

## Ethical Hacking & Website Exploitation

Deadline: March 22nd 2023

**Disclaimer:** It is strictly prohibited to test the following code on public websites. You can download [Metasploit](#), which provides access to websites designed for hacking.

You have been tasked with creating a Python script that can be run on a website to discover subdomains (e.g. admin-test.domain.com), directories (e.g. domain.com/for\_admins\_only/), and files (e.g. domain.com/user\_passwords.txt). This script will reveal hidden links that may lead to pages with vulnerabilities. Additionally, two text files are provided, which you will use to read potential subdomains and directories rather than generating them from scratch.

Your script should be executed from the terminal with the command "python test\_script.py" and should take an argument that represents the target website's URL. The script must store the input file values and write the output to the following files:

- subdomains\_output.bat
- directories\_output.bat
- files\_output.bat

You will need to inspect the HTML code to identify all possible unique links (usually stored in the "href" attribute inside the <a> tag). You are not allowed to use BeautifulSoup or other libraries to scrape a website. Instead, you must use regular expressions to extract information from the HTML file.

Furthermore, when reading text files, you may need to clean the string by removing extra spaces or trailing special characters. Please use regex to accomplish this, and do not use built-in functions.

### **Bonus:**

You can attempt to guess a website's login information by detecting the POST URL and using brute force to test several passwords.

### **Deliverables:**

- Publish your code on GitHub. Only atomic commits are allowed.
- Write a report detailing the project and the steps taken to complete it. The report should also include any challenges encountered and how they were overcome.