International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT232551

Osint Automation Application

Vinay Kumar Kureel[1], Manav Arya[2], Aditya Kini[3], Suraj Maurya[4], Rajesh Gaikwad[5]
[1-4]Department of Computer Engineering, Shree LR Tiwari College of Engineering, Mumbai, Maharashtra, India
[5]Assistant Professor Department of Computer Engineering, Shree LR Tiwari College of Engineering, Mumbai, Maharashtra, India

ARTICLEINFO

Article History:
Accepted: 07 April 2023
Published: 19 April 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
377-381

ABSTRACT

The amount of information produced by today's digital age is immense, and a major amount of it is available publicly, meaning that anybody can access it at any time, from anywhere on the Internet. In this way, Open-Source Intelligence (OSINT) is a branch of intelligence that gathers, analyses, and combines information from multiple sources across the entire cyberspace to provide information. In fact, because of recent technological developments, OSINT is receiving greater attention at an incredible rate, providing all the relevant information that eventually helps organizations or even individuals fight against cyber threats and cybercrime. This paper's objective is to outline the most recent advancements in OSINT technologies, which are now addressing data breaches and dispersed organizational responsibilities. We are building a web application that collects and processes publicly available information in an effective and efficient manner for providing insight into possible data leaks, vulnerability or any other sensitive Information using OSINT Framework.

Keywords: OSINT, Cybersecurity, Data Harvesting, Cyber Intelligence, Cyber Defense, Computational Intelligence, Data Privacy.

I. INTRODUCTION

The definition of open-source intelligence (OSINT) is the gathering, processing, and analysing of publicly available information from publicly accessible data sources, such as the news media, social networks, forums, and blogs, as well as publicly available government data, publications, or commercial data. OSINT continuously deepens the understanding of the target using a small amount of input data and the use of sophisticated collection and analysis techniques. In this approach, the information uncovered helps the investigative process move closer to the goal. OSINT is increasingly commonly used by governments and intelligence organisations to conduct investigations and fight cybercrime. Yet, in addition to those relating to state affairs, it also serves a number of other objectives.

A crucial tool for the intelligence community is quickly emerging extracting unique and valuable intelligence from public records to build thorough profiles of specific targets.

The OSINT provides opportunities for both defenders and attackers; you can identify a company's weakness and rectify it while also having the potential for the hole to be exploited.

II. LITERATURE SURVEY

Past studies demonstrate a variety of methods for obtaining publicly accessible information from the Internet that collect, analyse, and combine data from various sources throughout the entire cyberspace to offer information.

[1] The current state of OSINT is discussed in detail, and the paradigm is thoroughly examined, with a focus on the techniques and instruments developing the cybersecurity sector. On the flip side, we talk about its disadvantages. The present situation of OSINT was covered in this study. The ability to guarantee the necessary outcome for a particular purpose in an automated and self-driven way is OSINT's ultimate goal. This paper discusses the advantages and limitations of OSINT in the online world.

[2] The purpose of this article was to illustrate the value of OSINT in comparison to other intelligence techniques and to clarify various OSINT domain concepts. The report continued by discussing the advantages and disadvantages of using OSINT in cyberspace. According to researchers, it can be difficult to analyse the huge volume of data that OSINT provides. Before making any judgements, the sources of the OSINT data must also be confirmed.

[3] This study illustrates the various immediate benefits of using OSINT in exploratory large-scale data analysis. This project's goal was to show how well an automated system works for acquiring and analysing cybersecurity threat intelligence in order to undertake near-real-time information analysis.

[4] This study illustrates the various immediate benefits of using OSINT in exploratory large-scale data analysis. This project's goal was to show how well an automated system works for acquiring and analysing cybersecurity threat intelligence in order to undertake near-real-time information analysis.

[5] This study uses a number of approaches to collect a sizable amount of data and discusses key insights that may be used in a cyber operation. This is explored throughout the study as the researcher tries to show how useful OSINT is. Unfortunately, there are many readily available sources that include information that is freely available, which is not ideal and reduces the overall security of the system. The researcher then offered their own methods for discovering such data and fixing such flaws in advance of a cyberattack.

III.METHODOLOGY

The first step in every penetration testing strategy is reconnaissance (in other words, with OSINT). Open-Source Intelligence (OSINT) is divided into three categories: semi-passive, active, and passive.

