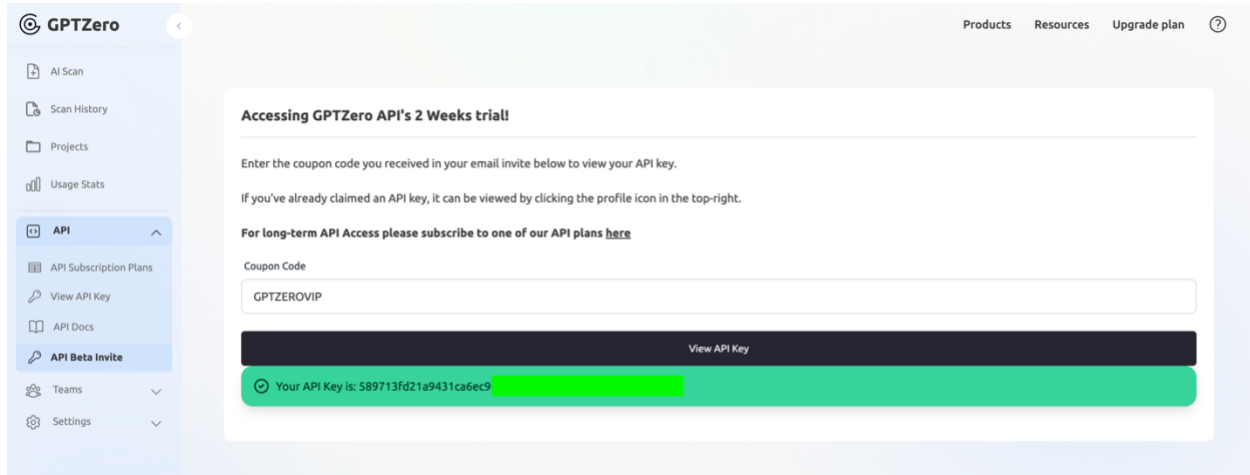


## Q1: Write a web scraper that scrapes something and runs the text through the GPTZero API.

Instructions:

1. Navigate to the GPTZero Dashboard ([app.gptzero.me](https://app.gptzero.me))
2. To get your free API key, navigate to the API -> API Beta Invite section and enter in the coupon code section: **GPTZEROVIP**
3. For the purposes of this assignment, we have rate limited your individual API key to 250 requests per hour active for two weeks.



4. Write a web-scraper in any programming language. If you need a place to start, Selenium is a great Python web scraper: <https://www.geeksforgeeks.org/selenium-python-tutorial/>
5. Access the GPTZero API Documentation here: <https://gptzero.stoplighlight.io/docs/gptzero-api/>. You can get starter code for any language via the request sample sections on the right. (see screenshot below)
6. Feel free to share the input and output of something you scraped, or your code. Either is fine.

The screenshot shows the GPTZero API documentation on the left and a code editor on the right. The documentation includes a sidebar with navigation links, a main content area with a description of the endpoint, and sections for Request, Headers, Body, and Responses. The code editor shows a Python script for sending a request to the GPTZero API, with a dropdown menu open over the 'import' statement, listing various programming languages and shells.

This endpoint takes in file(s) input and returns the model's result. By default, the maximum number of files that can be submitted simultaneously is 50, and the max file size for all files combined is 15 MB. Each file's document will be truncated to 50,000 characters.

### Request

#### Headers

**x-api-key** string  
Your GPTZero API key, which is necessary to use the API after testing. See how to get one at <https://gptzero.stoplight.io/>

#### Body

multipart/form-data

Takes in a file(s) as input with content type being "multipart/form-data"

**files** array(string) required  
A list of files. For an example of how to attach a list of files in python, see <https://stackoverflow.com/a/18184882>

**version** string  
Optionally, you can specify the desired version of the detector. We only maintain the latest version which is currently "2024-01-09"  
Default: 2024-01-09

#### Responses

200 400 404 429 500

A list of document perplexities. Returns DocumentPredictions object containing "documents" of each file

Request Sample: Python / Python 3

```
import os
import requests
import json

conn = requests.Connection("api.gptzero.me")
payload = {"files": [{"path": "111000001101001\r\nContent-Disposition: multipart/form-data; boundary=---\r\nContent-Type: application/json"}]}

headers = {"x-api-key": "your-api-key", "Content-Type": "multipart/form-data; boundary=---\r\nContent-Type: application/json"}

conn.post("/predict/files", payload, headers)

response = conn.json()

print(response)

# Example of a response: [{"document": "your document content", "perplexity": "0.9999999999999999"}]
```

*Screenshot of someone accessing the sample code for the GPTZero API. We are mostly interested in using the 'AI' 'human' or 'mixed' classification outputs of the API.*

## Q2: Using your web scraper, find an interesting result on the internet and share here.

Please feel free to use your discretion of what is considered 'interesting' but here are some examples of results that would knock this assignment out of the park.

- Uncovering published articles that might be AI written, like this [Futurism](#) investigation on Sports Illustrated, or analyzing and comparing AI probabilities of different content platforms.
- Scraping 100+ Substack articles and showing how some authors might have more AI likely posts after 2022.
- Scraping reviews of a sales website and finding AI-generated or mixed reviews
- Scraping Telegram or social media posts and uncovering networks that are repeatedly sending AI generated messages.

**Q3: If you have one, please share your writing portfolio or some samples of past writing (eg. articles, student newspaper, blog posts, personal pieces, etc.)**