https://doi.org/10.58245/ipsi.tir.2502.05

Application of OSINT Methods in Ensuring Cybersecurity

Sabina Szymoniak; Kacper Foks; and Aleksandra Pyrkosz-Dziubczyk

Abstract: ***Open-source intelligence (OSINT) is a key aspect of contemporary cybersecurity, allowing the collection and analysis of publicly available data to discover potential threats and vulnerabilities within organizations. The increased availability of open data—everything from social media usage to hacked credentials—has made OSINT a must-have for both cyber defence and offence. However, one of the most significant challenges comes from the unregulated nature of open-source data. While existing cybersecurity solutions rely on OSINT for threat intelligence, they are often marred by information overload, dubious data sources, and legal or ethical concerns. Furthermore, the same OSINT methods are leveraged by cybercriminals to gather intelligence on prospective victims to make it easy for them to conduct advanced attacks like spear phishing and reconnaissance-driven intrusions. This paper examines how OSINT can effectively enhance cybersecurity without its drawbacks. By learning from various OSINT tools, techniques, and practical applications, we propose a structured approach to filtering and verifying intelligence to improve threat detection efficiency. Unlike existing solutions, this approach includes leveraging automated processes and contextual verification to limit false positives and stay ethical.***

Index Terms: ***cybersecurity, open-source intelligence, OSINT techniques, OSINT tools***

1. Introduction

Cybersecurity is a constant battle between defenders and attackers. Organizations face new threats daily, and cybercriminals improve their attack techniques. One key tool that can tip the scales in favour of defence is Open-Source Intelligence (OSINT)—the analysis of publicly available data, which allows for identifying threats and discovering vulnerabilities before unauthorized persons use them [14, 57, 15, 35].

OSINT is about collecting information and allowing it to be filtered and interpreted. In a world where a huge part of our lives occurs online, publicly available data can reveal a surprising amount—from an organization's internal structures through information about its employees to potential entry points for attackers. Cybercriminals also use OSINT to prepare phishing attacks, analyse victim systems, or search for accidentally disclosed confidential information. This is why the conscious and effective use of OSINT in cybersecurity is more important than ever today [33, 20, 19, 55].

In this article, we will examine OSINT methods used in cybersecurity—both defensive and offensive actions. We'll discuss the tools used to collect and analyze data, the challenges of ethics and disinformation, and the legality of using these techniques. OSINT is a powerful weapon—the question is, can we use it wisely?

Our work is based on a review of open data sets, OSINT automated tools, and an analysis of cyber attacks in which the OSINT played a key role. We analyze the effectiveness of different OSINT techniques in detecting security risks and discuss their application's effect on detectors and attackers.

The report acknowledges OSINT's two-sided nature as a benefit and a risk in cybersecurity. Although it has a considerable positive impact on threat intelligence, malicious actors' use indicates that the area needs additional verification mechanisms and ethical data handling.

S. Szymoniak is with the Department of Computer Sciences, Czestochowa University of Technology, Poland (e-mail: sabina.szymoniak@icis.pcz.pl).

A. Pyrkosz-Dziubczyk is with the Department of Computer Sciences, Czestochowa University of Technology, Poland (e-mail: a.pyrkosz-dziubczyk@pcz.pl).

With artificial intelligence and big data analytics being developed in the evolving field of OSINT, future use of OSINT in cybersecurity will only increase, necessitating a specific OSINT approach with the organization to keep up with emerging threats.

This paper is a revisited and extended version of the conference paper [58]. The rest of this paper is organized as follows. Section 2 provides an overview of OSINT, discussing its key principles, applications, and significance in various domains. Section 3 explores OSINT methods used in cybersecurity, including data collection techniques, tools, and their role in identifying threats and vulnerabilities. In Section 4, we examine the challenges and threats associated with OSINT, such as privacy concerns, misinformation, and the potential for misuse by cyber criminals. Section 5 discusses the future of OSINT in cybersecurity, emphasizing advancements in AI, automation, and integration with emerging technologies. Finally, Section 6 summarizes our study's key findings and implications, highlighting the need for a balanced and ethical approach to OSINT in cybersecurity.

2. Open-Source Intelligence

An OSINT practitioner gathers legally and morally sound information from publicly available sources such as social media, government records, commercial databases, and online scanning utilities. Illegitimate practices like these are precluded from OSINT, password cracking, or manipulation [34, 7, 9, 43].

OSINT methods are used in business, law enforcement, and national security. Business firms use them to check the credibility of their partners by studying the legal, financial, and economic factors, etc. Government agencies, meanwhile, use OSINT to evaluate political dynamics and formulate state policy. OSINT assists in the reconnaissance phase, particularly cybersecurity, through target analysis and anomaly identification throughout penetration testing [32, 28, 8].

Attacks carried out by hackers and other offenders sometimes utilize OSINT techniques to collect information made in social media, enabling them to perform tailored attacks, like spear phishing. Spear phishing is phishing designed to target a particular person or organization. The attacks depend on public records, like birthdays or project information, to validate the messages. Such attacks can lead to data theft, identity theft, or the introduction of malware [23, 65].

Within OSINT, several categories can be distinguished, each focusing on publicly available information sources. SOCMINT (Social Media Intelligence) involves X, Facebook, LinkedIn, Instagram, and TikTok data. Coursework undertaken will allow students to transform geospatial data into actionable insights with practical applications in a variety of fields of interest, including defence, spatial epidemiology, urban planning, environmental monitoring, and many more. Users of this category collect information on the movements of armed forces, critical infrastructure, and natural catastrophes and geo-locate them. Much of the human intelligence (HUMINT) data is published for public consumption, from blogs and forums to interviews and podcasts. You will have access to information covering a more excellent geographical range. This will allow for a more complete view of the global situation [27, 50, 22, 21].

