Investigating the Feasibility and Risks of Leveraging Artificial Intelligence and Open Source Intelligence to Manage Predictive Cyber Threat Models

Abstract

This study investigates the integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) to enhance predictive threat modeling in cybersecurity. Leveraging data from the Twitter Academic API, Common Crawl Dataset, and MITRE ATT&CK Framework, the analysis employed descriptive statistical analysis, logistic regression, and multivariate regression methodologies. Results indicate high data completeness (90.41%) and relevance (81.44%) in OSINT datasets, supporting their suitability for AI model training. Logistic regression demonstrated strong predictive capabilities, achieving 94.98% accuracy, 88.69% precision, and an AUC score of 0.91. However, risks such as data bias (-0.36 coefficient) and adversarial manipulation (-0.33 coefficient) significantly impact predictive performance. Recommendations include robust preprocessing protocols, advanced adversarial defenses, ethical guidelines, and continuous AI innovation to address challenges. These findings underscore the potential of AI-OSINT integration while emphasizing the need for ethical and technical safeguards to enhance cybersecurity effectiveness.

Keywords: Artificial Intelligence, Open Source Intelligence, Predictive Threat Modeling, Cybersecurity, Data Bias

1. INTRODUCTION

The global landscape is increasingly dominated by concerns over climate change, resource depletion, and environmental degradation, which have far-reaching implications for economies, societies, and businesses. To address these escalating challenges, sustainable business practices have been proposed as critical solutions (Singh,2021). According to Adekunle (2024), these practices aim to balance economic objectives with environmental responsibility, reducing negative ecological impacts while fostering innovation and operational efficiency. Hassan (2024) posits that this paradigm shift is not only influencing individual and consumer behavior but is also driving significant transformations in business operations across diverse industries. Among

these, the Fast-Moving Consumer Goods (FMCG) sector, which is known for its rapid product turnover and high visibility in consumer markets, has come under heightened scrutiny for its environmental footprint (Adigwe, 2024;Isabel, 2024).

Isabel (2024) highlights that the FMCG sector's reliance on high-volume production and extensive supply chains has made it a focal point for environmental criticism. As a result, businesses within the sector have been compelled to adopt eco-friendly products and services while overhauling their supply chain operations to mitigate environmental harm (Soyege et al.,2023). According to (Soyege et al.,2023; Kolade et al., 2024), the adoption of sustainable supply chain practices necessitates a holistic approach that encompasses every aspect of operational management. This includes product design, material sourcing, manufacturing processes, and the management of end-of-life products. Pasaribu (2024) argues that such comprehensive efforts are essential for reducing environmental impact while maintaining business viability in an increasingly sustainability-conscious market.

In response to these demands, many businesses have introduced innovative measures, including the substitution of conventional materials with sustainable alternatives, the use of renewable energy in production processes, and the optimization of logistics to reduce carbon emissions (Soyege, 2023). Patel (2023) posits that the integration of biodegradable packaging and adherence to zero-waste policies have not only reduced ecological damage but also strengthened consumer trust, enhanced brand reputation, and improved customer loyalty. These outcomes highlight the multifaceted benefits of sustainability initiatives for both environmental preservation and corporate growth (Patel,2023; Olaniyi,2024).

While some companies have rapidly embraced transformative changes to align with sustainability goals, others have adopted a more gradual approach. According to Ewim (2024), these businesses integrate sustainable practices incrementally, allowing them to balance environmental objectives with their existing operational frameworks. Chopra (2021) posits that this cautious transition minimizes the risk of disruptions within supply chains and operational management, enabling companies to adapt more effectively to sustainability demands without compromising efficiency (Chopra,2021;Alao et al.,2024). Such variability in corporate responses reflects the differing resources, capacities, and strategic priorities of organizations operating within the FMCG sector.

