

Cyber Attack Surface Management System

Vindhya L, Mahima B Gowda , Gowramma Gaari Sindhu, Keerthan V

Department of Information Science and Engineering
SJC Institute of Technology, Chikkaballapura, Karnataka, India

Abstract: *Defenders struggle to keep up with the pace of digital transformation in the face of an expanding modern enterprise attack surface and more sophisticated adversaries. A conceptual framework for relating attack surface management (ASM) to vulnerability management and cyber threat intelligence (CTI) improves cyber defense. The framework explains how ASM improves cyber resiliency in proactively detecting and responding to weaknesses that adversaries could exploit to cause unacceptable harm. Defenders should prioritize ASM aligning with the business continuity and enterprise risk management functions. A CTI-driven ASM conceptual framework (CTI-ASM) helps defenders achieve decision clarity on how best to prioritize preventing the most impactful exploitations based on adversaries' capabilities, opportunities, and intent. Security researchers have applied decision analysis methodology to solve various security challenges generally. Applying decision analysis methodology to CTI-ASM may improve the quality of its implementation and support higher quality CTI. Potentially helpful decision analysis tools and concepts include relevance diagrams, possibility and probability trees, sensitivity analysis, corporate risk attitudes, weighing imperfect information, and accounting for cognitive biases.*

Keywords: Threat Intelligence, Attack Surface Management (ASM), OpenSource Intelligence(OSINT), Penetration Testing, Spiderfoot (Tool Used)

I. INTRODUCTION

This commentary suggests that a cyber threat intelligence (CTI)-driven attack surface management (ASM) conceptual framework (CTI-ASM) may improve cyber resiliency against a continuously expanding enterprise modern attack surface (AS) and more sophisticated adversaries.

After discussing desirable features of CTI-ASM and possible technical tools, methods, and architectures to implement CTI-ASM, the commentary suggests that applying decision analysis methodology (DA) could improve both CTI-ASM and CTI. DA is a scientific method combining systems analysis and statistical decision theory for making rational decisions in complex, dynamic, and uncertain situations. Security researchers have applied DA to physical security systems, document trustworthiness, hardware security, counterfeit electronics detection, cyber system upgrades and maintenance, and intrusion detection architectures. They have more recently applied DA with graph analytics for cyber system resiliency, rank-weight methods for multi-criteria decision analysis, dynamic information processing, and simulations to improve risk thinking.

Open sources possess much of the information needed to understand the physical and human factors of the operational environment of unified land operations. Physical and human factors of a given operational environment can be addressed utilizing publicly available information to satisfy information and intelligence requirements and provide increased situational awareness interrelated with the application of technical or classified resources.

The world is being reinvented by open sources. Publicly available information can be used by a variety of individuals to expand a broad spectrum of objectives. The significance and relevance of open-source intelligence (OSINT) serve as an economy of force, provide an additional leverage capability, and cue technical or classified assets to refine and validate both information and intelligence.

Many organizations considered the data that accumulated on web space as a result of digital transformation to be a gold mine because it could have been transformed into knowledge and intelligence. As the number of open source data is available on the public domain rose, the necessity of artificial intelligence in the OSINT process became evident.

Penetration Testing is also referred to as pen test. National Institute of Standards and Technology (NIST) defines Penetration Testing as the security testing which imitates cyber-attacks to identify the vulnerabilities of a system or a network before they can be taken advantage of by adversaries in the real world. Weissman (1995) called Penetration

Testing “a pseudo-enemy attack by a friendly evaluation team on a computer system of interest to discover ways to breach the system's Security controls, to penetrate the security perimeter of protection to obtain sensitive information, to obtain unauthorized services, or to cause damage to the system that denies service to legitimate users”. The UK National Cyber Security Center defines Penetration Testing as "A method for gaining assurance in the security of an IT system by attempting to breach some or all of that system's security, using the same tools and techniques as an adversary might" (“Penetration Testing”, 2017). As pointed out by Bishop (2007), the “aspect” being tested need not be a computer system or network. It can also be a building, or a combination of people, an office, and a computer system.