3. OSINT Methods in Cybersecurity

OSINT is highly important in cybersecurity because it enables organisations to identify threats, analyse risks, and monitor cybercriminal activities. Using public data, an organisation can prepare for eventual attacks and reduce the chances of data breaches or leakage. Sources on hacker forums, blogs, or even social media may be monitored for a fast indication about new attack methods, cybercriminal tools, and software vulnerabilities that will allow a faster reaction to new emerging threats [45, 47].

A significant way to do this is by investigating the disclosure of a company's information, which includes employees or customers, on the surface, deep or dark web. This way, the leaks can be detected before they are used for phishing or infrastructure compromise. They also watch for new vulnerabilities from repositories

like the National Vulnerability Database or from sites like GitHub [27, 64].

OSINT is also used to watch cybercriminal groups, including APTs, by analysing publicly available information, usually obtained from dark web forums. We can use this information to predict the likelihood of various targets and helps organisations to get ready for future attacks. Furthermore, OSINT assists in managing the organisation's reputation by identifying and responding to potential crises through monitoring social media, blogs, and forums [67, 19, 63, 46].

Image search is another form of OSINT that can be used to confirm the authenticity of an image, detect tampering, or identify a person or an event. These tools are helpful in the fight against fake news and the verification of images, especially in cases of social manipulation or cyberattack misinformation. Farhangian et al. in [17] compared the effectiveness of reverse image search engines Google, Bing and Yandex and identified their effectiveness and challenges in different situations.

Cybercriminals' use of fake news presents a significant challenge in the form of wrong directions of public opinion or creating chaos. Graph Attention Neural Event Embedding and Convolutional Neural Networks are AI methods that effectively identify and refute fake news. These tools enable organisations to distinguish between real threats and risks from misinformation and address them [39, 12, 26, 42].

In the end, OSINT methods assist in identifying the traffic that is produced by Domain Generation Algorithms, which are used in malware campaigns to avoid detection. Explainable AI combined with OSINT, as introduced by Suryotrisongko et al., enhances the detection of DGA-based traffic and produces effective and explainable solutions for this cybersecurity problem [56].

In cybersecurity, many OSINT tools help gather and analyze information from public sources to identify vulnerabilities, evaluate risks, and determine possible attack points. Amass is a tool that assists in domain enumeration and infrastructure mapping to find subdomains, DNS records, IP addresses and network data from various sources. This tool is most useful for penetration testers when carrying out a pre-attack survey of large networks. Another popular option is Censys. It is also a search engine like Shodan but is specific to understanding the Internet's physical structure. It collects data on SSL certificates, network services and device configurations and helps identify misconfigured or vulnerable assets [25, 3].

Moreover, Metagoofil is a tool used to gather metadata from public documents (for example, PDF, DOC) on websites. It can reveal information such as users, paths to files, or versions of the used software. Datasploit is one more OSINT tool that provides tools for unified data collection from email, social media, domains and IP addresses and for creating threat intelligence reports. Finally, IVRE (Interactive Visual Reconnaissance Environment) provides extra capabilities for network scanning and data visualization, which makes it a perfect choice for mapping the infrastructure and potential risks in an enterprise network [31, 19]

Other OSINT tools in cybersecurity are Maltego, theHarvester, SpiderFoot, OSINT Framework, and Recon-ng [58].

4. OSINT Challenges and Threats

We can identify several key challenges and threats associated with OSINT. OSINT is based on collecting information on individuals or companies from publicly available sources. However, gathering and analysing in mass can violate people's privacy. Social media posts, forum discussions, photos, and application data are sources of information that can be used to build a detailed user profile. It can lead to cases where this data is abused, for example, in phishing campaigns, identity theft or harassment. In addition, if the material gathered from the public domain is combined with proprietary information (for instance, data leaks), it can paint a comprehensive picture of a person's life or the works of an organisation.

Table 1 summarizes the key challenges and threats identified in OSINT and their potential to cause harm. Although OSINT is the process of deriving information from the collection of publicly available information, it poses a severe privacy risk, especially if combined with

Table 1: Key challenges and threats related to OSINT based on the literature

Challenge/Threat	Description	Ref.
Privacy Violations	OSINT enables mass collection and analysis of publicly available data, leading to detailed profiles of individuals or organizations. This data can be misused for phishing, identity theft, or harassment. When combined with private information (e.g., data leaks), it provides a comprehensive picture of a person's life or organizational structure.	[66]
Legal and Ethical Issues	OSINT poses legal and ethical challenges, such as GDPR violations and jurisdictional differences. Misuse of illegally shared data, e.g., from forums containing stolen information, raises questions about legality. Subtle boundaries exist between legal data collection and illegal activities like password cracking.	[59, 29, 54]
Credibility of Information	Open-source intelligence can contain false, outdated, or manipulated data. Cybercriminals or nation-states may spread disinformation, leading to incorrect cybersecurity decisions. Public sources alone may provide insufficient or misleading analyses without integration with other forms of intelligence.	[68]
Challenges of Automation	Automation increases efficiency but introduces risks, such as information overload, difficulty in filtering critical data, and the potential misuse of automated tools by cybercriminals for attacks like mass phishing campaigns.	[30]
Resource Intensity	Effective OSINT requires significant resources, including specialized tools and experts. Automation cannot fully replace human analysis. Tools require updates, management, and training, incurring hidden costs.	[13]
Escalation of Cyber Conflicts	Misuse of OSINT, such as monitoring competitors, can lead to unintended consequences like increased cybercrime or surveillance races.	[66]
Misuse by Cybercriminals	Cybercriminals exploit OSINT to gather information about potential targets, identify vulnerabilities, and understand organizational technologies and strategies for malicious purposes.	[30]

