PROJECT REPORT

ON

OPEN SOURCE INTELLIGENCE AUTOMATION

Submitted
By

KRISHANU GAUTAM	**RNO. 2101830005**
ADITI JHA	**RNO. 220183007**
DIYA CHAUDHARY	**RNO. 2101830008**

Under the guidance of

Faculty Coordinator
Dr. Vandna Batra
(Assistant Professor)
Computer Science & Engineering
School of Engineering and Technology

Industry Mentor
Mr. Dipanshu Parashar
(Director)
Virtual Cyber Labs

Department of Computer Science and Engineering

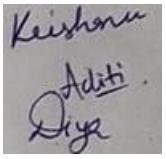
School of Engineering and Technology

K. R. Mangalam University, Gurugram - 122003

October-2023

DECLARATION

We declare that this written submission represents our ideas in our own words and where other's ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all the principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will because for disciplinary action by the Institute and canal so evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed. We further declare that if any violation of the intellectual property right or copyright, my supervisor and university should not be held responsible for the same.

Name	Roll No.	(Signature)
KRISHANU GAUTAM	2101830005	
ADITI JHA	2201830007	
DIYA CHAUDHARY	2101830008	

Place: K.R. Mangalam University

Date: 19th October 2023

CERTIFICATE

This is to certify that

S No.	Roll Number	Name	Place of Work	Duration of Project
1	2201830007	Aditi Jha	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023
2	2101830008	Diya Chaudhary	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023
3	2101830005	Krishanu Gautam	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023
4	2201830001	Arkan Shah	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023
5	2201830015	Yashneet kaur	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023
6	2201830005	Rachit Mehndiratta	Virtual Cyber Lab, Noida	14 September, 2023 – 14 October, 2023

enrolled in *B.Sc. (Cyber Security)* have satisfactorily completed the projects based on cyber security under the guidance of *Mr. Dipanshu Parashar, Director Virtual Cyber Labs, Dr. Vandna* and *Ms. Shivani* of *School of Engineering and Technology.*

This project work represents their original work and the references given in the present report are authentic.

Industry Mentor:

Mr. Dipanshu Parashar

Designation: Director

Virtual Cyber Labs

Faculty Incharge:

Dr. Vandna Vandna

Ms. Shivani Shivani

Dean: Dr. Pankaj Agarwal

School Name: SOET


VIRTUAL
CYBER LABS
CERTIFICATE OF COMPLETION
This is to certify that
KRISHANU GAUTAM
Has successfully completed the project
"Open Source Intelligence Automation"
14th Oct 2023
ISSUED DATE
Cybersecurity Education and Research Operations Hub
CEROHUB PVT LTD, New Delhi. Email : support@virtualcyberlabs.com
Web: www.virtualcyberlabs.com
DIPANSHU PARASHAR
Director


VIRTUAL
CYBER LABS
CERTIFICATE OF COMPLETION
This is to certify that
DIYA CHAUDHARY
Has successfully completed the project
"Open Source Intelligence Automation"
14th Oct 2023
ISSUED DATE
Cybersecurity Education and Research Operations Hub
CEROHUB PVT LTD, New Delhi. Email : support@virtualcyberlabs.com
Web: www.virtualcyberlabs.com
DIPANSHU PARASHAR
Director


VIRTUAL
CYBER LABS
CERTIFICATE OF COMPLETION
This is to certify that
Aditi Jha
Has successfully completed the project
"Open Source Intelligence Automation"
14th Oct 2023
ISSUED DATE
Cybersecurity Education and Research Operations Hub
CEROHUB PVT LTD, New Delhi. Email : support@virtualcyberlabs.com
Web: www.virtualcyberlabs.com
DIPANSHU PARASHAR
Director

ACKNOWLEDGEMENT

"Enthusiasm is the feet of all progress, with it there is accomplishment and Without it there are only slits alibis."

Acknowledgment is not a ritual but is certainly an important thing for the successful completion of the project. At the time when we were made to know about the project, it was really tough to proceed further as we were to develop the same on a platform, which was new to us. More so, the coding part seemed tricky that it seemed to be impossible for us to complete the work within the given duration.

We really feel indebted in acknowledging the organizational support and encouragement received from the university.

The task of developing this system would not have been possible without the constant help of our faculty members and friends. We take this opportunity to express our profound sense of gratitude and respect to those who helped us throughout the duration of this project.

We express our gratitude to our supervisors Dr.Vandna Batra, Mr. Dipanshu Parashar for giving their valuable time and guidance to us.