II. LITERATURE SURVEY

Since last two decades, a large number of researchers have published papers in the Open source Intelligence domain. With the advancement of technology, opensource intelligence also evolved. The massive data that has accumulated in the public domain because of the evolution of social mediaplatform has enticed many actors including business corporations, antisocial elements, government agencies, law enforcement agencies etc., to integrate opensource intelligence for their benefit. Thanks to Internet Accessibility that people can now readily find and post any type of information (Edwards et al., 2017) [26].

Lee & Shon (2016) [27] proposed a new framework for cyber security threat inspection of critical infrastructure based on an Open source intelligence. This framework included four steps: developing an opensource intelligence plan, preparing opensource intelligence, gathering information from open source platforms, and producing security intelligence.

Hayes & Cappa (2018) [28] have demonstrated that OSINT may be used to do risk assessments for the company in order to prevent potential cyber-attacks on its critical infrastructure, which was part of the US electrical grid. A vulnerability assessment and various such open-source intelligence analysis procedures were carried out in order to profile the company's network, applications, devices, and critical IT resources.

Johnsen & Franke (2019) [31] discussed their research on text preprocessing needs and document formation for the Latent Dirichlet Allocation (LDA) a popular topic model algorithm, in which they propose repeated preprocessing processes, such as removing common terms until the result comprises cohesive and clear subjects. Data cleansing at preprocessing stage would help opensource intelligence to analyses and produce a reliable result.

Herrera-Cubides et al., (2020) [32] conducted a study with an aim to investigate the evolution of production of research and study material in OSINT platform. This analysis looks at two of the material sources of OSINT such as research knowledge distribution databases and repositories pertaining to educational resources. This study provides academics with a roadmap to the current level of OSINT research and teaching, as well as a valuable metadata description in order to make resources more accessible and reusable in the educational ecosystem.

Fleisher (2008) [33] presented a conceptual paper on how the growing popularity of open-source data and information affects competitive and marketing intelligence. This is a descriptive conceptual article that builds arguments from a review of three uncategorized collections of material in competitive and marketing intelligence, processing of intelligence, and analysis of market. The problems they confront in exploiting this data are described in this article, as well as the effective strategies that certain firms have exhibited in incorporating and integrating open sources in the analysis processes of competitive and marketing intelligence field. It can be seen that the study was conducted from the standpoint of a marketing analyst and the usefulness of intelligence derived from OSINT for the benefit of improving marketing efforts, rather than from the viewpoint of the individual who specialize in collecting the said data.

Magalhães&Magalhães (2019) [73] suggested TExtractor, an OSINT tool that will make gathering details concerning cyber threats easier. TExtractor is a tool that extracts text from video/audio in public sources and searches for keywords associated with harmful actors' activities. The findings are provided in the study, and they reveal that a tool such TExtractor can discover allusions to cyberattacks on audio/video sources in real time with such an accuracy of 60% to 70%.

TExtractor could also be used to keep track of a brand or automate the clipping process, which involves finding brand or product references in audio or video channels.

Kanta et al., (2020) [74] studied the possibility for Open Source Intelligence (OSINT) to be used for greater effective password cracking is investigated in this paper. A detailed review of the literature on strong passwords, cracking of

password, and OSINT is presented, as well as the legal issues that these topics raise. A study of password complexity as well as demographic characteristics that influence password choosing is also offered. Finally, the impact of OSINT by a law enforcement on password cracking is explored.

Kang (2020) [75] in order to quantify cyber threats, the authors offer the assessment variables for cyber threats among cyber-attack databases and analyze the priority of those elements. As evaluation variables for cyber threats, he choose the objective of the assault, attack type, target, convenience of attack, attack durability, frequency of OSINT database, and elements of the lowest layer of each component. The priority of each element is assessed only by using the analytic hierarchy process after it has been chosen.

III. METHODOLOGY

The use of Artificial Intelligence in OSINT not only improved the reliability but also speed of the process. Some AI sub functions have been employed in OSINT tools, however the degree of implementation and optimal use of AI capability must be investigated. The purpose of the literature review is to gain a better understanding of OSINT's current state and to identify the applications that rely on it. The integration of artificial intelligence with OSINT, as well as its role in cyber security, were also investigated. The goal of the project was to see how prepared OSINT is for Web 3.0 which will allow for the creation of a global data warehouse in which any data format may be shared and interpreted irrespective of any device over any network .

