

# TP : Trouver l'API cachée d'un site

## Objectif

Beaucoup de sites modernes sont des **SPA** : le navigateur charge une page quasi vide, puis récupère tout son contenu en appelant une API. Cette API n'est pas documentée publiquement, mais elle est sous vos yeux.

En groupe, vous allez **découvrir l'API cachée d'un site réel**, puis **rejouer une de ses requêtes hors du navigateur** avec `curl`.

On observe, on ne casse rien : uniquement des requêtes de lecture, à la main, sans boucle ni script. Les conditions d'utilisation de certains sites interdisent l'usage automatisé de leurs API internes.

## Consignes

### 1 - Choisir un site

Choisissez un site parmi la liste, ou proposez-en un autre au formateur.

- **Plutôt faciles** : Wikipédia (barre de recherche), Open Food Facts, Steam (suggestions de recherche), Deezer (recherche), adresse.data.gouv.fr
- **Plus coriaces** : Leboncoin, SNCF Connect

### 2 - Observer le trafic

- Ouvrez les DevTools (`F12`), onglet **Réseau** (*Network*).
- Filtrez sur **Fetch/XHR** pour ne garder que les appels d'API.
- Utilisez le site normalement (recherche, filtres, pagination...) et regardez quelles requêtes apparaissent.

### 3 - Analyser une requête

Pour une requête intéressante, relevez :

- l'URL et la méthode HTTP ;
- les paramètres (dans l'URL ou dans le body) ;
- les headers qui semblent importants ;
- le code de statut et le format de la réponse (onglets **Aperçu** et **Réponse**).

### 4 - Rejouer la requête avec curl

- Clic droit sur la requête → **Copier** → **Copier en tant que cURL**.
- Collez la commande dans un terminal et exécutez-la : vous devez obtenir la même réponse.
- Supprimez les headers **un par un** et relancez à chaque fois. Gardez la **commande la plus courte qui fonctionne encore**.
- Modifiez un paramètre (le terme recherché, la page...) et vérifiez que la réponse change.

Si la requête rejouée est refusée (403, captcha...), le site est sans doute protégé par un anti-bot. Ce n'est pas un échec : notez quels headers ou cookies semblent nécessaires et pourquoi.

## Résultat attendu

Une fiche par groupe, présentée à l'oral en quelques minutes :

- le site choisi et ce que fait la requête étudiée ;
- l'URL, la méthode et les paramètres ;
- un extrait de la réponse (format, structure) ;
- la **commande curl minimale** qui fonctionne, ou l'explication de ce qui bloque ;
- votre avis : cette API est-elle « propre » ? Les URL sont-elles lisibles ?