**README**

**Overview**

This project contains two Python scripts:

1. **excel.py** - A utility for creating and managing Excel spreadsheets.
2. **ida\_selenium.py** - A Selenium-based web scraper designed to analyze web pages, extract network requests, and store relevant data in an Excel file.

**Prerequisites**

Before running these scripts, ensure you have the following dependencies installed:

* Python 3.x
* openpyxl (for working with Excel files)
* selenium (for web automation)
* pychrome (for Chrome DevTools integration)
* htmlreport (for generating HTML reports)

Install them using pip:

pip install openpyxl selenium pychrome htmlreport

For Selenium, ensure you have the correct **Chrome WebDriver** installed and accessible in your system's PATH.

**Script Details**

**1. excel.py**

This script provides functionality to create and manipulate Excel spreadsheets. It checks if an Excel file (Data.xlsx by default) exists and either loads or creates it. The script defines methods for adding website analysis data such as:

* Website name
* Date and time of analysis
* Website link
* Third-party trackers
* User consent details
* Privacy policy status

**Usage**

This script is primarily used as a helper module within ida\_selenium.py. However, it can also be tested independently:

python excel.py

**2. ida\_selenium.py**

This script uses Selenium to automate web browsing and analyze network requests. It logs third-party requests that contain a given search term and stores findings in an Excel file (testExcel.xlsx).

**Features**

* Opens and navigates web pages using Selenium.
* Extracts network requests using Chrome DevTools (pychrome).
* Filters requests containing user-specified search terms.
* Logs third-party requests into an Excel spreadsheet.
* Generates HTML reports for analyzed websites.

**Usage**

Run the script by providing a text file containing a list of URLs to analyze:

python ida\_selenium.py urls.txt

The script will prompt you for a search term and begin analyzing the specified URLs.

**Example urls.txt:**

https://example.com

https://anotherwebsite.com

**Running the script:**

python ida\_selenium.py urls.txt

This will:

1. Open each URL in Chrome.
2. Perform a search for the provided term.
3. Log any third-party tracking data.
4. Store results in testExcel.xlsx.
5. Generate an HTML report.

**Notes**

* The script may require **Chrome remote debugging** to be enabled.
* If the Chrome WebDriver version is incompatible with your browser, update it accordingly.
* The script may require administrator or root permissions to modify browser settings.