

Exercice : Requête HTTP

Exercice 1

On donne la capture des en-têtes d'une requête HTTP effectuée par un navigateur.

- 1) Quelle est la méthode utilisée par cette requête ?
- 2) La ressource demandée a-t-elle été renvoyée par le serveur ? Justifier.
- 3) Quelle est la taille des données ?
- 4) Quel est le type de données envoyées par le serveur ?
- 5) Quel est le domaine du site visité ?
- 6) Quelle est l'URL affichée dans la barre d'adresse du navigateur ?
- 7) Quel est le serveur web utilisé par ce site ?

▼ En-têtes de la requête (520 o)

```
GET /le-modele-turing/ HTTP/1.1
Host: interstices.info
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64;
Accept: text/html,application/xhtml+xml,application/
Accept-Language: fr,fr-FR;q=0.8,en-US;q=0.5,en;q=0.3
Accept-Encoding: gzip, deflate, br
```

▼ En-têtes de la réponse (576 o)

```
HTTP/1.1 200 OK
Date: Wed, 17 Mar 2021 16:49:10 GMT
Server: Apache
X-Frame-Options: SAMEORIGIN
Expires: Thu, 19 Nov 1981 08:52:00 GMT
Cache-Control: private, must-revalidate
Pragma: no-cache
X-Pingback: https://interstices.info/xml
Link: <https://interstices.info/wp-json/
Link: <https://interstices.info/?p=15005
Vary: Accept-Encoding,User-Agent
Content-Encoding: gzip
X-XSS-Protection: 1; mode=block
X-Content-Type-Options: nosniff
Content-Length: 18220
Connection: close
Content-Type: text/html; charset=UTF-8
```

Exercice 2

On donne la capture d'une requête HTTP effectuée par un navigateur.

- 1) Quelle est la méthode utilisée par cette requête ?
- 2) Des données ont été envoyées au serveur par le client. Combien et comment ?
- 3) Quelle réponse a été fournie par le serveur et sous quelle forme ?
- 4) On peut effectuer cette requête avec la méthode GET. Comment se réalise-t-elle ?

► POST https://www.httpbin.org/post

État	200 OK ?
Version	HTTP/2
Transfert	1,17 Ko (taille 954 o)
Politique de référent	unsafe-url

▼ En-têtes de la réponse (246 o)

```
? access-control-allow-credentials: true
? access-control-allow-origin: https://www.httpbin.org
? content-length: 954
? content-type: application/json
? date: Wed, 17 Mar 2021 16:42:19 GMT
? server: gunicorn/19.9.0
```

▼ Données de formulaire

```
page: "numerique_futur"
titre: "cryptographie"
auteur: "yvan"
```

Exercice 3

Le logiciel **curl** ou **cURL** est une application multi-plateforme. Cette application est un client HTTP qui effectue des requêtes HTTP en ligne de commande.

Cette application s'exécute avec différentes options :

- **curl -i** : Affiche le contenu complet de la réponse HTTP. On visualise ainsi la ligne de statut et les en-têtes retournés par le serveur.
- **curl -I** : Affiche l'en-tête de la réponse HTTP retourné par le serveur.
- **curl -v** : Mode verbeux affichant notamment la requête émise et la réponse reçue.
- **curl -X [méthode]** : Permet de préciser la méthode HTTP à utiliser dans la requête (GET par défaut).
- **curl -H "[en-tête]"** : Permet d'ajouter ou de modifier un en-tête de la requête.
- **curl --data-binary [data]** : Ajoute les données spécifiées comme corps du message
- **curl -o [nom de fichier]** : Permet de sauvegarder la réponse retournée par le serveur dans un fichier.

- 1) Ouvrez l'application **cmd** de windows et vérifiez que le logiciel **curl** est bien installé en tapant la commande **curl --version**.
- 2)
 - a) Effectuer une requête HTTP vers l'adresse <https://interstices.info> et afficher l'en-tête de la réponse HTTP.
 - b) Enregistrer la page d'accueil du site avec curl et ouvrez-la dans un navigateur.
- 3) Choisissez une image ou photo sur le site **wikimedia commons** et en récupérez son url.
 - a) Effectuer une requête HTTP avec curl sur cette image pour en récupérer sa taille en octets.
 - b) Enregistrer cette image avec curl. Vérifier le fichier et sa taille en octets.
 - c) Enregistrer l'image en utilisant le navigateur. A-t-on la même taille ?
- 4) Parmi les adresses suivantes, donner les codes d'état des réponses HTTP suite à une requête HTTP avec curl sur les url suivantes :
 - a) <http://oups.fr>
 - b) <https://faillite.com>
 - c) <http://cretin.fr>
 - d) <https://www.fier-comme-un-coq.com/>
 - e) <https://developer.mozilla.org/be/>

