

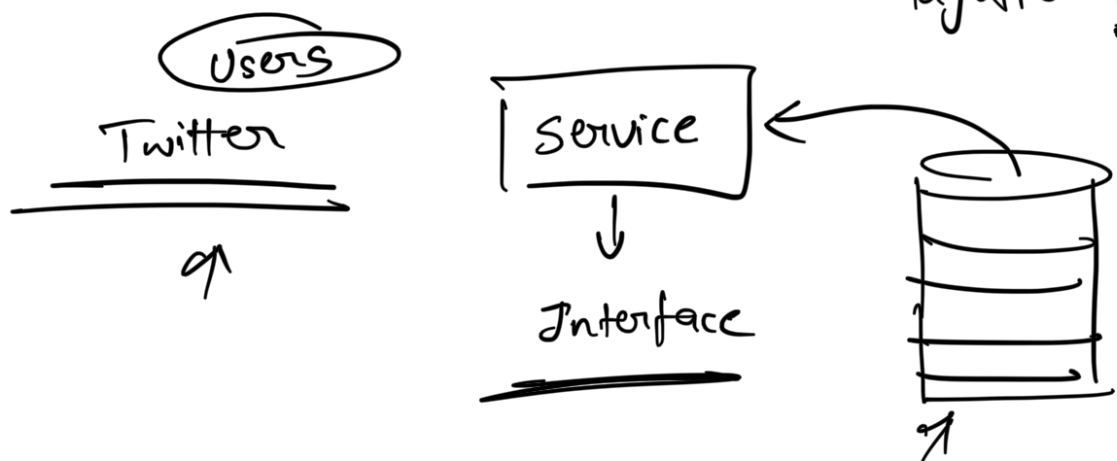
Web API & Scraping

- ① Scarcity of relevant data
 - ② Raw / useless data
- } ✓

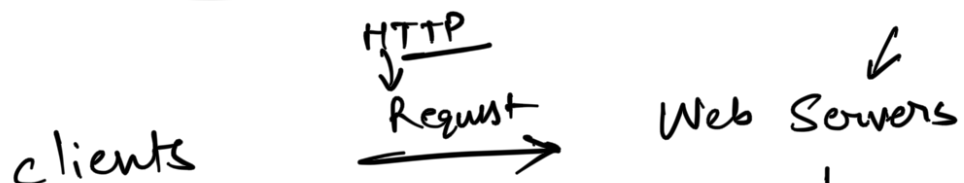
Common Data Sources

- ① Files / Public repositories ✓
 - ② DBMS (SQL)
 - ③ Web APIs (JSON) ✓
 - ④ Web Scraping
- Focus [

Problem: How to get all the tweets which are talking about 'Omicron' 'layoffs'?