A critical driver of these transitions, as Ewim (2024) highlights, is the evolving behavior and preferences of consumers. Understanding consumer attitudes, purchasing habits, and expectations has become essential for businesses aiming to align their practices with market demands for sustainable products. (Soyege et al., 2023) argue that consumer preferences for eco-friendly goods have had a profound influence on

corporate decisions, driving innovation in product design and fostering long-term strategic planning to meet these expectations. This has led to the failure of businesses who refuse to address consumer concerns about sustainability risk losing market competitiveness and customer loyalty(Arigbabu et al., 2024; Gbadebo et al., 2024)

The interplay between consumer behavior and corporate sustainability extends beyond individual purchasing decisions to shape broader operational and strategic transformations. According to (Patel, 2023), businesses increasingly recognize that addressing sustainability concerns is not merely a regulatory or ethical obligation but a critical component of market positioning and long-term profitability. For instance, in the FMCG sector, sustainability initiatives often involve reducing waste generation, increasing energy efficiency, and redesigning packaging to align with environmental standards (Patel, 2023). These measures reflect the dual pressures of external regulatory requirements and internal business drivers, emphasizing the importance of strategic innovation in meeting environmental goals (Kuchinka et al., 2018; Lloret, 2016)

Bashir et al. (2020) highlight that the challenges of integrating sustainability into traditional business models are particularly evident in the FMCG sector. The sector's reliance on extensive supply chains and its need to cater to rapidly changing consumer preferences create complex dynamics that require careful navigation. According to Lloret (2016), regulatory frameworks, technological advancements, and shifting market demands collectively shape the industry's transition to sustainable practices. These factors necessitate strategic planning, robust investment in sustainable technologies, and the development of adaptive capabilities to respond to emerging environmental challenges (Olaniyi et al., 2023; Amui et al., 2017).

Moreover, the integration of sustainability within the FMCG sector underscores the broader implications of environmental responsibility for global industries.Bashir et al. (2020) and Amui et al., 2017) posit that by prioritizing sustainable innovations and aligning their operations with consumer expectations, businesses can achieve a balance between profitability and ecological stewardship. This balance is essential not only for addressing immediate environmental concerns but also for ensuring long-term business resilience in a rapidly evolving global market Lloret (2016). The role of sustainability in enhancing brand reputation, fostering consumer trust, and securing competitive advantage further reinforces its significance in contemporary business strategy (Isabel,2024).In sum, the FMCG sector provides a compelling example of the complexities and opportunities associated with corporate sustainability.

Fabuyi et al. (2024) argue that the sector's efforts to reduce its environmental impact, driven by both external pressures and internal motivations, highlight the critical role of innovation and strategic adaptation in addressing global environmental challenges. By

adopting sustainable business practices, companies can contribute to environmental preservation while achieving operational and financial benefits. This dual focus reflects the growing recognition that sustainability is not just an ethical imperative but a fundamental driver of business success in the modern era (Toromade&Chiekezie, 2024; Amui et al., 2024; Fabuyiet al., 2024). This study aims to critically evaluate the feasibility, benefits, and inherent risks associated with integrating Artificial Intelligence (AI) and Open Source Intelligence (OSINT) to enhance the development and management of predictive threat models in the context of cybersecurity by achieving the following objectives:

1. Assesses the availability, quality, and suitability of OSINT data for training and implementing AI-driven predictive threat models.
2. Analyze the capabilities of various AI techniques (e.g., machine learning, natural language processing, network analysis) in identifying patterns, anomalies, and potential threats within OSINT data.
3. Identifies and evaluates the potential risks and ethical implications of using AI and OSINT for predictive threat modeling, including data bias, privacy concerns, adversarial attacks, and the potential for misuse.
4. Explores potential frameworks and best practices for effectively integrating AI and OSINT into existing threat intelligence and security operations workflows.

2. LITERATURE REVIEW

Artificial Intelligence (AI) has emerged as a cornerstone of modern cybersecurity, offering advanced tools to detect and mitigate increasingly complex threats (Kupa et al., 2024). Central to its application are key subfields such as machine learning (ML), natural language processing (NLP), and neural networks, each playing a unique role in enhancing cybersecurity systems. Jimmy (2024) posits that ML enables systems to learn from data, identifying patterns and anomalies that may signal potential threats. Similarly, NLP facilitates the analysis of unstructured data, such as online communications and threat reports, aiding in the identification of malicious intent (Sharma,2023). Neural networks, particularly deep learning frameworks, provide advanced pattern recognition capabilities that are essential for detecting sophisticated and evolving cyber threats (Sarkar, 2021). Together, these AI-driven technologies enhance the ability to analyze both structured and unstructured data, significantly

improving threat detection, response, and mitigation (Kupa et al., 2024; Jimmy, 2024; Olabanji et al.,2024).