private or leaked data. The use of its legal and ethical issues is further complicated by jurisdictional variations and regulations such as GDPR [18] as barriers. However, the information from the open source is also relatively untrustworthy and may be ancient, fake or otherwise manipulated. Cybercriminals can misuse the efficiency gain acquired through automation in OSINT, but that comes at the price of overwhelming amounts of information. Using OSINT also entails a significant resource investment, requiring unique tools and skilled personnel to analyze them, which may entail hidden costs. In addition, OSINT can be misused to escalate cyber wars or for evil use, thus highlighting the need to use OSINT prudently and moderately.

It should also be noted that the misuse of OSINT can worsen cyber wars. For instance, a corporation can gain knowledge of its competitor's actions through OSINT, which may have unforeseen consequences such as an increase in cybercrime or a competitive surveillance race.

OSINT techniques are not just employed by cybersecurity entities or for lawful objectives. Cybercriminals employ identical techniques to collect intelligence on prospective targets, including corporations and their personnel, identify publicly accessible vulnerabilities within an organization's systems and infrastructure, or acquire information regarding technologies and strategies utilized by rival companies or nations.

Cyber threats have changed dramatically, and OSINT has had to adapt alongside them. How attackers and defenders have used open-source intelligence over time gives us a better idea of where the field is headed. Here, we can point out some significant periods connected with OSINT evolution.

The Early Days (Pre-2010). OSINT was mostly manual—security researchers and hackers had to handly sift through public records, forums, and websites. Social engineering was rudimentary, and phishing relied on generic email scams rather than personalized attacks.

Automation Takes Over (2010-2020). This period saw a major shift as OSINT tools like Shodan, Maltego, and Censys became widely available. Organizations began using automated scanning to find security gaps, while attackers improved their reconnaissance, crafting targeted phishing campaigns based on social media data.

AI and Disinformation (2020-Present). AI-driven OSINT tools now process vast amounts of data in real-time, helping analysts detect threats faster. However, cybercriminals have also stepped up their game, using deepfake technology and large-scale misinformation campaigns to manipulate public perception and bypass traditional security measures.

The Next Phase (Beyond 2025). Future OSINT will likely involve AI-powered intelligence continuously monitoring emerging threats. Blockchain analytics will play a bigger role in tracking cybercrime, and IoT devices will create new attack surfaces that require smarter OSINT methodologies. The line between OSINT, SIGINT, and CYBINT will blur as intelligence sources become more interconnected.

The challenge now is ensuring that OSINT remains effective in this rapidly changing landscape without becoming just another tool that attackers can manipulate to their advantage.

4.1 OSINT and Disinformation and Information Operations

Modern information operations use OSINT to both analyze and counter disinformation campaigns. Cybercriminals, APT (Advanced Persistent Threats) groups and state actors use open sources of information to manipulate public opinion, conduct social engineering attacks, and destabilize politics. OSINT enables the identification and investigation of disinformation operations. OSINT investigations, aimed at confirming or denying information, use the previously mentioned OSINT categories.

The first group of OSINT methods used in such situations is social network analysis. This method helps investigators identify bots [61], troll farms [41], or patterns of spreading false content [38]. Rao et al. in [44] proposed a machine learning-based approach to distinguishing real tweets from fake ones. Thus, Ilias et al. in [24] proposed two methods that improve these processes. The authors extract features that identify accounts that post automated messages in the first method. In the second method, the authors present a deep learning architecture that aims to determine whether tweets were posted by real users or generated by bots. Other solutions were proposed in [60, 5, 40]. Shet et al. in [52] proposed modelling trolling as a graph where user interactions and their content on social media are represented as a heterogeneous graph, allowing for better capture of troll behaviour patterns. The authors used Graph Neural Networks (GNN) to identify trolls based on their behavior and published content.

Fake news detection is a broadly analysed theme in cybersecurity, especially in the case of military operations, politics, or global tragedies and events (for example, [2]). Detecting fake news increases society's resilience to cyber threats and protects against information manipulation, which is crucial for cybersecurity. It is worth integrating such systems with real-time content analysis tools to counter threats effectively.

Soga et al. in [53] used the similarity of

opinions between users to detect fake news based on graph analysis, where users and their interactions are nodes and edges of the graph. Their method analyzes users' attitudes towards news articles and interactions between posts. This allows for detecting fake news and information manipulation based on how users react to the content. Brinda et al. in [10] proposed a heterogeneous deep learning model that incorporates richer information beyond text, particularly the features of the dissemination structure from social networks (facets, user interactions, sharing patterns, and network topology). With these features, the authors can derive insights into the dissemination of information. Abualigah et al. in [1] used a convolutional neural network, deep neural network, and long short-term memory for fake news detection. Esteban et al. in [16] focused on the methodology of predicting the virality of fake news based on its content.

Other methods used to prevent disinformation are as follows:

- linguistic and stylometric analysis, which allows detecting recurring patterns in fake news narratives,
- geolocation and visual analysis, which allows verifying the authenticity of photos and video recordings,
- link analysis, which allows identifying sources of disinformation.

We can use earlier tools, such as Google Reverse Image Search, Maltego, or SpiderFoot, to use these methods.

