

What is an API (basic logic)?

- **API = Application Programming Interface**
- It's a way for one program to talk to another (usually over the web).
- Most modern APIs are **web APIs** that use **HTTP requests** (the same protocol your browser uses).

Think of it like a **restaurant menu**:

- The **menu** = API (list of available actions you can request).
- You (the client) place an **order** = API request.
- The **kitchen** (server) prepares it and gives you back a **dish** = API response (usually in JSON, CSV, or XML).

Anatomy of a web API request

Typical parts:

1. **Endpoint (URL)** → e.g. `https://api.gdc.cancer.gov/projects`
2. **Method** → e.g. GET (fetch data), POST (send data), etc.
3. **Parameters / filters** → e.g. `?project=TCGA-BRCA`
4. **Response** → structured data (JSON, CSV).

Example: NCI GDC API (raw HTTP)

Let's say you want to list all cancer projects in GDC.

Raw API request (in a browser or with `curl`):

`https://api.gdc.cancer.gov/projects`

Response (JSON snippet):

```
{
  "data": {
    "hits": [
      {"id": "TCGA-BRCA", "name": "Breast Invasive
Carcinoma"},
```

```

    {"id": "TCGA-LUAD", "name": "Lung Adenocarcinoma"},
    ...
  ]
}

```

That JSON is the **API response**. It's raw but machine-readable.

Same query in R with **TCGAbiolinks**

```
library(TCGAbiolinks)
```

```
# This calls the API for you
projects <- getGDCprojects()
```

```
head(projects[, c("project_id", "name")])
```

Output (cleaned up for R):

```

      project_id      name
1  TCGA-BRCA  Breast Invasive Carcinoma
2  TCGA-LUAD   Lung Adenocarcinoma
...

```

👉 Here, `getGDCprojects()` is just an **R wrapper around the API endpoint**. Instead of parsing raw JSON yourself, the package gives you a nice `data.frame`.

API logic in 3 steps

1. **You send a request** → usually a URL + method.
2. **Server processes it** → applies filters, retrieves data.
3. **You get a response** → often JSON/CSV, which you parse or an R package parses for you.

✅ So the logic is the same whether you do it raw (**curl**/browser) or via R (**TCGAbiolinks**, **GEOquery**) — the package just makes life easier.