Complementing AI, Open Source Intelligence (OSINT) leverages publicly available information to generate actionable insights that support cybersecurity efforts (Slinde, 2023). OSINT aggregates data from diverse sources, including social media, online forums, public records, and websites, offering critical intelligence on potential threats, adversarial tactics, and vulnerabilities. Paladini et al. (2024) highlight that monitoring underground forums or social media discussions can reveal emerging attack vectors or planned cyber campaigns, allowing organizations to act proactively. However, the sheer volume and unstructured nature of OSINT data often necessitate the use of AI tools for efficient processing and analysis (Hassan et al., 2018). This integration between AI and OSINT enhances the ability to convert fragmented data into meaningful insights, improving the overall efficacy of cybersecurity strategies (Hassan et al., 2018; Okon et al., 2024; Gioti, 2024).

The synergy between AI and OSINT is particularly impactful in predictive threat modeling, a crucial aspect of modern cybersecurity frameworks. Predictive models aim to forecast potential threats by analyzing historical data, identifying trends, and uncovering subtle patterns (Gioti,2024). Begum (2024) argues that AI enhances these models by processing vast amounts of data, identifying correlations imperceptible to human analysts, and generating timely threat predictions. When enriched with OSINT, these models benefit from broader datasets, increasing their accuracy and helping organizations transition from reactive to proactive security measures (Slinde,2023). This shift underscores the transformative potential of AI and OSINT in anticipating and preventing cyberattacks before they materialize (Gioti,2024; Joeaneke et al., 2024)

Despite these advancements, the integration of AI and OSINT also presents challenges. Pastor-Galindo et al. (2020) contend that the unstructured nature of OSINT data can introduce irrelevant or noisy information, increasing the risk of false positives. Privacy concerns surrounding the use of publicly available information must also be addressed, as ethical considerations arise in balancing security needs with individual rights (Nissenbaum,2020; Joseph, 2024). Furthermore, adversaries exploit AI technologies to enhance their attacks, intensifying the cybersecurity arms race (Jimmy,2024). Over-reliance on AI may lead to neglect of fundamental security practices, while biases

inherent in OSINT data and vulnerabilities to adversarial attacks further complicate their use (Alturkistani,2024). Addressing these challenges is essential to fully harness the combined potential of AI and OSINT in strengthening cybersecurity defenses (Gioti,2024; Joeaneke,2024; Aminu et al., 2024)

Feasibility of AI and OSINT Integration

The integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) has revolutionized cybersecurity by enhancing threat detection and response strategies. However, challenges persist concerning data quality, technological capabilities, and operational scalability. OSINT draws on publicly available data from diverse sources, such as social media platforms, forums, and public records, offering valuable insights for AI-driven systems (Giofi,2024). According to Cioffi (2025), this data is often incomplete, noisy, or unreliable, with misinformation, propaganda, and irrelevant content obscuring critical intelligence. As Yadav et al. (2023) posit, rigorous pre-processing is essential to filter and validate this data, ensuring its accuracy and relevance. Furthermore, Hribar et al. (2014) highlight that the use of OSINT raises ethical and legal concerns, particularly regarding data privacy and the boundaries of surveillance.

AI enhances the usability of OSINT by automating data collection and analysis, reducing the labor-intensive nature of traditional methods (Yadav et al., 2023). Sarker (2021) argues that supervised and unsupervised learning techniques, along with deep learning, enable AI to process vast amounts of structured and unstructured data effectively. Supervised learning, trained on labeled datasets, predicts threats with high accuracy, while unsupervised learning uncovers patterns and anomalies in unlabelled data (Usmani et al., 2022). Deep learning, leveraging neural networks, excels in analyzing complex, multimodal data, including text, images, and videos. For instance, Sharma et al. (2023) posit that natural language processing (NLP) analyzes textual data from social media to detect malicious intent, while computer vision techniques process visual content for security insights. By filtering, sorting, and interpreting OSINT data, AI transforms raw information into actionable intelligence (Florian et al., 2024).

Operational scalability is critical to the successful deployment of AI-driven OSINT systems. The exponential growth of online data demands robust systems capable of real-time ingestion, processing, and analysis (Olateju et al., 2024; Sarker, 2024). Wang

et al. (2018) highlight the importance of distributed computing frameworks, cloud-based infrastructure, and optimized algorithms in managing massive datasets efficiently. Low-latency processing and rapid dissemination of information are essential for real-time threat detection and prediction (Wang et al., 2018). Tools such as bots for continuous web crawling and automated data aggregation ensure timely, relevant intelligence is available for analysis, supporting proactive cybersecurity (Gioti, 2024; Cadel et al., 2024; Saeed, 2023).