Obtaining passive information Generally speaking, passive information collection is only beneficial if it is absolutely required for the target to remain unaware of the operations. This form of profiling is technically difficult to carry out because we never send traffic to the target organisation from one of our hosts or from "anonymous" hosts or services across the Internet. This implies that we are only able to access and collect data that has been saved or archived. We are only able to use data that has been collected from a third party, thus this data could be out-of-date or incorrect.

Semi-passive Information Gathering's goal is to profile the target utilising methods that appear to be consistent with regular Internet behaviour and usage. We only query the public name servers for data, we don't perform exhaustive reverse lookups or brute-force DNS requests, we don't search for "unpublished" servers or directories, etc. We are not actively searching for secret information; we are merely examining the metadata in published papers and files. We don't use crawlers or portscanners at the network level. Here, it's crucial to keep our actions hidden from view. The target might be able to observe the reconnaissance efforts in retrospect, but they shouldn't be able to attribute the action to a particular individual.

Active Information Gathering: The target should be able to recognise active information gathering if there is any suspicious or damaging behaviour. At this stage, we are actively mapping the network architecture (think full port scans with nmap -p1-65535), actively identifying the open services and/or testing them for vulnerabilities, and aggressively searching for unpublished servers, files, and folders. Most of this activity falls under the "reconnaissance" or "scanning" categories typical of traditional pentests.

The Open-Source Intelligence (OSINT) outlines a 5-step process:

1. Source Identification

Determine where to find the information for the specific intelligence requirement.

2. Harvesting

Gather relevant information from the identified source.

3. Data Processing

Process the identified source's data and extract meaningful insights.

4. Analysis

Combine the processed data from multiple sources.

5. Reporting

Create a final report on findings

Collection From Open Source : The investigator will investigate their sources for information such as emails, phone numbers, usernames, given names, addresses, etc. to compile a file on the case at hand, adhering to the maxim "start with what you know or what you have."

Filtering: Many of the discoveries will prove to be inconsistent with the research or irrelevant, and they must be put away in order to make the proper judgements. Dealing with incorrect information will inevitably result in the wrong decision being made regarding a case.

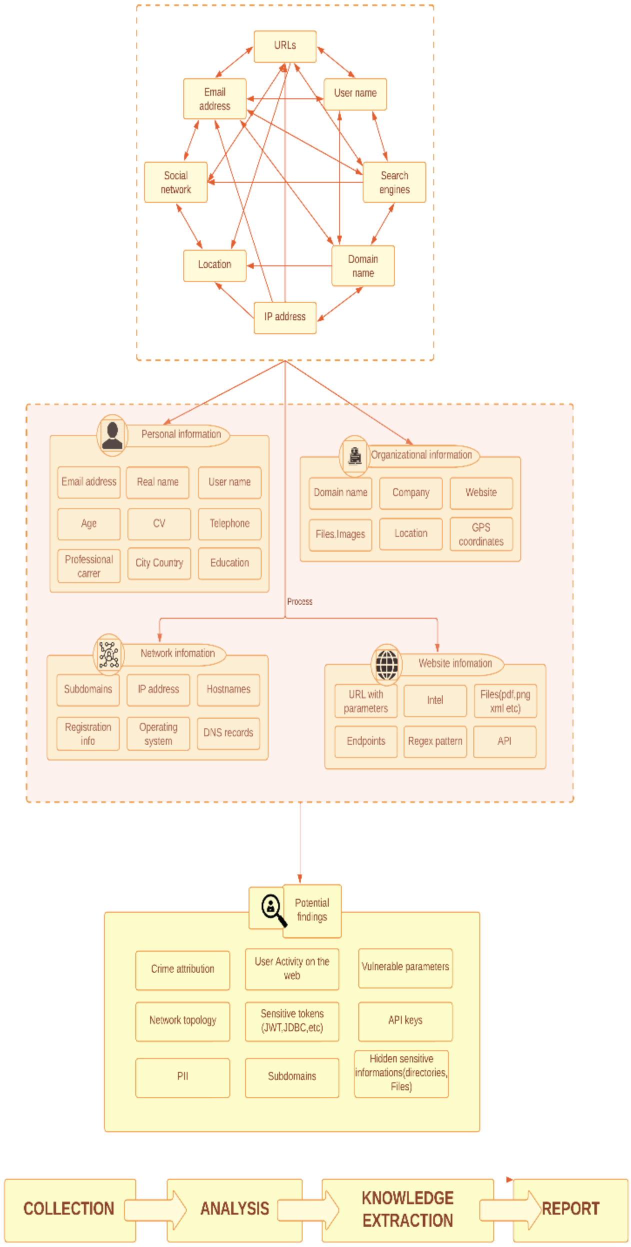

Fig 1. OSINT workflow

Information analysis: The analyst or investigator looks at the data and develops a theory based on what they observe, striving to produce actionable insights. This is known as bottom-up logic (or inductive reasoning). This might be as basic as calling a spade a spade or as complex as discovering extensive plans and machinations.