Fake news campaigns are mainly carried out as disinformation operations to influence elections. In addition, they can be used to attack the reputation of companies. Spreading false information about organizations is aimed at destabilizing the market. These campaigns use deepfakes [62] and social engineering. Here, one method is to make a fake video of the company's CEO giving false statements. In the fight against disinformation, identifying sources and verifying content are essential parts of the process, which is accomplished, more often than not, by using OSINT tools to analyze the source and detect fake accounts. You can also use AI methods to detect disinformation automatically.

OSINT is a key tool in the fight against disinformation. Analyzing content, sources, and social networks effectively identifies and combats manipulative campaigns. In the future, the development of artificial intelligence and the improved integration of OSINT with cybersecurity tools will facilitate even more effective combating of false information.

4.2 OSINT Automation – AI and LLM in Cybersecurity

OSINT automation using large language models (LLM) [6] opens up new possibilities in analysing open information sources. Their ability to process vast amounts of text data allows for lightning-fast searches of forums, social media, and the dark web for significant threats. Traditional methods required manual content filtering, which was time-consuming and prone to human error. LLMs such as GPT-4 [51] or Claude [48] allow for automatic information classification, identification of key threads, and generating precise summaries that help analysts make quick decisions.

One of the key applications of AI in OSINT is detecting disinformation and manipulation in the media space. Language models can analyse text structure and detect characteristic propaganda patterns, such as the repetition of false narratives or emotional overtones of the message. Moreover, advanced algorithms can compare analysed content with credible sources, assessing its probability of falsehood. In the face of the growing number of deepfakes and manipulated video materials, AI is becoming a key tool in the fight against digital fraud. Models based on neural networks can detect subtle irregularities in images and recordings that are invisible to the human eye [4, 37].

Equally important is the automatic generation of reports, which allows for effective information management in cybersecurity. Thanks to the integration of AI with SIEM (Security Information and Event Management) systems, anomalies can be detected in real-time, and threats can be prioritised based on their potential impact. Automating this process speeds

up incident analysis and reduces the number of false alarms, which are often a significant problem in operational environments [49, 36, 36].

Though AI is beneficial in utilising OSINT, it also has its dangers. Among them is that high-quality models can create more realistic deepfakes or tailored narratives, making them more complex and more challenging to spot. Additionally, over-reliance on automated procedures for analysis leads to blind acceptance by decision-makers of conclusions generated by algorithms, robbing them of their ability to think critically about what they are presented with for information. Thus, one of the main challenges in the upcoming years will be achieving a balance between AI and human verification to ensure that OSINT continues to be an efficient and trustworthy tool in cybersecurity.

5. The Future of OSINT in Cybersecurity

Future growth of OSINT techniques will revolutionize cybersecurity as the online threat landscape changes, the data volume expands, and technology evolves. The complexity and variety of expanding data sources require automated data gathering and analysis by sophisticated algorithms. Automation facilitates fast processing and threat detection of high-priority threats and enhances response quality and speed [11].

The use of artificial intelligence and machine learning will enhance OSINT capabilities. They can recognize anomalies and patterns and, accordingly, identify false news, monitor your social media activities for suspicious behaviour, and find out the patterns of phishing attacks. In practice, AI-powered chatbots that monitor social media and forums in real time and escalate emerging threats to security teams around the clock are becoming popular.

The growth of IoT devices will revolutionize OSINT. IoT-enabled smart homes, cars, cities, and other platforms will generate a never-before-seen amount of public data. This data type from surveillance and networked sensors will be helpful in threat assessment and monitoring. With the recent advancement in big data technologies, data availability will exponentially increase, enabling in-depth and practical analysis of big data sets such as social media posts, news stories, and domain name registrations.

Another critical focus is the creation of better tools that can collect information from several sources and analyze them in more detail to find correlations and dependencies in large data sets. OSINT will be combined with other intelligence sources, such as HUMINT, SIGINT, and CYBINT, which will give a more accurate picture of the threats and help make better decisions.

OSINT will change from being a passive indicator of threats to an active one. This will enable organisations to create new forms of surveillance that can continuously monitor open sources and give an early sign of the evolution of threats with AI and automation. These systems can provide specific threat intelligence to organizations or industries—say, informing banks of credit card fraud threats. Hence, the statement is true because AI enhances the efficiency of OSINT in both ways mentioned earlier. Thus, we can state that OSINT integrated with AI increases the chances of success in both ways: during investigation and prevention. The future of OSINT lies in greater automation, integration with AI and IoT, more advanced analytical platforms, and more proactive and personalized threat detection and prevention.

Even though they are decentralized and anonymous, Blockchain technologies are increasingly becoming the focus of great scrutiny in OSINT and cybersecurity contexts. Cryptocurrency transactions are easier to track on blockchains, but the latter also help identify cybercriminal activities, terrorist financing, and money laundering. OSINT analysts are, therefore, able to spot suspicious patterns, map out wallet connections, and even follow the movement of funds across different exchanges because all transactions on public blockchains like Bitcoin or Ethereum are solidified forever. Unlike traditional financial systems, where banks can block access to information, blockchain provides an open, transparent source of data that, when properly analyzed, can provide key evidence in digital investigations.

Using OSINT tools to analyze blockchains is already conventional in the work of law en-

forcement and cybersecurity teams. Platforms such as Chainalysis, Elliptic, and CipherTrace assist in identifying suspicious addresses, real-time analysis of transactions and tracking wallet interactions with cryptocurrency exchanges. These tools employ AI and machine learning to identify links between illegal transactions, even if the funds have been tumbled through a series of wallets to hide their origin. When combined with other forms of OSINT, for example, data from the darknet forums, analysts can link blockchain addresses to particular entities or persons, thus enhancing the effectiveness of cybercrime operations.