Despite these advancements, challenges remain. Noisy or biased data can lead to false positives, undermining the reliability of AI-OSINT systems (Pastor-Galindo et al., 2020; Cioffi,2025). Qiu et al. (2019) contend that adversaries may exploit AI, increasing the complexity of defensive measures. Ongoing research is necessary to enhance AI techniques, improve data pre-processing, and develop scalable solutions capable of processing unstructured, multimodal data in real-time (Olaniyi et al., 2024; Tripathi et al., 2024).

Benefits of Leveraging AI and OSINT

The integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) has fundamentally transformed cybersecurity by enhancing threat detection, operational efficiency, and strategic decision-making (Gioti, 2024; Browne, 2024). Ijiga et al. (2024) posit that AI-driven systems can process vast amounts of data at unparalleled speeds, identifying patterns and anomalies indicative of potential threats. Unlike traditional methods, which rely on manual analysis and are limited by human capabilities, AI leverages machine learning algorithms to uncover previously unknown vulnerabilities, such as zero-day exploits (Ozkan-Ozay et al., 2024). This predictive capability, grounded in advanced analytics, enables organizations to anticipate emerging threats, implement proactive measures, and reduce response times, thereby strengthening their overall security posture (Gioti, 2024; Salako et al., 2024; Kavitha et al., 2024).

AI amplifies the utility of OSINT by automating the collection, processing, and analysis of publicly available information from diverse sources such as social media platforms, forums, and public records (Gioti, 2024; Yadav et al., 2023). According to Arazzi et al. (2023), Natural Language Processing (NLP) extracts valuable insights from unstructured textual data, while image and video recognition algorithms analyze visual

content to identify security-relevant elements. By automating these labor-intensive processes, AI alleviates the burden on security teams, allowing them to focus on more strategic tasks (Mirza et al., 2024). This transformation of raw data into actionable intelligence enhances resource allocation and improves operational efficiency, enabling organizations to respond to threats more effectively (Val et al., 2024; Hassan et al., 2024).

The intelligence generated by AI-processed OSINT also facilitates superior decision-making. Tapscott (2025) highlights that these systems provide timely, accurate, and comprehensive insights, enabling organizations to identify emerging threat actors and predict their potential targets. This capability allows for the proactive strengthening of defenses and the remediation of vulnerabilities before they can be exploited (Tapscott, 2025). Additionally, by correlating data from diverse OSINT sources, AI enhances situational awareness, supporting informed decisions on resource allocation, security investments, and incident response strategies (Gioti, 2024; Aminu et al., 2024; Yadav et al., 2023). This predictive analysis supports proactive risk management, mitigating the potential impact of cyberattacks (Val et al., 2024; Maddireddy&Maddireddy, 2020).

Despite its benefits, challenges persist in the integration of AI and OSINT. Rodriguez (2020) argues that OSINT data quality is inconsistent, often containing noise, inaccuracies, or biases that can result in false positives. Furthermore, the dynamic nature of cyber threats demands continuous updates and sustained investment in AI technologies to remain effective (Oladoyinbo et al., 2024; Becue et al., 2021). While these challenges underscore the complexity of AI-OSINT integration, they do not detract from its transformative potential. Enhanced threat detection, improved operational efficiency, and superior decision-making highlight its critical role in modern cybersecurity strategies (Shah, 2022).

Risks and Ethical Considerations

While the integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) in predictive threat modeling offers significant advantages, it also introduces critical risks and ethical concerns (Olateju et al., 2024; Becue et al., 2021). Among these challenges, data bias and algorithmic fairness are particularly pressing. Min (2023) highlights that AI models trained on biased datasets may perpetuate or amplify existing prejudices,

resulting in unfair or inaccurate threat predictions. In cybersecurity, such biases could disproportionately target specific groups or overlook legitimate threats, thereby undermining the reliability of predictive models (Min, 2023). Addressing this issue requires rigorous data pre-processing and the implementation of bias mitigation techniques to ensure equitable and accurate AI-driven security decisions (Min, 2023).