WWW — World Wide Web

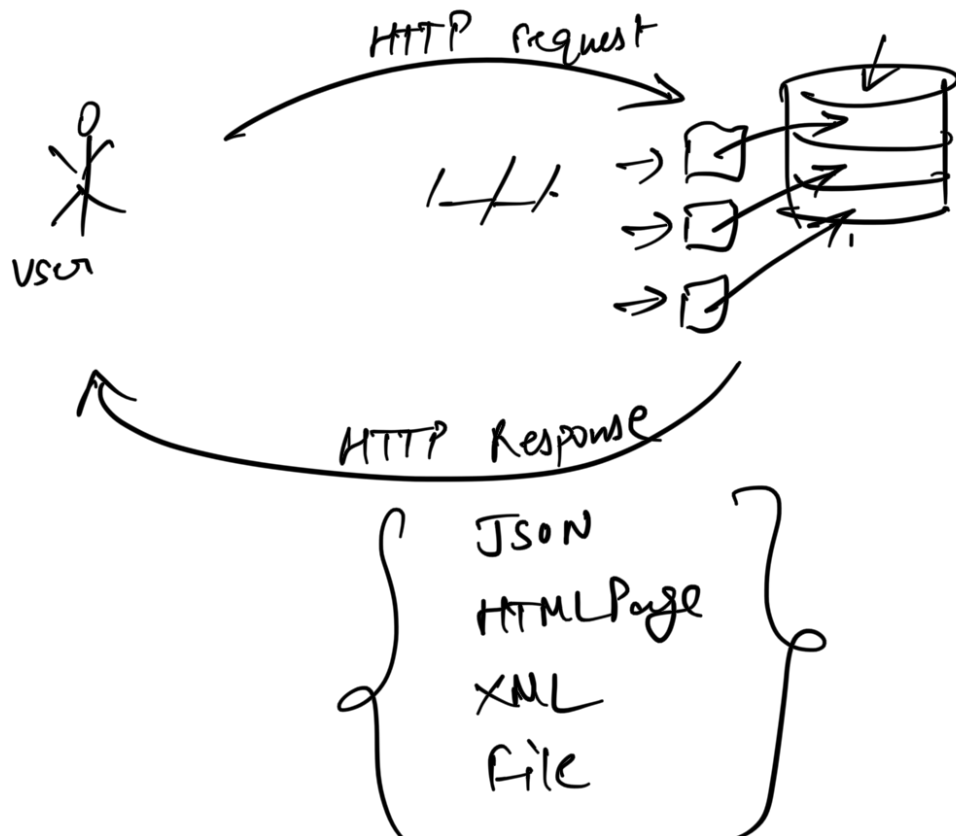
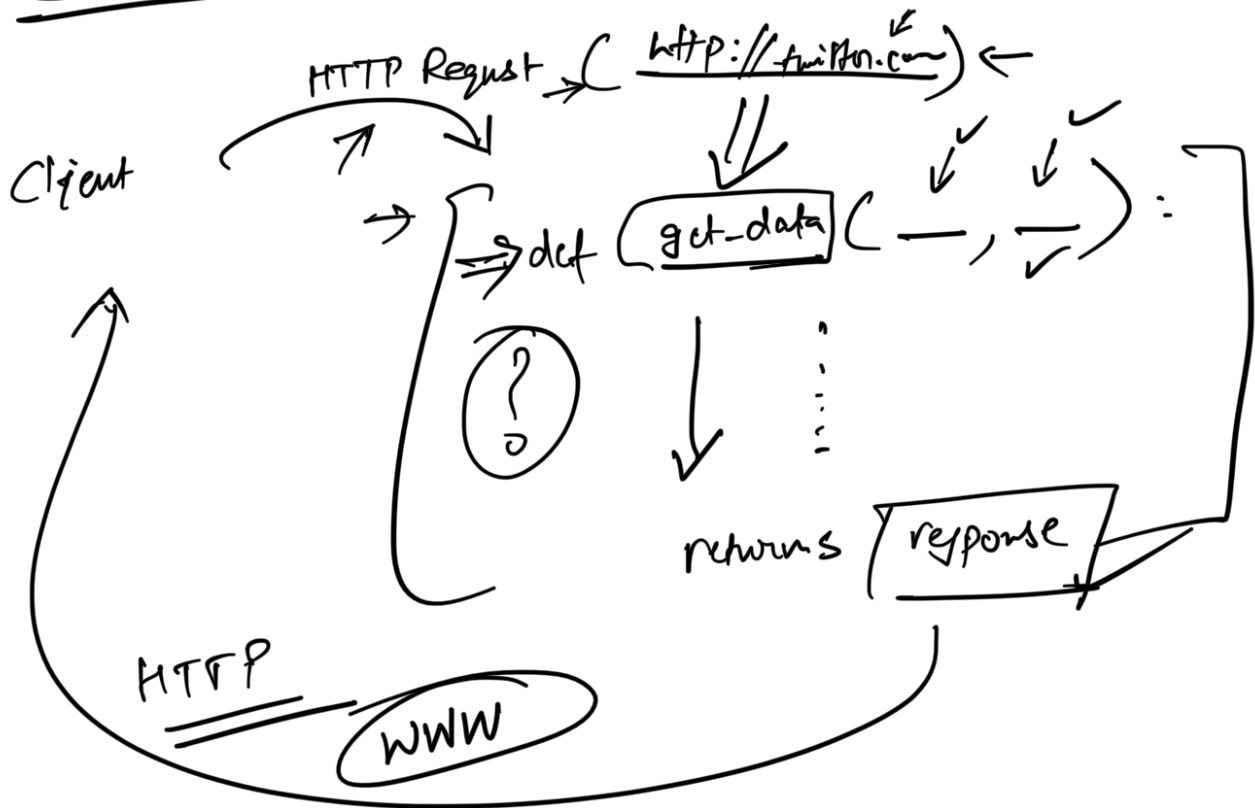


↳ Web Browsers
(Chrome, Safari, Edge)

← Response (HTTP)

on cloud

HTTP — HyperText Transfer Protocol



Types of Request (HTTP)

- ① GET → To request data from a specified resource.
- ② POST
- ③ UPDATE
- ④ DELETE
- ⑤ PUT

GET → Parameters are sent in URL

"/test/demo-form.php?name1=value1
& name2=value2"

Notes on GET

- ① Cached
- ② Browser History
- ③ Bookmark
- ④ Sensitive data shouldn't be sent on a GET request.