Gaining Results : The investigator might submit their reasoning and the pertinent data at the conclusion of the investigation to offer a recommendation on what should be done with the given case. A second researcher should be brought in at this point to rule out any potential bias.

Sr.no.	OSINT Techniques	Findings
1.	Subdomain Enumeration	Valid (resolvable) subdomains for one or more domain(s).
2.	User Enumeration	Social media accounts, emails, phone number, sensitive information related to user, etc.
3.	URL Footprinting	Vulnerable parameters, sensitive tokens, hidden directories, etc.
4.	Github Dorking	Fetches an individual or organization's sensitive information leaked on their public repositories.
5.	Image lookup	EXIF data, location, hidden data, etc.

Fig 2. OSINT Techniques

IV. SYSTEM MODEL

Step 1: User will register or Sign-up to the web application providing his user information like Name, Email address, Phone number, Organization, Designation, etc.

Step 2: User will login to the web application with his credentials and will be redirected to the dashboard.

Step 3: The dashboard will help user to create new projects each containing basic information like Project Name, Description, Date of Creation, Tags, etc.

Step 4: Next, the web application will display all OSINT modules available to the user. The user will then select(s) all the OSINT modules as per requirement. After that, user will provide all the appropriate inputs required by each module(s) selected by the user.

Step 5: After that, the module(s) will collect, process, and analyse the inputs provided by the user and will do operations accordingly based on the data/information available publicly on the Internet.

Step 6: At last, the web application will finally generate a report that will be available to the user profile and saved to the database.

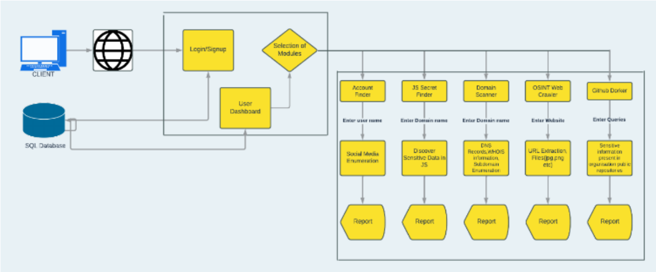

Fig 3. System Architecture

V. CONCLUSION

This project gives insight on the fundamentals of OSINT, the various types of OSINT, who uses it, and how it may be applied by various parties to gather intelligence in various situations. The project goes a little deeper into the topic and shows how various OSINT tools and approaches may be used to find information online. By the end, our goal is that every individual or organization will acquired the necessary skills to be a powerful OSINT investigator and understand how to use this Open-Source Intelligence (OSINT) automation application to gather intelligence for their own organization, intelligence that can be effectively used to further achieve their objectives.

VI. REFERENCES

[1]. Pastor-Galindo, Javier; Nespoli, Pantaleone; Gomez Marmol, Felix; Martinez Perez, Gregorio (2020). The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends. IEEE Access, 8(), 10282–10304.

[2]. Hwang, Y.W.; Lee, I.Y.; Kim, H.; Lee, H.; Kim, D. Current Status and Security Trend of OSINT. Wirel. Commun. Mob. Comput. 2022, 2022, 14.

[3]. Hoppa, Mary Ann, et al. "Twitterosint: Automated open source intelligence collection, analysis & visualization tool." Annual Review of Cybertherapy And Telemedicine 2019 121 (2019).

[4]. OSINT Framework. Available online: https://osintframework.com/

[5]. Qusef, A.; Alkilani, H. The effect of ISO/IEC 27001 standard over open-source intelligence. PeerJ Comput. Sci. 2022, 8, e810.

Cite this article as :

Vinay Kumar Kureel, Manav Arya, Aditya Kini, Suraj Maurya, Rajesh Gaikwad, "Osint Automation Application", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.377-381, March-April-2023. Available at doi :
https://doi.org/10.32628/CSEIT232551
Journal URL : https://ijsrcseit.com/CSEIT232551