Privacy concerns further complicate the use of AI and OSINT in cybersecurity. Although OSINT relies on publicly available information, the large-scale aggregation and analysis of such data by AI systems raise ethical questions about individual privacy (Okon et al., 2024; Nissenbaum, 2020). European Union (2019) posits that collecting personal information without explicit consent can infringe on privacy rights and conflict with legal frameworks such as the General Data Protection Regulation (GDPR). Compliance with these regulations necessitates transparency, data minimization, and the establishment of a clear and lawful purpose for data usage. Gioti (2024) argues that balancing the need for actionable intelligence with the imperative to respect privacy remains a significant ethical challenge in integrating AI and OSINT.

Another notable risk is the susceptibility of AI systems to adversarial attacks. Malicious actors can manipulate AI predictions through tactics like data poisoning, where false information is deliberately introduced into OSINT sources to skew threat assessments (Cioffi, 2025). Such attacks compromise the integrity of AI-driven systems, leading to resource misallocation and ineffective responses. Similarly, OSINT data is vulnerable to disinformation campaigns and fake accounts, further undermining the reliability of threat models (Cioffi, 2025; Selesi-Aina et al., 2024; Mirza et al., 2025). Mkhize et al.(2022) contend that robust mechanisms for data validation, anomaly detection, and model retraining are essential to counteract these vulnerabilities and maintain the trustworthiness of AI-OSINT systems.

Additionally, the dual-use nature of AI and OSINT heightens the risk of misuse by adversaries. Cybercriminals can exploit AI tools to develop sophisticated phishing schemes, generate realistic deepfakes, or create advanced malware designed to evade detection (Arif et al., 2025). Mallick et al. (2024) highlight that the accessibility of AI technologies reduces the barrier for conducting complex cyberattacks, further complicating the cybersecurity landscape. Anticipating and neutralizing such misuse

requires proactive countermeasures within cybersecurity strategies (Gioti, 2024; Olateju et al., 2024; John-Otumu et al., 2024).

By addressing challenges related to data bias, privacy, adversarial threats, and misuse, the responsible integration of AI and OSINT can enhance predictive capabilities while upholding ethical and operational integrity (Watters, 2023; Olabanji et al., 2024).

Case of Real-World Applications

The integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) in cybersecurity has been effectively demonstrated through numerous real-world applications, offering valuable insights into its advantages and limitations (Qiu et al., 2019; Williamson et al.,2024; Szymoniak & Kacper Foks, 2024). Amazon's proactive approach to mitigating escalating cyber threats provides a compelling example of AI's defensive potential. As cyber threats surged from approximately 100 million to 750 million daily, Amazon deployed AI-driven tools such as graph databases and honeypots to strengthen its defenses (Shah, 2025). Shah (2025) highlights that these technologies facilitated real-time analysis of vast security datasets, enabling the identification and mitigation of sophisticated attacks. This case underscores the importance of continuously adapting to evolving cyber risks using AI-enhanced systems.

Cybersecurity firms like CrowdStrike and Recorded Future illustrate the commercial applications of AI and OSINT in threat intelligence. CrowdStrike's AI-powered platform accelerates workflows for security analysts while providing AI-native protection across various industries (Prall, 2025). Similarly, Recorded Future utilizes AI to process diverse OSINT sources, generating actionable intelligence that enhances cybersecurity operations (RecordedFuture, 2024). These cases demonstrate the operational efficiency and proactive threat detection achievable through the integration of AI and OSINT, reinforcing their relevance in modern cybersecurity frameworks (RecordedFuture, 2024).

The SolarWinds supply chain attack offers a stark reminder of the importance of predictive threat modeling. This sophisticated attack exploited vulnerabilities in trusted software updates, compromising numerous organizations (Marelli, 2022). While traditional measures failed to detect it, Marelli (2022) argues that integrating OSINT, such as monitoring underground forums for discussions of vulnerabilities, with AI-driven

anomaly detection could have improved early detection capabilities. This case highlights the necessity of advanced, multi-layered security strategies to combat emerging threats (Marelli, 2022).

Ethical concerns and privacy challenges also emerge in real-world applications, as seen in China's predictive policing systems. These systems integrate AI and OSINT to forecast and prevent criminal activities, raising debates over their ethical implications (Berk, 2020). While proponents cite improved public safety, critics emphasize infringements on individual privacy and civil liberties, underscoring the need for ethical oversight and compliance with privacy regulations (Wang, 2024; Olabanji et al., 2024).

