**TP n°1 : Découverte de HTTP**

L’objectif de ce TP est de vous faire découvrir, par la pratique, la manière dont un client Web (votre navigateur) et un serveur, échangent des données sur le Web, *via* le protocole HTTP.

**La console de développeur**

Pour ce TD, vous allez vous familiariser avec la console de développeur, disponible dans la plupart des navigateurs modernes. Les instructions qui suivent concernent Firefox et Chrome ; sur d’autres navigateurs, YMMV.

| **Firefox** | **Chrome** |
| --- | --- |
| [Console développeur sous Firefox](https://perso.liris.cnrs.fr/pierre-antoine.champin/2019/progweb-python/_images/console_firefox.png) | [Console développeur sous Chrome](https://perso.liris.cnrs.fr/pierre-antoine.champin/2019/progweb-python/_images/console_chrome.png) |

* Pour ouvrir/fermer la console, tapez F12.
* La console peut apparaître à droite de la page, plutôt qu’en dessous.
* La console comporte plusieurs onglets. Celui qui nous intéresse dans cette séance est l’onglet Network/Réseau.
* Assurez vous que la case “Disable cache” n’est **pas** cochée.

**Navigation**

* Ouvrez [cette page](https://perso.liris.cnrs.fr/pierre-antoine.champin/2019/progweb-python/_static/td1/) **dans un nouvel onglet**.
* Dans cet onglet, ouvrez la console sur l’onglet navigation, puis cliquez sur le lien vers la page n°1.
* Vous constaterez qu’un certain nombre de lignes apparaissent dans la console. Chaque ligne représente un échange entre votre navigateur et un serveur.
* En passant le pointeur de la souris sur la colonne File/Fichier (sous Firefox) ou Name/Nom (sous Chrome), vous voyez apparaître l’URL complète.

**Questions**

* + Faites une copie d’écran.
  + Combien d’échanges constatez vous lorsque vous cliquez sur le lien ?
  + Comment expliquez-vous ce nombre ? Vous pouvez vous aider en affichant le source HTML de la page (CTRL+U).
* Cliquez maintenant sur le lien vers la page n°2.

**Questions**

* + Faites une copie d’écran.
  + Combien d’échanges constatez vous cette fois ?
  + Quelles différences voyez-vous par rapport à la page précédente ? (indice : regardez notamment la colonne Status/État)
  + Que signifie le code 404 ?
* Cliquez maintenant sur le lien vers la page n°3.

**Questions**

* + Cela confirme-t-il votre réponse sur le code 404 ?
* Revenez en arrière pour ré-afficher la page n°2, et cliquez sur le lien vers la page n°1.

**Questions**

* + Faites une copie d’écran.
  + Comparez le résultat avec votre première copie d’écran.
  + Recharchez la page avec CTRL+R. Que constatez-vous ? Regardez notamment dans les colonnes Status/État, et Transferred/Transfert (Firefox) ou Size/Taille (Chrome).
  + Recharchez la page avec CTRL+SHIFT+R. Que constatez-vous ?

**Indice**

Le *cache* est une mémoire dans laquelle le navigateur stocke temporairement les informations récupérées sur le réseau, afin de gagner du temps lorsqu’il doit les réutiliser plusieurs fois de suite.

Nous reparlerons plus en détail de cette notion dans une prochaine séance.

**Requête et réponse**

* Chaque échange entre le navigateur et le serveur est constitué de deux messages :
  + une **requête** (*request*), envoyée par le navigateur au serveur,
  + suivie d’une **réponse** (*response*), envoyée par le serveur au navigateur.
* Recharchez la page avec CTRL+SHIFT+R, puis cliquez sur la première ligne (page1.html). Un nouveau panneau apparaît, affichant le détail de la requête et de la réponse, et comportant plusieurs onglets.
* Cliquez sur l’onglet Response/Réponse. Cet onglet contient le contenu de la réponse du serveur.

