



Scottish Graduate School  
of Social Science  
*Sgoil Cheumnaichean Saidheans*  
*Sòisealta na h-Alba*

## Collecting Digital Data

What Are APIs and Why Use Them?

Dr Diarmuid McDonnell

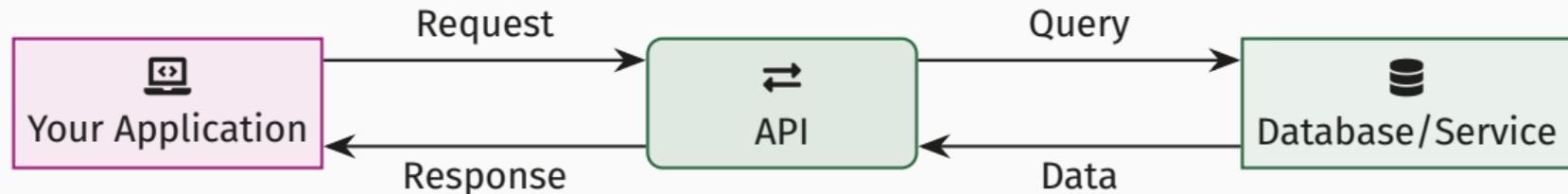
Braw Data Ltd · Gradel Institute of Charity, University of Oxford

24 February 2026

## **What Are APIs and Why Use Them?**

---

# What is an API?



**API = Application Programming Interface**

Think of it as a **translator**: your application doesn't need to know *how* the database is organised – it just asks the API in a standard way, and the API returns structured data.

# APIs vs Web Scraping

	<b>Web Scraping</b>	<b>APIs</b>
<b>Data format</b>	Unstructured HTML	Structured JSON/XML
<b>Reliability</b>	Breaks when site changes	Versioned, documented
<b>Legality</b>	Grey area	Usually explicit ToS
<b>Rate limits</b>	You must self-impose	Enforced by provider
<b>Authentication</b>	Usually none	Often API key

## Rule of thumb

If an API exists, use it. Web scraping is a last resort.

# Anatomy of an API Call

`https://data.police.uk/api/stops-force?force=city-of-london&date=2024-01`

Base URL

Endpoint

Query parameters  
key=value pairs joined by &



**Headers**  
Accept: application/json



**Authentication**  
API key, OAuth token, or none

The UK Police API is **open** — no authentication needed. Many other APIs require an API key sent as a header or query parameter.

# Reading API Documentation

## What to look for when you open the docs:

- ✓ **Endpoints** — what URLs are available and what data do they return?
- ✓ **Authentication method** — API key, OAuth, or open access?
- ✓ **Rate limits** — how many requests per minute/day?
- ✓ **Response format** — JSON, XML, CSV?
- ✓ **Example requests & responses** — try them in a browser first!

## UK Police API — [data.police.uk/docs/](https://data.police.uk/docs/)

- ✓ Open access (no key needed)
- ✓ Returns JSON
- ✓ Rate limit: 15 requests per second
- ✓ Endpoints: stop-and-search, crimes, neighbourhoods, forces

# JSON Deep-Dive

```
{  
  "type": "Person",  
  "age_range": "18-24",  
  "legislation": [  
    {  
      "name": "John Doe",  
      "age": 21  
    }  
  ],  
  "outcome": true  
}
```

## JSON types:

- **Object** {} – key-value pairs
- **Array** [] – ordered list
- **String, Number, Boolean, null**

JSON is the most common format for API responses. Each record is a set of **key-value** pairs. Values can be strings, numbers, booleans, arrays, or nested objects.

— Keys — Values

### Practical 2: UK Police API

- Make your first API request (in a browser, then in code)
- Parse JSON responses into tabular data
- Explore stop-and-search data for a chosen police force
- Handle pagination and query parameters

Open the **Practical 2** notebook in Google Colab  
(Python or R – your choice!)