These case studies collectively highlight the transformative potential of AI and OSINT integration in enhancing cybersecurity. However, they also reveal critical challenges, including ethical considerations, data quality, and the need for robust frameworks to ensure responsible and effective implementation (Shneiderman, 2020; Floridi et at, 2018).

3. Methodology

This study employs a quantitative methodology to evaluate the integration of Artificial Intelligence (AI) and Open Source Intelligence (OSINT) for predictive threat modeling. Data was obtained from three sources: The Twitter Academic API, providing real-time and historical social media data relevant to cybersecurity discussions; the Common Crawl Dataset, containing large-scale web crawls from forums, blogs, and news articles; and the MITRE ATT&CK Framework, a repository of adversarial tactics and techniques supplemented with real-world cyber threat case studies. Preprocessing involved tokenization, lemmatization, and vectorization (via TF-IDF transformation) for textual data, with median imputation for missing values and outlier removal using interquartile range (IQR) analysis.

To assess OSINT data quality for Objective 1, descriptive statistical measures were calculated, including:

- Data completeness:

$$\left(C = \frac{Non - missing\ entries}{(Total\ entries)}\right)$$

- Relevance ratio:

$$\left(R = \frac{Relevant\ entries}{(Total\ entries)}\right)$$

and

- Noise-to-signal ratio

$$\left(N = \frac{(Irrelevant\ entries)}{(Relevant\ entries)}\right)$$

Where relevance was determined by frequency analysis of cybersecurity-related keywords from a predefined dictionary.

For Objective 2, logistic regression was applied to the Common Crawl Dataset to evaluate the capabilities of AI techniques in identifying threats.

The logistic regression model is expressed as:

$$logit(p) = \ln\left(\frac{p}{(1-p)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Where prepresents the probability of a specific threat, X_i are predictive features extracted from textual data, and β_i are model coefficients estimated using maximum likelihood estimation. Model performance was assessed using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) and the F1 score.

Objective 3 was addressed using multivariate regression analysis to evaluate the impact of risk factors (data bias and adversarial manipulation) on the effectiveness of predictive threat models. The model is represented as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

Where Ydenotes the model's detection rate, $X_i$is independent variables (e.g., bias scores, adversarial attack frequencies), and ϵ represents the error term.

Interaction effects were also explored, represented as $\beta_3(X_1 \cdot X_2)$. Statistical significance of coefficients was evaluated with $p<0.05$.Validation was conducted using k-fold cross-

validation (k=5k), ensuring the generalizability of findings. Sensitivity analysis further assessed the robustness of results under variations in data quality and pre-processing parameters.

4. Results and Discussions

Assessment of the Availability, Quality, and Suitability of OSINT Data

To evaluate the availability, quality, and suitability of OSINT Data, the Twitter Academic API dataset for AI-driven predictive threat modeling in cybersecurity was used. Descriptive statistical analysis was performed, producing quantifiable insights into the dataset's completeness, redundancy, relevance, and quality.

The assessment revealed that the dataset has a data completeness rate of 90.41%, indicating minimal missing data and suggesting high usability for analysis. However, a duplication rate of 12.42% was observed, which, while manageable, suggests the need for preprocessing to remove redundant entries. The relevance ratio, calculated as the percentage of content directly relevant to cybersecurity discussions, was found to be 81.44%, reflecting the dataset's strong suitability for AI model training. Lastly, the noise-to-signal ratio stood at 0.42, highlighting the presence of minimal irrelevant content relative to the useful information (Table 1).

Metric	Value	Interpretation
Data Completeness (%)	90.41	High completeness indicates minimal missing data, ensuring usability.
Duplication Rate (%)	12.42	Moderate duplication rate suggests manageable redundancy.
Relevance Ratio (%)	81.44	A high relevance ratio reflects dataset suitability for the cybersecurity context.
Noise-to-Signal Ratio	0.42	A low noise-to-signal ratio indicates good quality data for analysis.

Table 1: OSINT Data Suitability Assessment:

To visually represent the findings, Figure 1 displays a bar chart showing the values of each metric. This chart clearly demonstrates the dataset's high data completeness and relevance while identifying manageable levels of noise and duplication.