In the Web3 era, OSINT is becoming a key tool in combating the growing number of frauds, rug pulls, and attacks on decentralized finance platforms. Hackers increasingly use smart contracts to carry out complex fraud schemes that leave traces only on the blockchain. Using tools to visualize transactions and monitor blockchain networks in real-time, OSINT analysts can identify criminal groups' patterns and warn users about suspicious projects. In the future, integrating OSINT with blockchain technology could also help create more tamper-proof digital identification systems and increase transparency in the cryptocurrency sector.

As OSINT continues to evolve, certain elements will be critical in determining its effectiveness in cybersecurity:

- There is too much data for humans to analyze alone. AI will help process and filter intelligence, but analysts must remain involved to catch false positives and biased outputs.
- Attackers are improving at spreading false information, and OSINT needs to adapt. AI-powered verification tools will be crucial for identifying fake social media profiles, manipulated videos, and misleading narratives.
- Cryptocurrencies are a double-edged sword. While they enable financial privacy, they also facilitate cybercrime. OSINT tools capable of analyzing blockchain transactions will be vital for tracking illicit activities.
- With smart cities, connected cars, and industrial IoT devices becoming the norm, OSINT must expand its focus beyond traditional networks. This means developing tools to identify vulnerabilities in IoT ecosystems before attackers exploit them.

OSINT offers enormous opportunities but has many challenges that must be addressed to remain an effective and responsible tool. One of the main concerns is ethical and legal issues. Although data is publicly available, it should not always be used without consideration. There is a fine line between legally collecting information and violating privacy, especially in the context of personal data protection. Regulations such as GDPR are increasingly influencing how OSINT can be used, meaning that organizations must balance security and respect for privacy.

Another big problem is disinformation. Cybercriminals and state-sponsored groups increasingly use OSINT techniques to manipulate public opinion and spread false narratives. Without proper verification mechanisms, OSINT can become a tool for amplifying fake news instead of combating it. Therefore, developing tools to recognize deepfakes, detect social media bots, and assess information sources' credibility is crucial.

Automating OSINT is a big step forward but carries certain risks. More and more analyses are being performed by AI algorithms, which can make mistakes or be susceptible to manipulation. If organizations blindly trust automated reports, they may miss essential threats or – even worse – make decisions based on incorrect data. Therefore, the balance between automation and human analysis will be crucial in the OSINT process. Machines can help filter huge amounts of information, but the conclusions should always be left to the experts.

OSINT is used not only by security specialists but also by cybercriminals. Hackers use the same techniques to gather information about potential attack targets, identify vulnerabilities in company systems, and match phishing campaigns to their victims. Organizations should know how much information about them is publicly available and conduct regular security audits. Limiting the network's available data can

significantly hinder attackers' actions.

OSINT will increasingly need to integrate with other intelligence disciplines, such as SIGINT and HUMINT. Combining different data sources will give a more complete picture of threats and better decision-making.

The next step should be developing AI systems that collect information and help distinguish real threats from disinformation. In an era of deepfakes and increasingly sophisticated manipulation campaigns, quickly verifying content will be crucial.

The role of education should also not be forgotten. Even the best OSINT tools will not be practical if people do not know how to use them. Organizations should invest in training that helps analysts recognize new threats, avoid the pitfalls of automation, and apply best practices in ethics and law.

The future of OSINT also involves moving from a reactive approach to a proactive one. Instead of analyzing incidents after the fact, the goal should be to detect and neutralise threats before they become a real problem. The development of real-time threat monitoring tools and AI to predict cyberattacks can make OSINT an even more powerful tool in the fight against cybercrime.

In summary, OSINT must evolve to keep up with the changing landscape of cyber threats. The key will be to combine automation with human analysis wisely, develop tools to combat disinformation and increase awareness of threats. If we can achieve these goals, OSINT will remain an indispensable element of a cybersecurity strategy.

6. Conclusion

This study shows how OSINT will be inevitable in cybersecurity. In recent years, collecting and analyzing information from the public domain has become essential in identifying potential risks, tracking cybercriminal activities and enhancing cyber defence. The study's outcomes mentioned in this paper show that when OSINT techniques are correctly implemented, they can significantly improve the effectiveness of security operations by offering real-time and relevant information.

On the other hand, OSINT comes with its challenges. Although it is a valuable instrument in the hands of security specialists, criminals also use it to gather information and launch social engineering attacks. Automation and AI in OSINT have been improved to conduct searches faster and discover threats faster. However, as with any tool, they create new risks, such as the possibility of incorrect or intentionally wrong data provided by the adversary. Hence, experts should check the use of automated tools to avoid wrong information being used as intelligence from open sources.

Another critical topic is the ethical and legal aspects of OSINT. Issues regarding collecting and analysing publicly available information, such as privacy, data protection laws, and the correct intake of intelligence collection, exist. Security professionals will face these challenges as the field grows while using OSINT correctly to help people and organizations and not violate people's rights or ethical principles.

OSINT is expected to become even more significant in the future of cybersecurity. Applying AI, big data, and blockchain forensics will enhance the potential of OSINT, leading to better and faster analysis of threats. However, as cyber threats increase, including deepfakes and AI-generated misinformation, OSINT techniques must follow the trend. The further development of tools for intelligent analysis and the right approach to intelligence gathering will guarantee the further importance of OSINT in cybersecurity.