3.1 Open Source Intelligence (OSINT)

Opensource intelligence (OSINT) is intelligence gathered from publicly available data sources such as academic publications, journals, social media sites, online communities, and newspapers, among others. OSINT has gained popularity in the modern era as a result of the internet revolution, which has resulted in the accumulation of massive amounts of data on the Internet. Social media posts, blogs, journals, published articles, newspapers, video-audio files, online forums, discussion groups, company websites, government documents, maps (Klaus et al., 2020) [13] (John et al., 2007) [14], and so on are all sources of data for modern-day OSINT.

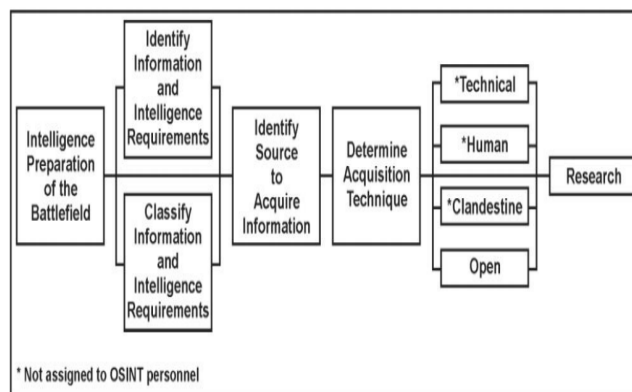


Fig 1 – Process for collecting openly available information

Governments and intelligence services are increasingly relying on OSINT to conduct investigations and combat cybercrime (Nouh et al., 2019) [15]. Open source intelligence not only finds and gathers information, but it also searches, chooses, and extracts relevant material from microblogging sites, and then analyses the information to provide an intelligent report (Koops et al., 2013) [16]. OSINT employs a systematic methodology to extract valuable intelligence from raw data and present it in the form of a usable intelligent report. The entire OSINT activity can be divided into four parts: data collection, data processing, data exploitation, and data production or extraction.

In this paper as shown in Fig.1 below four major steps, collection, processing, exploitation, and production, have been considered and explored for the purpose of a literature review (Williams 2018) [34].

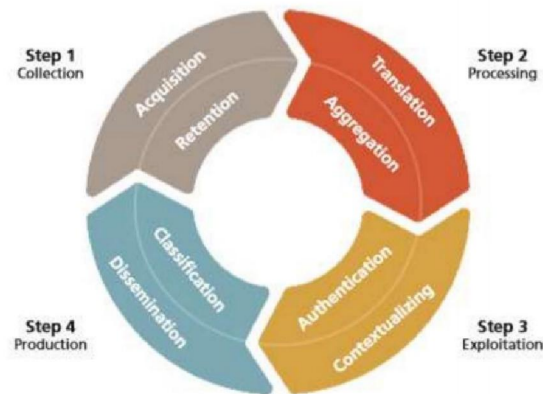


Fig 2 – OSINT Operation Cycle.

Collection of Data

Data is an important asset for carrying out any kind of intelligence activity. Just like any other intelligence method, OSINT heavily depends on the data, which are extracted from publicly available sources. In this phase one has to identify the source of data and type of data to be collected

Processing of Data

During the data collection phase in order to make it usable for analysis. Filtering out irrelevant data, translating texts from another language into English, converting photos, audio, and video files into useful data, and so on are all tasks carried out during the processing phase. Among the most prominent parallel processing models, such as MPI, General Purpose GPU (GPGPU), MapReduce, and MapReduce-like, a work has studied MapReduce to learn how to improve its performance when processing large amounts of data. It is also covered how to employ algorithms and parallelization techniques to improve scalability and performance when processing big data.

Exploitation of Data

Exploitation also known as analysis phase is in charge of determining whether the material processed in the preceding phase is what it claims to be and how valuable it is to the Intelligence community. Exploitation phase comprise of threesteps such as authentication, credibility evaluation and contextualization. Verification of authenticity and credibility of information plays a key role in developing a trustworthy knowledge. Contextualizing entail assembling several open source information pieces from any source into an output that gives a comprehensive understanding of a subject (Williams2018) [34]. Most commonly used methods for analysis are lexical analysis, semantic analysis, geospatial analysis, and social media analysis.