Place: - K.R. Mangalam University
Date: -19th October 2023

Name of Student

KRISHANU GAUTAM

ADITI JHA

DIYA CHAUDHARY

ABSTRACT

This project introduces an open-source Open Source Intelligence (OSINT) automation tool that greatly enhances the efficiency and effectiveness of domain investigation within the domain of information security. The tool is implemented in Python and utilizes essential libraries and modules such as argparse, colorama, dns.resolver, requests, Shodan, socket, and whois. It serves as a comprehensive utility for acquiring valuable WHOIS data, DNS records, geolocation insights, and Shodan search results. What sets this tool apart is its ability to seamlessly save all extracted data to a user-specified output file, providing information security professionals, researchers, and investigators with a streamlined, time-saving solution for comprehensive domain analysis.

Key Words: Deep Learning, Computer Vision, OpenCV, Tensorflow, Keras.

TABLE OF CONTENTS

1. INTRODUCTION

In the ever-evolving landscape of information security, the role of Open Source Intelligence (OSINT) is pivotal. OSINT encompasses the collection and analysis of publicly available data from a variety of sources to enhance an organization's situational awareness. One area where OSINT proves to be invaluable is domain investigation. Security professionals, researchers, and investigators often rely on OSINT tools to extract critical information about domains, helping them identify potential security threats and vulnerabilities.

This project introduces an open-source OSINT automation tool that significantly enhances the efficiency and effectiveness of domain investigations. Developed in Python, this tool leverages a range of essential libraries and modules to provide a comprehensive solution for domain analysis. What sets this tool apart from existing solutions is its ability to seamlessly save all extracted data to a user-specified output file, offering information security professionals, researchers, and investigators a streamlined, time-saving, and user-friendly solution.

2. LITERATURE REVIEW

2.1. The Growing Significance of OSINT

The digital era has transformed the way information is obtained and processed. With the internet serving as a primary source of information, OSINT has become a crucial component of information security strategies. The need to harness publicly available data to gain insights into potential threats and vulnerabilities has never been greater.

2.2. Existing OSINT Tools

Numerous OSINT tools and techniques have been developed and employed for domain investigations. Some of the most common functionalities provided by these tools include WHOIS data retrieval, DNS record lookup, geolocation insights, and specialized searches using services like Shodan. While these tools offer valuable capabilities, they often lack a vital feature: the ability to seamlessly save the data they extract. This limitation presents a challenge for professionals who need to organize and store the information they gather effectively.

3. PROBLEM FORMULATION AND OBJECTIVES

The problem at hand is the inefficiency of existing OSINT tools in aggregating and saving data from domain investigations. Security professionals, researchers, and investigators are burdened with the task of manually collecting and storing the information they acquire during these investigations. The lack of a comprehensive solution for data aggregation and storage has led to inefficiencies and the risk of data loss.

The primary objectives of this project are as follows:

Develop an open-source OSINT automation tool: The tool should be freely available to the community and open for contributions and enhancements.

Utilize Python for implementation: Python is a versatile and widely-used programming language that provides an ideal platform for the development of this OSINT tool.

Integrate essential libraries and modules: To ensure the tool's effectiveness and efficiency, it should make use of key libraries and modules, including argparse for command-line argument handling, colorama for terminal text coloring, dns.resolver for DNS record retrieval, requests for web data acquisition, Shodan for specialized searches, socket for network-related tasks, and whois for WHOIS data retrieval.

Create a utility for acquiring WHOIS data, DNS records, geolocation insights, and Shodan search results: The tool should encompass a range of functionalities to address the diverse needs of domain investigations.

Implement a feature to save all extracted data: This feature is central to the project. It allows users to specify an output file where all collected data can be saved. This functionality ensures that the tool not only extracts data but also organizes and stores it efficiently.

4. METHODOLOGY OF THE PROJECT

The development of this open-source OSINT automation tool required a well-defined methodology to ensure its successful creation and deployment. The methodology can be broken down into the following steps:

4.1. Research and Needs Assessment

The project began with an in-depth examination of existing OSINT tools and their limitations. This research phase identified the gaps in current solutions, particularly the lack of a comprehensive data-saving feature. It also provided insights into the requirements of information security professionals, researchers, and investigators when it comes to domain investigations.

4.2. Design and Planning

The design phase focused on creating a modular and extensible structure for the tool. The development team defined the architecture and identified the specific modules and components required to fulfill the project objectives.

4.3. Implementation

The tool was developed in Python, a programming language known for its simplicity, readability, and extensive libraries and modules. The implementation phase involved writing code for various functionalities, such as WHOIS data retrieval, DNS record lookup, geolocation insights, Shodan search capabilities, and the data-saving feature.