**Questions**

* + Le contenu de l’onglet Réponse correspond-il à ce que vous attendiez ?
  + À votre avis, que contiendra cette onglet pour la ligne style.css ?
  + Cliquez maintenant sur la ligne style.css. Le contenu de l’onglet Réponse correspond-il à ce que vous attendiez ?
* Cliquez sur l’onglet Headers/En-têtes. Cet onglet les informations complémentaires includes dans les messages de requête et de réponse.

**Questions**

* + Identifiez les trois parties de cet onglet, contenant 1) les informations générales, 2) les en-têtes de la réponse, 3) les en-têtes de la requête.
  + Essayez de deviner, d’après leurs noms, lequel des en-têtes de requête (envoyés par le navigateur), indique au serveur quelle est la langue préférée de l’utilisateur.
  + Essayer de deviner, d’après leurs noms lequel des en-têtes de réponse (envoyés par le serveur), indique au navigateur la date de dernière modification du fichier envoyé.
  + Quelle est la méthode utilisée par cette requête (Request method/Méthode de requête) ? Voyez vous des requêtes utilisant d’autres méthodes ?

**Indice**

Au fur et à mesure de l’avancée du cours, nous découvrirons la signification précise d’un certain nombre d’en-têtes et de méthodes.

**Formulaires**

* Depuis la page n°1, cliquez sur le lien vers le formulaire.
* Remplissez le avec des données quelconques, puis cliquez sur Envoyer.
* Vous aboutissez sur une page récapitulant les données que vous avez saisies.

**Questions**

* + Quelle est la méthode utilisée par cette requête (Request method/Méthode de requête) ? Est-ce conforme avec ce que la page affiche.
* Une requête POST se différencie d’une requête GET par le fait que le navigateur **envoie** des données au serveur.
* Ces données sont visibles dans le panneau de détail :
  + dans l’onglet Parameters/Paramètres sous Firefox,
  + en bas de l’onglet Headers/En-têtes sous Chrome.

**Questions**

* + Affichez, dans l’onglet correspondant, le contenu de la requête. Est-ce conforme avec ce que la page affiche.
  + Revenez au formulaire et saisissez-y des informations différentes. Constatez dans la console que le contenu de la requête change en conséquence.

**Requêtes depuis un programme**

La navigateur n’est pas le seul client Web possible ; vous pouvez écrire votre propre programme qui interagira avec un serveur *via* HTTP.

* Créez un [environnement virtuel](https://perso.liris.cnrs.fr/pierre-antoine.champin/2019/progweb-python/annexes/a1_venv.html) python, et installez y le package [requests](https://perso.liris.cnrs.fr/pierre-antoine.champin/2019/progweb-python/annexes/a2_requests.html).
* Écrivez un programme qui, pour chacune des URLs ci-dessous, interroge l’URL (requête GET), et affiche :
* l’URL,
* le code de status (à trois chiffres) retourné par le serveur,
* le cas échéant, l’en-tête *content-type* de la réponse,
* le cas échéant, les vingts premiers octets du corps la réponse.

Liste des URLs à traiter :

URLS = [

'https://perso.liris.cnrs.fr/pierre-antoine.champin/2018/progweb-python/\_static/td1/page1.html',

'https://perso.liris.cnrs.fr/pierre-antoine.champin/2018/progweb-python/\_static/td1/no\_such\_page.html',

'https://perso.liris.cnrs.fr/pierre-antoine.champin/2018/progweb-python/\_static/td1/style.css',

'https://perso.liris.cnrs.fr/pierre-antoine.champin/2018/progweb-python/\_static/td1/sand.png',

'https://www.uniprot.org/',

'https://www.uniprot.org/uniparc/UPI000000001F',

'https://www.uniprot.org/uniparc/UPI000000001F.tab',

'https://www.uniprot.org/uniparc/UPI000000001F.truc',

'https://www.uniprot.org/images/logos/uniprot-rgb-optimized.svg',

]