3.2 Penetration Testing .

Penetration Testing is also referred to as pen test. National Institute of Standards and Technology (NIST) defines Penetration Testing as the security testing which imitates cyber-attacks to identify the vulnerabilities of a system or a network before they can be taken advantage of by adversaries in the real world.

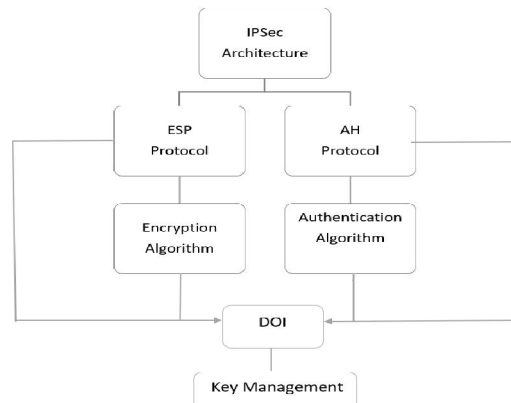


Fig 3 – Flow chart

It can also be a building, or a combination of people, an office, and a computer system. Osborne (2006) has defined pen testing in his book as “A test to ensure that gateways, firewalls and systems are appropriately designed and configured to protect against unauthorized access or attempts to disrupt services”. The goal of a Penetration Test is to certify the effectiveness of the security measures taken by an organization to protect their system. Penetration testing achieves this by discovering vulnerabilities by simulating an attack by adversaries.

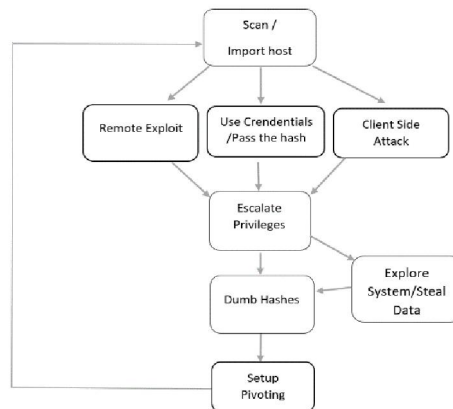


Fig 4 – System Architecture

Penetration Testing Processes

There are many different processes for penetration testing. Depending on the needs of the entity that requires the pen test, a specific process is chosen. According to Thorsen, Nufryk, & Taylor, (2019), there are eight phases in a traditional Penetration Testing Process:

Phase 1: Planning

This is the first step in the process of Penetration Testing. Scope of the Pen test is defined in this step. Tiller (2011) stated that the scope and scale of the test is decided based on factors like existing security policies, culture, laws and regulations, best practices and industry requirements. This is a very important step because it defines the entire test and guides the deliverable of the test.

Phase 2: Reconnaissance

This step is the information gathering stage where a pen tester gathers all the information he can about the organization or the system that is to be pen tested, in the hopes that this information can be useful during the attack. This information gathering can be passive information gathering and deliberate information gathering.

Phase 3: Scanning

Also known as vulnerability scanning, this stage is when a pen tester uses scanning tools to scan for vulnerabilities in a target system. (Thorsen, Nufryk& Taylor, 2019).

Phase 4: Gaining Access

Using the knowledge gained from reconnaissance and exploiting the vulnerabilities discovered in scanning, a pen tester starts attacking the target system to gain access into that system. (Thorsen, Nufryk& Taylor, 2019).

Phase 5: Maintaining Access

Once the pen testers gain access to the system in the previous stage, they use various In this stage, pen testers analyze all the information acquired during the testing process, along with the vulnerabilities discovered and also 1718 suggest remediation measures to counteract the identified vulnerabilities (Thorsen, Nufryk& Taylor, 2019).

Phase 8: Reporting

This is the stage where all the information collected in the previous stages is formally reported to the company stakeholders. This report usually consists of vulnerabilities discovered, sensitive data accessed, time taken for the pen test and suggested remediation solutions.