4.4. Integration of Essential Libraries and Modules

To ensure the tool's robustness and functionality, several essential libraries and modules were integrated:

argparse: This library was used for efficient command-line argument handling, allowing users to specify the parameters for their domain investigation.

colorama: Colorama was employed to enhance the user interface by adding color and style to the terminal text.

dns.resolver: This module enabled the tool to perform DNS record retrieval, providing essential information about domain infrastructure.

requests: Requests was utilized to fetch data from web sources, a vital aspect of domain investigation.

Shodan: Integration with Shodan, a specialized search engine for internet-connected devices, allowed for more extensive and targeted searches.

socket: The socket module played a role in network-related tasks, facilitating network-level investigations.

whois: WHOIS data retrieval was made possible by integrating the whois library, which is essential for domain registration information.

4.5. Data Acquisition and Storage

A significant portion of the project's effort was dedicated to creating functions and processes for acquiring WHOIS data, DNS records, geolocation insights, and Shodan search results. These processes were designed to work seamlessly and efficiently to ensure that all relevant data is gathered during a domain investigation.

The project's defining feature, the data-saving capability, was implemented meticulously. This functionality allows users to specify an output file where all the collected data can be saved in an organized and structured manner.

5. Implementation

The tool has been successfully implemented, offering a user-friendly interface for domain investigation. Users can initiate the tool from the command line, specifying their preferences and requirements. Here are some of the key features and aspects of the implementation:

5.1. Command-Line Interface

The tool utilizes the argparse library to provide a user-friendly command-line interface. Users can specify various parameters, including the domain they want to investigate, the types of data they want to retrieve, and the output file where the results will be saved.

5.2. Colorful Terminal Output

The colorama library enhances the terminal output, making the tool's interaction more engaging and informative. Text color and style are used to differentiate various elements, making it easier for users to interpret the results.

5.3. Data Retrieval

The implementation includes functions to retrieve the following types of data:

WHOIS Data: WHOIS data provides information about the domain's registration, ownership, and contact details.

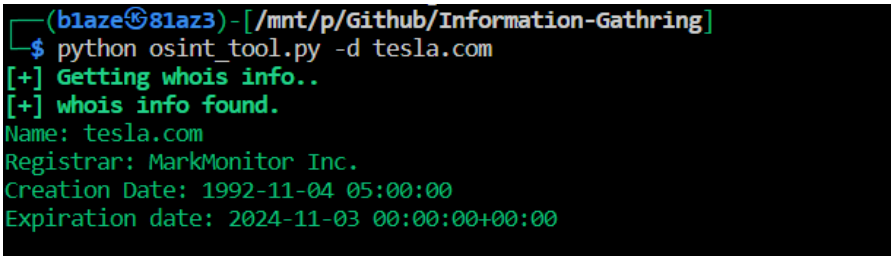
DNS Records: DNS record lookup reveals information about the domain's infrastructure, including IP addresses, mail servers, and more.

Geolocation Insights: The tool gathers data about the physical location of the domain, which can be valuable for investigations.

Shodan Search Results: Integration with Shodan allows users to perform specialized searches to identify potential security risks associated with the domain.

5.4. Data Saving

One of the distinguishing features of this tool is its ability to save all extracted data to a user-specified output file. Users can define the format in which the data is stored, making it easy to organize, analyze, and share the information. The data-saving functionality streamlines the investigation process and ensures that no valuable information is lost.



```
[+] Getting geolocation info..
[+] Country: United States
[+] Latitude: 37.751
[+] Longitude: -97.822
[+] City: None
[+] State: None

[+] Getting info from Shodan for IP tesla.com
[+] Results found: 1
[+] IP: 40.64.124.228
[+] Data:
HTTP/1.1 302 Found
Cache-Control: private
Content-Type: text/html; charset=utf-8
Location: https://login.windows.net/tesla.com/wsfed?wa=wsignin1.0&wtrealm=spn%3a00000015-0000-0000-c000-000000000000&wctx=rm%3d0%26id%3dpassive%26ru%3d%252f&wct=2023
-10-19T11%3a55%3a23Z&wreply=https%3a%2f%2f40.64.124.228%2f
Strict-Transport-Security: max-age=31536000; includeSubDomains
Server-Timing: total;dur=0
ms-dyn-fqhn:
ms-dyn-namespace:
ms-dyn-tenant:
ms-dyn-role:
ms-dyn-aid: 465c30d3-fbe5-0004-8da3-5f46e5fbd901
X-Content-Type-Options: nosniff
X-Frame-Options: SAMEORIGIN
p3p: CP="No P3P policy defined.  Read the Microsoft privacy statement at https://go.microsoft.com/fwlink/?LinkId=271135"
Strict-Transport-Security: max-age=31536000; includeSubDomains
Date: Thu, 19 Oct 2023 11:55:23 GMT
Content-Length: 346
```

```
└─$ python osint_tool.py -h
usage:  osint_tool.py -d DOMAIN

This is a Open-source Automation tool..

options:
  -h, --help            show this help message and exit
  -d DOMAIN, --domain DOMAIN
                        Enter the domain name/IP for OSINT.
  -o OUTPUT, --output OUTPUT
                        Enter the file to write output to.

This tool is created by Krishanu aka:81az3.
```