POST ✓ test/demo/ HTTP/1.1

→ Body of the request

Forms

body of

∴ Body ⇒ name1 = value1 & name2 = value2

HTTP Status Codes

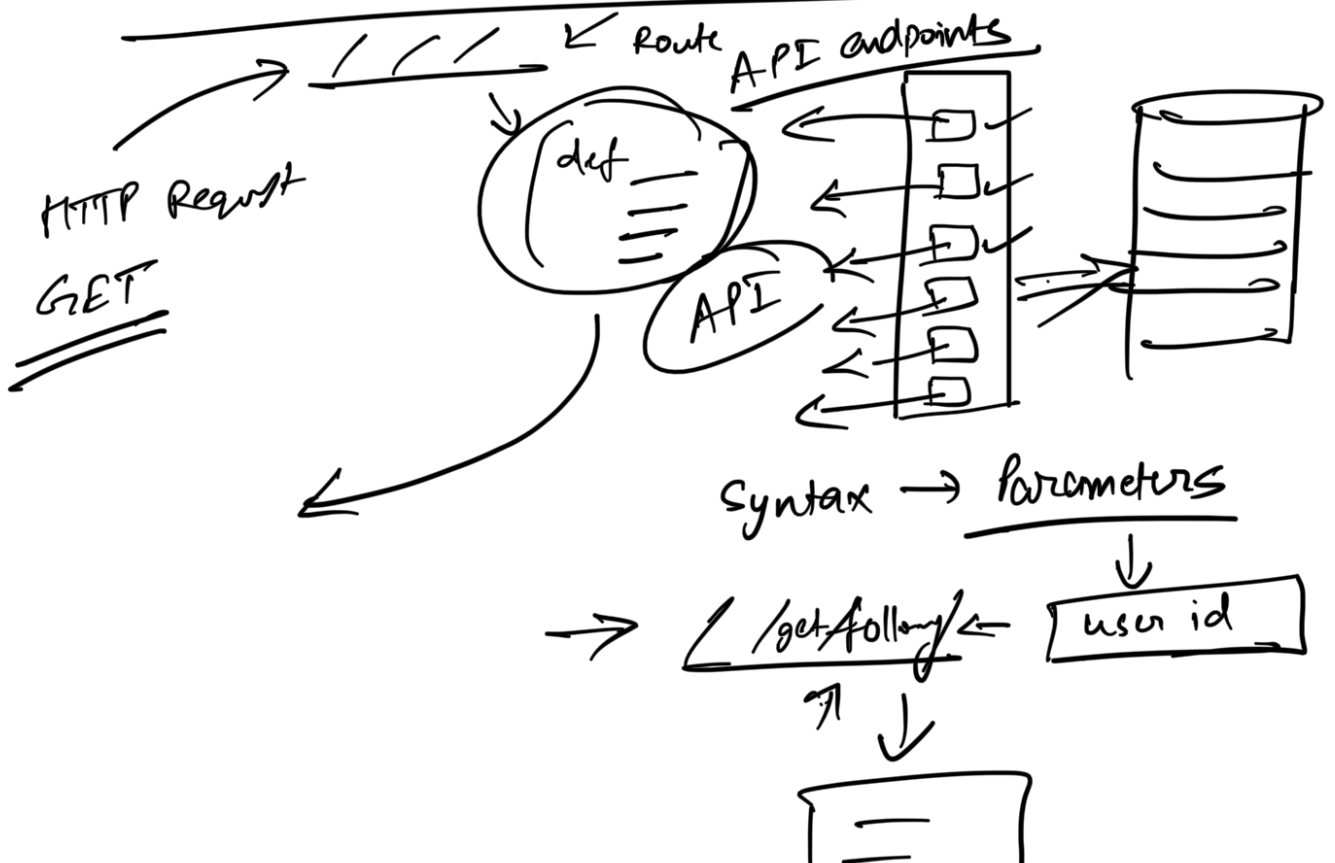
404 (Page not found)

200 - Success

✓ 500 - Internal Server Error

What is API?

Application Programming Interface



Types of APIs

Access Key
Access Token } Authentication ✓

Public APIs

↗
↓
Data