moteurs de recherches** sont nécessaires. Le site Web le plus visité est celui du moteur de recherche *Google search*, qui traite plusieurs milliards de requêtes par jour. Le nom Google a d'ailleurs pour origine *googol* qui désigne le nombre  $10^{100}$  (supérieur au nombre d'atomes dans l'univers!).

Une présentation en video et une activité Scratch sur l'algorithme Page Rank sur [https://frederic-junier.org/SNT/Theme1\\_Web/SNT\\_Activite4\\_Web.html](https://frederic-junier.org/SNT/Theme1_Web/SNT_Activite4_Web.html)

# Moteurs de recherche



1 Fonctionnement d'un moteur de recherche

source : « Le livre scolaire »

## Moteurs de recherche : exercice

- 1 Dans les moteurs de recherche **Google** et **Qwant**, saisir la requête `super ligue`. Obtient-on les mêmes résultats ? Comment repérer les liens sponsorisés ? Les URL des résultats de **Google** ont une particularité, laquelle ?
- 2 Comparer les résultats obtenus avec les requêtes : `super -ligue` puis `super OR ligue`.
- 3 Changer le moteur de recherche par défaut dans les options de votre navigateur.

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web**
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# HTML et CSS

Les **pages statiques** sont écrites en HTML et en **CSS** qui est un langage de feuille de style utilisé pour décrire la présentation d'un document écrit en HTML. La standardisation du CSS a souffert de la guerre des navigateurs à la fin des années 1990 mais CSS s'est imposé ensuite comme un complément indispensable de HTML. HTML décrit la structure du document et CSS définit son apparence. Le fond et la forme sont ainsi séparés. D'ailleurs, le code CSS, même s'il peut être intégré au code HTML, est plutôt dans une feuille de style séparé liée au document HTML par une balise `<link rel="stylesheet" href="feuille_style.css">`.

# HTML

1	<code>&lt;!DOCTYPE html&gt;</code>	→ Indique au navigateur les règles de mise en forme
2	<code>&lt;html&gt;</code>	→ Début du document html
3	<code>&lt;head&gt;</code>	→ Début de l'en-tête (description du document)
4	<code>&lt;link rel="stylesheet" href="style.css"/&gt;</code>	→ Lien vers la feuille de style CSS associée
5	<code>&lt;title&gt;Titre&lt;/title&gt;</code>	→ Nom visible dans l'onglet du navigateur
6	<code>&lt;/head&gt;</code>	→ Fin des informations de description du document
7	<code>&lt;body&gt;</code>	→ Début du corps de la page (éléments de la page)
8	<code>&lt;h1&gt;Ma première page&lt;/h1&gt;</code>	→ Titre de premier niveau
9	<code>&lt;img src="image.png"&gt;</code>	→ Image et lien
10	<code>&lt;p&gt;Une belle image&lt;/p&gt;</code>	→ Paragraphe
11	<code>&lt;/body&gt;</code>	→ Fin du corps de la page
12	<code>&lt;/html&gt;</code>	→ Fin du document html

1 Exemple de code HTML

source : « *Le livre scolaire* »

# HTML et CSS

Code source HTML

```
<!DOCTYPE html>
<html>
  <head>
    <meta charset="utf-8">
    <link href="style.css" rel="stylesheet" type="text/css" />
  </head>
  <body>
    <h1 class="titre">Langage HTML</h1>
    <h2 class="titre">Recommandation</h2>
    <p>Un <a id="monsite" href="https://frederic-junier.org/wp">Super site</a> à visiter absolument ! </p>
  </body>
</html>
```

Code CSS

```
p {
  font-style: normal;
}

#monsite {
  font-style: italic;
}

.titre {
  border: 2px solid black;
}
```

Affichage obtenu

**Langage HTML**

**Recommandation**

Un [Super site](https://frederic-junier.org/wp) à visiter absolument !

# HTML : exercice

Se connecter à l'ENT et accéder à l'exercice proposé dans le service Capytale (ressources numériques) avec le code **2f4c-15282**.  
Voici l'URL d'accès direct au service (connexion ENT nécessaire) :

<https://capytale2.ac-paris.fr/>

# CSS : exercice

Se connecter à l'ENT et accéder à l'exercice proposé dans le service Capytale (ressources numériques) avec le code **3838-15334**. Voici l'URL d'accès direct au service (connexion ENT nécessaire) :

<https://capytale2.ac-paris.fr/>

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques**
- 7 QCM de synthèse

## Cookies : document

Extrait de « *Cristaux liquides* » de Feu !Chatterton (2021).

*Oh, j'ai un cookie*

*Les navigateurs veulent mieux te connaître*

*Fais-leur un bisou*

*Fais-leur coucou*

*Oh, j'ai un cookie*

*Les navigateurs passent par la fenêtre*

*Les navigateurs sont à la fenêtre*

*Fais-leur coucou*

*Les navigateurs veulent mieux te connaître*

## Cookies : exercice

Vous utiliserez le navigateur Firefox.

- 1 Donnez votre définition d'un cookie de navigateur.
- 2 Comparez avec celle donnée dans <https://linc.cnil.fr/fr/une-petite-histoire-du-cookie>. Pourquoi les cookies ont-ils été inventés ?
- 3 Cliquez sur le cadenas devant l'URL pour effacer les cookies déjà déposés par le site dans votre navigateur. Rafraichissez la page, que se passe-t-il ? Quelles sont les informations données lorsque vous cliquez sur le bouclier à gauche de l'URL ?
- 4 Installez dans votre navigateur l'extension <https://privacybadger.org/> et analysez les traceurs listés pour des sites couramment visités.

## Mode de navigation privée : exercice

Vous utiliserez le navigateur Firefox.

- 1 Comment accède-t-on au mode de navigation privé ?
- 2 Permet-il de naviguer de façon anonyme ? et une connexion sécurisée en HTTPS ?
- 3 Quelle distinction faites-vous entre anonymat et pseudonymat ? Dans quelles situations l'anonymat sur le Web pourrait-il être important ?
- 4 Connaissez-vous des moyens de naviguer de façon anonyme ?

## Droit d'auteur : exercice

Vous utiliserez le navigateur Firefox.

- 1 Quelles sont les conditions de réutilisation de l'image de la Joconde téléchargeable sur le [site du musée du Louvre](#)? Et pour les contenus textuels de la notice?
- 2 Quelles sont les conditions de réutilisation des cours disponibles sur <https://frederic-junier.gitlab.io/parc-nsi/>?

# Licences Creative Commons

Licence CC	Dénomination
	Domaine public
	Zéro
	Attribution
	Partage avec la même licence
	Pas de modification
	Pas d'utilisation commerciale
	Pas d'utilisation commerciale Partage avec la même licence
	Pas d'utilisation commerciale Pas de modification

## 6 Licences Creative Commons

source : « *Le livre scolaire* »

# Plan

- 1 Web et Internet
- 2 Exercice
- 3 Les trois piliers du Web
- 4 Les moteurs de recherche
- 5 Les langages du Web
- 6 Aspects sociaux et juridiques
- 7 QCM de synthèse

# QCM

Répondre à ce QCM :

<https://www.quiziniere.com/#/Exercice/2BMZNY>