References

[1] Laith Abualigah, Yazan Yehia Al-Ajlouni, Mohammad Sh Daoud, Maryam Altalhi, and Hazem Migdady. Fake news detection using recurrent neural network based on bidirectional lstm and glove. *Social Network Analysis and Mining*, 14(1):40, 2024.

[2] Tanvir Ahammad. Identifying hidden patterns of fake covid-19 news: An in-depth sentiment analysis and topic modeling approach. *Natural Language Processing Journal*, 6:100053, 2024.

[3] Izzat Alsmadi, Zyad Dwekat, Ricardo Cantu, and Bilal Al-Ahmad. Vulnerability assessment of industrial systems using shodan. *Cluster Computing*, 25(3):1563–1573, 2022.

[4] Dipto Barman, Ziyi Guo, and Owen Conlan. The dark side of language models: Exploring the poten-

tial of llms in multimedia disinformation generation and dissemination. *Machine Learning with Applications*, page 100545, 2024.

[5] Maryam Bibi, Zahid Hussain Qaisar, Naeem Aslam, Muhammad Faheem, and Perveen Akhtar. Tl-pbot: Twitter bot profile detection using transfer learning based on dnn model. *Engineering Reports*, page e12838, 2024.

[6] Abeba Birhane, Atoosa Kasirzadeh, David Leslie, and Sandra Wachter. Science in the age of large language models. *Nature Reviews Physics*, 5(5):277–280, 2023.

[7] I.S. Black and L. Fennelly. *Investigations and the Art of the Interview*. Elsevier Science, 2020.

[8] Ludo Block. The long history of osint. *Journal of Intelligence History*, 23(2):95–109, 2024.

[9] Isabelle Böhm and Samuel Lolagar. Open source intelligence: Introduction, legal, and ethical considerations. *International Cybersecurity Law Review*, 2:317–337, 2021.

[10] BM Brinda, C Rajan, and K Geetha. Detecting evolving fake news in social media by leveraging heterogeneous deep learning model. In *2024 Second International Conference on Advances in Information Technology (ICAIT)*, volume 1, pages 1–5. IEEE, 2024.

[11] Thomas Oakley Browne, Mohammad Abedin, and Mohammad Jabed Morshed Chowdhury. A systematic review on research utilising artificial intelligence for open source intelligence (osint) applications. *International Journal of Information Security*, pages 1–28, 2024.

[12] Mariana Caravanti de Souza, Marcos Paulo Silva Gôlo, Alípio Mário Guedes Jorge, Evelin Carvalho Freire de Amorim, Ricardo Nuno Taborda Campos, Ricardo Marcondes Marcacini, and Solange Oliveira Rezende. Keywords attention for fake news detection using few positive labels. *Information Sciences*, 663:120300, 2024.

[13] Amir Djenna, Ezedin Barka, Achouak Benchikh, and Karima Khadir. Unmasking cybercrime with artificial-intelligence-driven cybersecurity analytics. *Sensors*, 23(14):6302, 2023.

[14] Jiří Dostál, Xiaojun Wang, William Steingartner, and Prasart Nuangchalerm. Digital intelligence-new concept in context of future school of education. In *Proceedings of ICERI2017 Conference 16th-18th November*, 2017.

[15] Joseph Downing. Social media, digital methods and critical security studies. In *Critical Security Studies in the Digital Age: Social Media and Security*, pages 71–108. Springer, 2023.

[16] Mercedes Esteban-Bravo, Lisbeth d. l. M. Jiménez-Rubido, and Jose M. Vidal-Sanz. Predicting the virality of fake news at the early stage of dissemination. *Expert Systems with Applications*, 248:123390, 2024.

[17] Faramarz Farhangian, Rafael MO Cruz, and George DC Cavalcanti. Fake news detection: Taxonomy and comparative study. *Information Fusion*, 103:102140, 2024.

[18] General Data Protection Regulation GDPR. General data protection regulation. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*, 2016.

[19] Devu Govardhan, Grandhi Guna Sai Hari Krishna, V Charan, Sribhashyam Venkata Anantha Sai, and Radhika Rani Chintala. Key challenges and limitations of the osint framework in the context of cybersecurity. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, pages 236–243. IEEE, 2023.

[20] Jarod Govers, Philip Feldman, Aaron Dant, and Panos Patros. Down the rabbit hole: Detecting online extremism, radicalisation, and politicised hate speech. *ACM Comput. Surv.*, feb 2023.

[21] Tony H Grubesic, Jake R Nelson, and Ran Wei. Drones and their future applications. In *UAVs for Spatial Modelling and Urban Informatics*, pages 149–167. Springer, 2024.

[22] Susan Henrico and Dries Putter. Intelligence collection disciplines—a systematic review. *Journal of Applied Security Research*, pages 1–25, 2024.

[23] Jhon Francined Herrera-Cubides, Paulo Alonso Gaona-García, and Salvador Sánchez-Alonso. Open-source intelligence educational resources: A visual perspective analysis. *Applied Sciences*, 10(21), 2020.

[24] Loukas Ilias and Ioanna Roussaki. Detecting malicious activity in twitter using deep learning techniques. *Applied Soft Computing*, 107:107360, 2021.

[25] Joy Winston James. Engineering the human mind: Social engineering attack using kali linux. *SN Computer Science*, 4(6):846, 2023.

[26] Kürşat Mustafa Karaoğlan. Novel approaches for fake news detection based on attention-based deep multiple-instance learning using contextualized neural language models. *Neurocomputing*, 602:128263, 2024.

[27] Varin Khera, Anand R Prasad, and Suksit Kwanoran. Open source intelligence (osint)–a practical introduction: A field manual. 2024.

