



Fundamentals of APIs

What is an API?

An Application Programming Interface (API) is:

...a set of functions and procedures allowing the creation of applications that access the features or data of an operating system, application, or other service.

(Oxford English Dictionary)

In essence: an API acts as an **intermediary** between software applications.

How does an API work?



Why collect data from APIs?

APIs can be an important source of publicly available information on social phenomena of interest.

APIs allow customised access to data resources.

Once collected, data can be reshaped into a familiar format (tabular) and linked to other sources of social science data.

What is the logic of using an API?

We need to **know** the following:

1. The location of the API (i.e., web address) through which the database can be accessed. For example, the UK Police API can be accessed via <https://data.police.uk/api>.
2. The terms of use associated with the API.
3. The location (*endpoint*) of the data of interest on the API.

What is the logic of using an API?

Then we need to **do** the following:

4. Register your use of the API (if required).
5. Request data from the endpoint of interest, supplying authentication if required. This process is known as making a call to the API.
6. Write this data to a file for future use.

What is the value of web-scraping?

The process of interacting with an API is a common and mature computational method, with lots of established packages (e.g., ``requests`` in Python), examples and help available.

APIs provide access to data that is intended to be shared.

The richness of some of the information and data stored on APIs is a point worth repeating.

APIs provide flexible, customisable access to data.

The data you need might only be available through an API.

What are the downsides of web-scraping?

APIs restrict the number of requests for data you can make.

The quality of an API's official documentation can vary wildly.

JSON!

An API is a product and you must comply with the Terms of Service/Use associated with it.

APIs can be updated on a frequent basis, resulting in changes to the rate limit, authentication requirements, endpoints etc.

Questions and Comments

Exercise

The Guardian API

<https://open-platform.theguardian.com/explore/>

Search for the following terms:

- Charity Commission
- Martin Thomas

What is JSON?

JSON (Javascript Object Notation) is a hierarchical data structure based on key-value pairs (known as items), which are separated by commas (Brooker, 2020; Tagliaferri, n.d.)

In Python a JSON data structure is known as a dictionary,

Questions and Comments