3.3 Spider Foot

Spiderfoot is a free and open-source tool available on Github. This tool is a framework written in the Python You must have python installed in your Kali Linux operating system to use this framework. Spiderfoot is used for reconnaissance. Spiderfoot uses different modules for information gathering. Spiderfoot is capable enough to gather information about the target host through active and passive scanning options available on the Spiderfoot framework. In the Spiderfoot framework different scanning options and modules available to set and scan the target host. Spiderfoot is an Open Source Intelligence and Information Gathering Tool. Spiderfoot is capable of doing everything almost you need for reconnaissance as per your need. Spiderfoot works as an open-source tool intelligence tool. It integrates with just about every data source available and utilizes a range of methods for data analysis, making that data easy to navigate. Spiderfoot has an embedded web server for providing an intuitive web-based interface, but you can also do the same using a command-line interface.

Features of Spiderfoot:

- Spiderfoot is a free and open-source tool available on Github.
- Spiderfoot works as a framework cum tool.
- Spiderfoot framework is written in python language.
- Spiderfoot can be used for reconnaissance.
- Spiderfoot contains many modules. As it's a framework that uses modules for information gathering.
- Spiderfoot works on the principles of OSINT.
- Spiderfoot is an automated OSINT Framework.
- Spiderfoot automates the reconnaissance processes.

Uses of Spiderfoot:

- Spiderfoot is used for reconnaissance.
- Spiderfoot is used for information gathering.
- Spiderfoot is working as a scanner for active and passive scanning on target.
- Spiderfoot can be used for domain footprinting.
- Spiderfoot can be used to find the phone numbers, email addresses of the target.
- Spiderfoot can be used to find bitcoin addresses.
- Spiderfoot can be used to save all the information gathering summary.
- Spiderfoot can be used to create graphs of scanning done by Spiderfoot.

- Spiderfoot can be used to automate GitHub all the information gathering processes.

IV. RESULT

The objective of this project is to test the effectiveness of the security measures in face of the attack and gives the information about the data that can be obtained when the attack is penetrated.

The results of the above scan is used to test the defensive parameters set by the product or the company by providing visibility.

This project is used to is to scan the target in terms of the number of iscovered and the time taken by the tool to discover those ports.

It helps in finding the flaws in the security measures that created by the company or organization.

The OSINT tools are used to obtain intelligence about their potential target during the research. When an analyst uses the correct OSNIT tool, he or she may give a more accurate intelligence report. The OSINT technologies employ artificial intelligence to locate confidential data on the internet.

The OSNIT tools fulfil three functions, however each one focuses on a different aspect. Firstly, locating assets that are visible to the public, then collecting sensitive details from outside the organization and finally turning it into meaningful intelligence.

Test that has been poorly scoped fails to achieve the goal of pen testing, even if it did meet a compliance or government requirement. However, some organizations with genuine problems like budget which make them compromise on their scope really suffer from getting the full benefit from the pen test. So, instead, if some or most of the tasks in pen testing were automated in the future, requiring very little to no human interaction, it could greatly benefit everyone who has a problem with this limitation.