[28] Alekya Sai Laxmi Kowta, Karan Bhowmick, Jeev Ratan Kaur, and N Jeyanthi. Analysis and overview of information gathering & tools for pentesting. In *2021 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–13. IEEE, 2021.

[29] Mariusz Kubanek and Sabina Szymoniak. Ethical challenges in ai integration: a comprehensive review of bias, privacy, and accountability issues. In *The Leading Role of Smart Ethics in the Digital World*, pages 75–85. Universidad de La Rioja, 2024.

[30] Philipp Kühn, Kyra Wittorf, and Christian Reuter. Navigating the shadows: Manual and semi-automated evaluation of the dark web for cyber threat intelligence. *IEEE Access*, 2024.

[31] Biresh Kumar, Sahil Prasad Bejo, Riya Kedia, Pallab Banerjee, Pooja Jha, and Mohan Kumar Dehury. Kali linux based empirical investigation on vulnerability evaluation using pen-testing tools. In *2023 World Conference on Communication & Computing (WCONF)*, pages 1–6. IEEE, 2023.

[32] Stefan Kutschera. Incidental data: observation of privacy compromising data on social media platforms. *International Cybersecurity Law Review*, 4(1):91–114, 2023.

[33] Soon Li Lee, Cai Lian Tam, and Sivakumar Thurairajasingam. Facebook depression with depressed users: The mediating effects of dependency and self-criticism on facebook addiction and depressiveness. *Computers in Human Behavior*, 139:107549, 2023.

[34] Xinyan Li, Dongxu Li, Zhihao Yang, Hui Zhao, Wei Cai, and Xi Lin. Nd-ner: A named entity recognition dataset for osint towards the national defense domain. In *International Conference on Neural Information Processing*, pages 361–372. Springer, 2022.

[35] D Manohari, ES Adithya, K Vijayakumar, et al. Information retrieval using osint and ghdb. In *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, pages 1–7. IEEE, 2023.

[36] Rahul Marri, Sriram Varanasi, and Satwik Varma Kalidindi Chaitanya. Integrating security information and event management (siem) with data lakes and ai: Enhancing threat detection and response. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023*, 6(1):151–165, 2024.

[37] Andrés Montoro Montarroso, Francisco Javier Cantón Correa, Paolo Rosso, Berta Chulvi Ferriols, Ángel Panizo, Javier Huertas Tato, Blanca Calvo Figueras, María José Rementeria Núñez, and Juan Gómez Romero. Fighting disinformation with artificial intelligence: fundamentals, advances and challenges. *El profesional de la información*, 32(3):2, 2023.

[38] Mary Luz Mouronte-López, Javier Gómez Sánchez-Seco, and Rosa M Benito. Patterns of human and bots behaviour on twitter conversations about sustainability. *Scientific Reports*, 14(1):3223, 2024.

[39] Anirban Mukhopadhyay, Sukrit Venkatagiri, and Kurt Luther. Osint research studios: A flexible crowdsourcing framework to scale up open source intelligence investigations. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW1):1–38, 2024.

[40] Lynnette Hui Xian Ng, Ian Kloo, Samantha Clark, and Kathleen M Carley. An exploratory analysis of covid bot vs human disinformation dissemination stemming from the disinformation dozen on telegram. *Journal of Computational Social Science*, pages 1–26, 2024.

[41] Serge Poliakoff. Trolls behind the mask of journalists: How yevgeny prigozhin's patriot media group was organized. *Problems of Post-Communism*, pages 1–13, 2025.

[42] Zhiguo Qu, Yunyi Meng, Ghulam Muhammad, and Prayag Tiwari. Qmfnd: A quantum multimodal fusion-based fake news detection model for social media. *Information Fusion*, 104:102172, 2024.

[43] Abdallah Qusef and Hamzeh Alkilani. The effect of iso/iec 27001 standard over open-source intelligence. *PeerJ Computer Science*, 8:e810, 2022.

[44] Paradesi Subba Rao, Farooq Sunar Mahammad, Parumanchala Bhaskar, M Shabarish, SV Kishore, T VarunKumar, B Chandra Sekhar, and SM Mansoor. Detecting malicious twitter bots using machine learning. In *AIP Conference Proceedings*, volume 3028. AIP Publishing, 2024.

[45] Thea Riebe. Privacy concerns and acceptance factors of osint for cybersecurity: A representative survey. In *Technology Assessment of Dual-Use ICTs: How to Assess Diffusion, Governance and Design*, pages 221–248. Springer, 2023.

[46] Thea Riebe. Values and value conflicts in the context of osint technologies for cybersecurity incident response. In *Technology Assessment of Dual-Use ICTs: How to Assess Diffusion, Governance and Design*, pages 157–190. Springer, 2023.

[47] Thea Riebe, Julian Bäumler, Marc-André Kaufhold, and Christian Reuter. Values and value conflicts in the context of osint technologies for cybersecurity incident response: A value sensitive design perspective. *Computer Supported Cooperative Work (CSCW)*, 33(2):205–251, 2024.

[48] Aya Sahbi, Céline Alec, and Pierre Beust. Semantic vs. llm-based approach: A case study of konpote vs. claude for ontology population from french advertisements. *Data & Knowledge Engineering*, 156:102392, 2025.

[49] Mary Jane C Samonte, Pritz Lorraine F Dorado, John Benjamin M Dulay, and Alain Marcus M Leyva. Enhancing threat detection in financial institutions with ai-driven security information and event management integration. In *2024 4th International Conference on Computer Systems (ICCS)*, pages 85–92. IEEE, 2024.

