

Instructions

This document provides instructions for using the web scraper which is designed to extract information from the India eTenders.

- **Set Up Your Python Environment:**

Ensure you have **VS Code** or any other IDE installed in your setup.

For creating an environment using bash:-

Refer to readme.md

For activating environment:-

Refer to readme.md

- **Getting Session Token:-**

1. Go to the etenders.gov.in website
2. Open developer tools
3. Click on Network tab in developer portal
4. Click on any link on the site
5. When the site loads you'll find "app?component=" in name on developer
6. Click on it, under headers tab find **Cookie:**
7. Copy example: `JSESSIONID=B9BE18334CEF09A405C8F5558FC57F60.geps1;`
8. Copy ignoring "JSESSIONID= and ;"

- **Install Required Libraries:**

Refer to readme.md

- **Run the Script:**

Execute this scraper script. This can be done in a few ways: Run it from the command line by navigating to the script's directory and running "python app.py"

Run it directly in your preferred Python IDE (e.g., VS Code or Jupyter Lab).

- **Wait for the Script to Complete:**

The script will scrape data from INDPPP URLs. It may take some time, depending on the number of URLs you're scraping and other factors. Be patient, and let the script run.

- **Data Output:**

The script will save the scraped data to a CSV file. By default, it's named "PPP_tenders.csv" in the script's directory. You can change the file name and path as needed.

- **Review the Scraped Data:**

Open the generated CSV file to review the data. If it's not accurate, check for errors in the log file.

Note:

Please make sure to comply with the website's terms of service and any legal or ethical guidelines when using this scraper.

The provided session token is for demonstration purposes.

You should obtain a valid session token if you intend to use the scraper for a real application.

Customization of the scraper may be required based on the specific website or data you want to extract.

Feel free to ask if you have any specific questions or need further assistance with any part of the code.

Email : nilaykr27@gmail.com