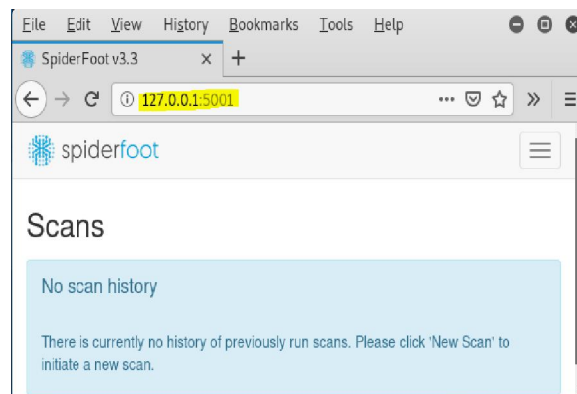


Fig 5 – Result 1

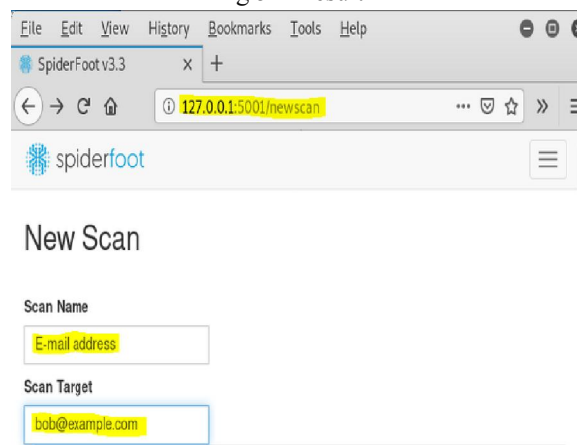
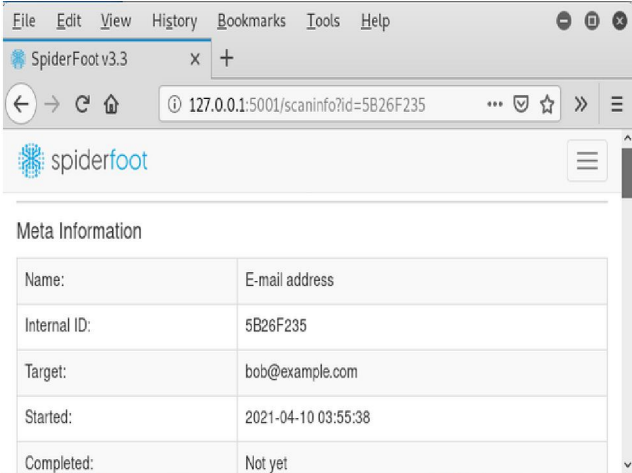


Fig 6 – New Scan

You can see a web page has been opened. This is a tool that is running on port 127.0.0.1:5001. There is a dashboard of the tool. The dashboard contains scan history, new scan, and setting options. For fresh installation, there is no

previous scan history. If we click the new scan tab, we see option to start the new scan along with the target seed field. The target seed field can be a target IP address, a domain name, or a sub-domain name. There are 3 types of configuration settings to define the scanning process. These are scan-by-use cases, required data, or modules. Each configuration setting has a number of options to choose from. For example, scan by use cases allows both, active and passive scanning of the target. It also gives the option to scan for all possible information or a range of information about the target.



Meta Information	
Name:	E-mail address
Internal ID:	5B26F235
Target:	bob@example.com
Started:	2021-04-10 03:55:38
Completed:	Not yet

Fig 7 – Scan Information

V. CONCLUSION

OSINT techniques for basic searches were presented, as well as the most sophisticated OSINT tools available today for advanced investigations. The study revealed that the degree of AI implementation still needs to be improved in order to achieve a fully automated solution with no human intervention in the decision-making stage. The goal is to effectively implement Artificial Intelligence in OSINT practice in order to improve overall performance and minimize misinterpretations that may occur as a result of the limitations that we reviewed. It is used along with the penetration testing tools and helps in the testing the security parameters that set the company. Different pieces of information using this framework. Spiderfoot is open source intelligence tool. Spiderfoot is capable of doing everything almost you need for reconnaissance as per your need. Spiderfoot works as open-source tool intelligence tool. It integrates with just about every data source available and utilizes a range of methods for data analysis, making that data easy to navigate. Spiderfoot has an embedded web server for providing an intuitive web-based interface but you can also do the same using the command-line interface.

VI. ACKNOWLEDGEMENT

We would like to express our gratitude to our College, SJC Institute of technology our Guide prof. Vindhya L ma'am and our project coordinator Aravind Tejas Chandra who have provided us with the opportunity to work on this project and given us support with guidance to make this project a success. We would also like to thank our teammates for their

REFERENCES

- [1]. HussenMaulud, D., Zeebaree, S. R., Jacksi, K., Mohammed Sadeeq, M. A., & Hussein Sharif, K. (2021). State of art for semantic analysis of natural language processing. *Qubahan Academic Journal*, 1(2), 21-28.
- [2]. Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y., Gelbukh, A., & Zhou, Q. (2016). Multilingual sentiment analysis: State of the art and independent comparison of techniques. *Cognitive Computation*, 8(4), 757-771.
- [3]. Charalambous, E., Kavallieros, D., Brewster, B., Leventakis, G., & Koutras, N. (2016). Combatting Cybercrime and Sexual Exploitation of Children: An Open Source Toolkit. In *Open source intelligence investigation: From strategy to implementation* (pp. 233-249). essay, Springer.