[50] Marco San Biagio, Roberto Acquaviva, Valentina Mazzonello, Ernesto La Mattina, and Vito Morreale. A new socmint framework for threat intelligence identification. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 692–697. IEEE, 2021.

[51] Katharine Sanderson. Gpt-4 is here: what scientists think. *Nature*, 615(7954):773, 2023.

[52] Advaith Shet, Deeksha Jatangi D, Nevasini Sasikumar, Satvik Agrawal, and Arti Arya. Detroll—leveraging graph neural networks with attention mechanism to detect state-sponsored trolls. In *International Conference on Information and Communication Technology for Intelligent Systems*, pages 141–151. Springer, 2024.

[53] Kayato Soga, Soh Yoshida, and Mitsuji Muneyasu. Exploiting stance similarity and graph neural networks for fake news detection. *Pattern Recognition Letters*, 177:26–32, 2024.

[54] William Steingartner and Darko Galinec. National cybersecurity strategy action plan for cyber resilience: Qualitative data and achievements. In *Smart Ethics in the Digital World: Proceedings of the ETHICOMP 2024. 21th International Conference on the Ethical and Social Impacts of ICT*, pages 233–236. Universidad de La Rioja, 2024.

[55] William Steingartner, Darko Galinec, Dávid Val'ko, and Norbert Ádám. Disinformation campaigns: Battling misinformation for resilience in hybrid threats model. *Acta Polytechnica Hungarica*, 21(10), 2024.

[56] Hatma Suryotrisongko, Yasuo Musashi, Akio Tsuneda, and Kenichi Sugitani. Robust botnet dga detection: Blending xai and osint for cyber threat intelligence sharing. *IEEE Access*, 10:34613–34624, 2022.

[57] Sabina Szymoniak and Kacper Foks. Open source intelligence opportunities and challenges–a review. *Advances in Science and Technology. Research Journal*, 18(3), 2024.

[58] Sabina Szymoniak and Kacper Foks. Role of open source intelligence methods in cybersecurity. In *Proceedings of the INFORMATICS 2024, IEEE 17th International Scientific Conference on Informatics, Poprad, Slovakia, 13-15 November 2024*, pages 401–406. IEEE, November 2024.

[59] Sabina Szymoniak and Mariusz Kubanek. Ethics in internet of things security: challenges and opportunities. In *The Leading Role of Smart Ethics in the Digital World*, pages 123–133. Universidad de La Rioja, 2024.

[60] Kunal Umbrani, Deven Shah, Amit Pile, and Anamika Jain. Fake profile detection using machine learning. In *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, pages 966–973. IEEE, 2024.

[61] Gitika Vyas, Piyush Vyas, Prathamesh Muzumdar, Anitha Chennamaneni, Anand Rajavat, and Romil Rawat. Extracting and analyzing factors to identify the malicious conversational ai bots on twitter. *Conversational Artificial Intelligence*, pages 71–83, 2024.

[62] Tianyi Wang, Xin Liao, Kam Pui Chow, Xiaodong Lin, and Yinglong Wang. Deepfake detection: A comprehensive survey from the reliability perspective. *ACM Computing Surveys*, 57(3):1–35, 2024.

[63] Esther Nanjala Wekesa, Casimer DeCusatis, and Andy Zhu. A black box comparison of machine learning reverse image search for cybersecurity osint applications. *Electronics*, 12(23):4822, 2023.

[64] Ashok Yadav, Atul Kumar, and Vrijendra Singh. Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber security. *Artificial Intelligence Review*, 56(11):12407–12438, 2023.

[65] Muhammad Mudassar Yamin, Mohib Ullah, Habib Ullah, Basel Katt, Mohammad Hijji, and Khan Muhammad. Mapping tools for open source intelligence with cyber kill chain for adversarial aware security. *Mathematics*, 10(12), 2022.

[66] Sai Sukeerthi Yarlagaada and Manas Kumar Yogi. Building a comprehensive cyber defense: The synergy of osint and traditional cybersecurity measures. *Journal of Hacking Techniques, Digital Crime Prevention and Computer Virology*, 1(2):13–23, 2024.

[67] Yangzong Zhang, Wenjian Liu, Kaiian Kuok, and Ngai Cheong. Anteater: Advanced persistent threat detection with program network traffic behavior. *IEEE Access*, 2024.

[68] Lijie Zhou, Yiran Qin, Shoujun Yan, Guoqing Zhang, and Wenjing Hu. Application analysis of generative artificial intelligence in the field of open source intelligence. In *2024 39th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 1603–1608. IEEE, 2024.

Sabina Szymoniak received an M.Sc. degree in computer science from the Czestochowa University of Technology, Czestochowa, Poland, and a PhD degree in computer science from the Faculty of Mechanical Engineering and Computer Science, Czestochowa University of Technology. She is currently an Assistant Professor at the Czestochowa University of Technology. She has authored or co-authored 50 research articles on verifying security protocols and cryptography.

Kacper Foks received an M.Sc in computer science from Czestochowa University of Technology. He is currently an IT Security Specialist at ZF Automotive Systems Poland. Kacper Foks deals with the security of computer systems, especially with the security of web applications. He has co-authored four research articles on open-source intelligence themes.

Aleksandra Pyrkosz-Dziubczyk received an M.Sc in computer science from Czestochowa University of Technology in 2023. She is currently an IT Security Specialist at ZF Automotive Systems Poland. Aleksandra Pyrkosz-Dziubczyk deals with the security of computer systems, especially with the security of web applications. He has co-authored three research articles on the security theme.