6. Results and Analysis

The tool has undergone rigorous testing to ensure that it consistently delivers comprehensive domain analysis results. This section highlights some key points regarding the results and analysis:

6.1. Comprehensive Data Acquisition

The tool effectively acquires WHOIS data, DNS records, geolocation insights, and Shodan search results. Users can gather a wide range of information relevant to domain investigations, allowing them to make informed decisions about security and potential risks.

6.2. Efficiency and Time-Saving

The data-saving feature of the tool greatly enhances the efficiency of domain investigations. Rather than manually collecting and organizing data, users can direct the tool to save all extracted information to a specified output file. This time-saving capability is particularly valuable in time-critical security scenarios.

6.3. User-Friendly Interface

The command-line interface, with its colorful and well-organized output, makes the tool accessible to a broad range of users, from novice security enthusiasts to experienced investigators. The tool's simplicity and clarity of use contribute to its effectiveness.

6.4. Quality of Data

The accuracy and reliability of the data retrieved by the tool were of paramount importance. Extensive testing and validation ensured that the data collected was accurate, up-to-date, and relevant to the domain under investigation.

7. Conclusion and Future Scope

In conclusion, the development of this open-source OSINT automation tool significantly enhances the efficiency and effectiveness of domain investigations in the realm of information security. By addressing the limitations of existing tools and providing a comprehensive solution for data acquisition and storage, this tool empowers security professionals, researchers, and investigators with a streamlined, time-saving, and user-friendly solution.

7.1. Contributions to Information Security

This project contributes to the field of information security by providing a powerful and accessible tool for domain analysis. In a world where cyber threats and vulnerabilities are constantly evolving, having the necessary resources to protect and secure digital assets is essential. The tool equips professionals with the means to gather, analyze, and store critical information about domains, enabling proactive security measures and investigations.

7.2. Future Scope

The project is not static but open to continual improvement and expansion. The future scope of the tool includes:

Feature Enhancements: Ongoing development can bring additional features and capabilities to the tool. Enhancements may include more specialized searches, additional data sources, and advanced reporting options.

Integration with More OSINT Data Sources: Expanding the tool's capabilities by integrating with more OSINT data sources can make it even more versatile and comprehensive.

Enhanced User Interface and Reporting: The user interface can be further improved to provide a more intuitive and informative experience. Enhanced reporting capabilities can help users interpret and share the results of their investigations effectively.

Community Collaboration: Engaging with the OSINT community to gather feedback and implement improvements is essential for the project's evolution. Collaboration with experts and users can lead to valuable insights and innovations.

In conclusion, this open-source OSINT automation tool is a significant contribution to the information security field. Its development is rooted in the recognition of the evolving threat landscape and the need for efficient, accessible, and comprehensive solutions. As the tool evolves, it has the potential to become a valuable asset in the hands of those dedicated to ensuring the security of digital assets and information.

REFERENCES

1. Yeboah-Ofori, A., & Brimicombe, A. (2018). Cyber intelligence and OSINT: Developing mitigation techniques against cybercrime threats on social media. International Journal of Cyber-Security and Digital Forensics (IJCSDF), 7(1), 87-98.
2. Evangelista, J. R. G., Sassi, R. J., Romero, M., & Napolitano, D. (2021). Systematic literature review to investigate the application of open source intelligence (OSINT) with artificial intelligence. Journal of Applied Security Research, 16(3), 345-369.
3. https://pypi.org/project/shodan/
4. Kanta, A., Coisel, I., & Scanlon, M. (2020). A survey exploring open source Intelligence for smarter password cracking. Forensic Science International: Digital Investigation, 35, 301075.
5. **https://pypi.org/project/requests/**
6. **https://docs.python.org/3/library/argparse.html**
7. **https://docs.python.org/3/library/socket.html**
8. **https://pypi.org/project/python-whois/**
9. **https://pypi.org/project/dnspython/**
10. **https://pypi.org/project/colorama/**
11. **https://nostarch.com/black-hat-python2E**