- [4]. Pastrana, S., Hutchings, A., Caines, A., & Buttery, P. (2018). International Symposium on Research in Attacks, Intrusions, and Defenses. In *Research in attacks, intrusions, and Defenses: 21ST International Symposium, RAID 2018, Heraklion, CRETE, Greece, September 10-12, 2018, proceedings* (Vol. 11050, pp. 207–227). Cham, Switzerland; Springer.
- [5]. Ponder-Sutton, A. M. (2016). The Automating of Open Source Intelligence. In *Automating open source intelligence: Algorithms FOR OSINT* (pp. 1–20). essay, Elsevier/Syngress.
- [6]. Benes, L. (2013). OSINT, New Technologies, Education: Expanding Opportunities and Threats. A New Paradigm. *Journal of Strategic Security*, 6(3), 22–37.
- [7]. Layton, R., & Watters, P. A. (2016). The Automating of Open Source Intelligence. In *Automating open source intelligence algorithms FOR OSINT* (pp. 1–17). essay, Syngress.
- [8]. Santarcangelo, V., Oddo, G., Pilato, M., Valenti, F., & Fornaro, C. (n.d.). Social Opinion Mining: An Approach for Italian Language. In *Future internet of things and Cloud (FICLOUD), 2015 3rd International conference on* (pp. 693–697). Rome, Italy.
- [9]. Hassan, N. A., & Hijazi, R. (2018). The evolution of open SourCeintelligenCe. In *Open source intelligence methods and tools a practical guide to online intelligence* (pp. 11–11). essay, APRESS.
- [10]. Azevedo, R., Medeiros, I., & Bessani, A. (2019). PURE: Generating Quality Threat Intelligence by Clustering and Correlating OSINT. In *2019 18th IEEE International Conference on Trust, Security and Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (pp. 483–490).
- [11]. Bruwer, R. (H.), & Rudman, R. (2015). Web 3.0: Governance, risks and safeguards. *Journal of Applied Business Research (JABR)*, 31(3), 1037.
- [12]. Glassman, M., & Kang, M. J. (2012). Intelligence in the internet age: The emergence and evolution of open source Intelligence (OSINT). *Computers in Human Behavior*, 28(2), 673–682.
- [13]. Klaus, S., Franziska, S., & Reiner, C. (2020). Conception and implementation of professional laboratory exercises in the field of open source intelligence (OSINT). *Society for Imaging Science and Technology*, 2020(3), 1-99.
- [14]. John, D. S. M., Goodchild, M. F., & Longley, P. (2007). In *Geospatial analysis: A guide to principles, techniques and software tools* (pp. 39–39). essay, Matador.
- [15]. Nouh, M., Nurse, J. R. C., Webb, H., & Goldsmith, M. (2019). Cybercrime investigators are users too! understanding the socio-technical challenges faced by law enforcement. *Proceedings 2019 Workshop on Usable Security*, 1-11.
- [16]. Koops, B.-J., Hoepman, J.-H., & Leenes, R. (2013). Open-source intelligence and privacy by design. *Computer Law & Security Review*, 29(6), 676–688.
- [17]. Layton, R., & Watters, P. A. (2016). The limitations of automating OSINT: understanding the question, not the answer. In *Automating open source intelligence algorithms FOR OSINT* (pp. 159–169). essay, Syngress.
- [18]. Bar-Ilan, J. (2001). Data collection methods on the Web for infometric purposes — A review and analysis. *Scientometrics*, 50(1), 7–32.
- [19]. Gibson, H., Ramwell, S. S., & Day, T. (2016). Analysis, Interpretation and Validation of Open Source Data. In *Open source intelligence investigation from strategy to implementation* (pp. 95– 110). essay, Springer-Verlag.
- [20]. Gibson, S. D. (2014). Exploring the Role and Value of Open Source Intelligence. In *Open source intelligence in the twenty-first century: New approaches and* (pp. 9–23). essay, Palgrave Macmillan.