Exercice 4

- 1) Aller sur le site **interstices.infos** puis effectuer une recherche sur le mot **internet** (avec la loupe).
- 2) On va effectuer cette même requête avec le client **curl**.
 - a) Afficher dans un premier temps l'en-tête de la réponse pour vérifier son succès.
 - b) Enregistrer dans le fichier **internet.html** la page web renvoyée par le serveur et vérifier qu'elle est identique à celle affichée par le site.
- 3) On affine la recherche en sélectionnant le domaine **algorithmes** et le type de contenu **articles**. Réaliser cette recherche sur le site.
- 4) On va effectuer cette même requête avec le client **curl**.
 - a) Effectuer une requête avec la méthode GET. On affichera l'en-tête de la réponse pour vérifier son succès.
 - b) Effectuer la requête en utilisant la méthode POST avec l'url "<https://interstices.infos>".
On enregistrera dans le fichier **internet.html** la page web renvoyée par le serveur et on vérifiera qu'elle est identique à celle affichée par le site.

Exercice 5

Le module **requests** de Python permet de réaliser des requêtes HTTP.

On utilisera principalement les méthodes suivantes :

- **get** et **post** pour effectuer nos requêtes HTTP :
rep=requests.get("url_du_site") ou rep=requests.post("url_du_site")
- **text** permet de récupérer le contenu renvoyé par le serveur.
print(rep.text) affiche le contenu.
- **headers** donne le contenu de l'en-tête de la réponse HTTP renvoyé par le serveur. print(rep.headers) affiche l'en-tête de la réponse HTTP.

Vous trouverez la documentation sur le web à l'adresses <https://fr.python-requests.org/en/latest/user/quickstart.html#contenu-de-la-reponse>.

Nous allons effectuer les mêmes requêtes que l'exercice précédent en Python.

Vous pouvez utiliser le **notebook jupyter** ou l'idle **Thonny** (il sera peut-être nécessaire d'installer le paquet requests).

- 1) Importer le module **requests** dans votre feuille de programmes.
- 2) Réaliser une première requête HTTP sur le site **interstices.infos** avec la méthode **get**. Afficher le code d'état de la réponse HTTP.
 - a) Enregistrer dans la variable **entete** l'en-tête de réponse du serveur.
 - b) Enregistrer dans la variable **page** le contenu de la réponse du serveur.
 - c) Quelle est la taille et le type du contenu renvoyé par le serveur ?
- 3) La variable **page** contient le contenu de la page web. On donne ci-dessous une fonction qui écrit dans un fichier le contenu d'une variable.

```
def page_html(text):  
    with open('ma_page.html', 'w', encoding='utf8') as page:  
        page.write(text)  
        page.close()
```

- a) Recopier cette fonction sur votre feuille de programmes.
 - b) Créer le fichier **page.html** avec le contenu de la variable **page** puis vérifier qu'elle s'affiche correctement dans un navigateur.
- 4) Effectuer une requête GET en passant en paramètre le mot **internet**. Afficher le code d'état de la réponse HTTP.
 - a) Enregistrer dans la variable **page** le contenu de la réponse du serveur.
 - b) Afficher dans le navigateur le contenu de la variable **page**.
- 5) On va effectuer une requête POST en passant les paramètres de recherche suivants :
 - mot cherché : internet
 - domaine : algorithmes
 - niveau de lecture facile

Pour réaliser cette requête, il faut au préalable enregistrer dans une variable les paramètres de recherche en respectant la structure suivante : 'clef 1' : 'valeur 1', 'clef 2' : 'valeur 2', ... , 'clef n' : 'valeur n' où les clefs et les valeurs correspondent aux paramètres de recherche.

- a) Enregistrer dans la variable **p** les paramètres de recherche.
- b) Effectuer une requête POST en passant en paramètre la variable **p**.
- c) Enregistrer dans la variable **page** le contenu de la réponse HTTP du serveur.
- d) Afficher dans le navigateur le contenu de la variable **page**.