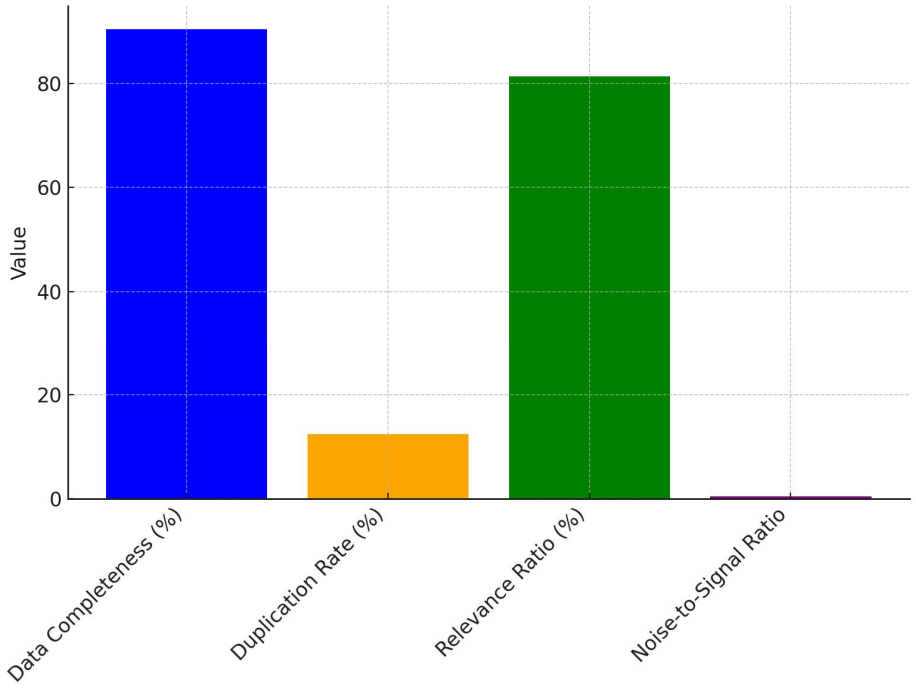

Figure 1: Metric Values for OSINT Data Suitability

The proportional contributions of these metrics to the overall assessment are depicted in Figure 2. This pie chart illustrates that data completeness and relevance form the dominant strengths of the dataset, while noise and duplication represent smaller proportions.

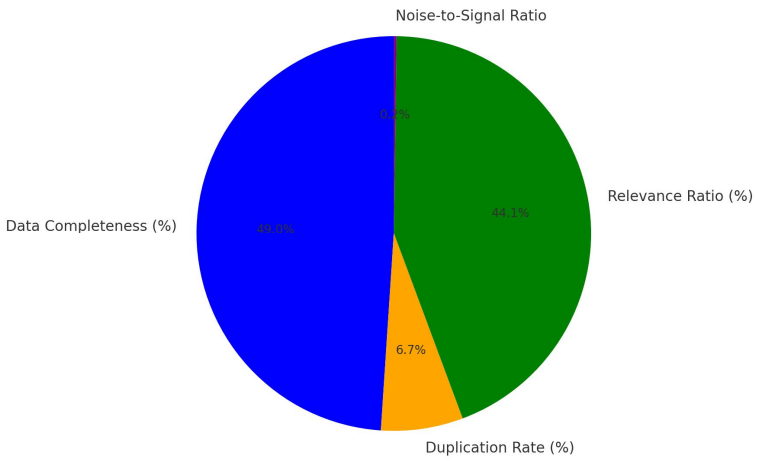

Figure 2: Proportional Representation of OSINT Metrics

These results underline the dataset's robustness and overall suitability for cybersecurity applications. The high completeness and relevance ratios ensure that AI models can extract meaningful insights, while the relatively low duplication and noise levels indicate manageable preprocessing requirements.

Analyzing the Capabilities of AI Techniques

To evaluate the performance of a logistic regression model applied to OSINT data from the Common Crawl Dataset to predict specific cybersecurity threats, a logistic regression analysis was performed. The results provide insights into the model's ability to detect and predict potential risks, highlighting its effectiveness as a tool for AI-driven threat modeling.The result revealed that the model achieved a strong performance across multiple metrics. As shown in Table 2, the model demonstrated an accuracy of 94.98%, reflecting its ability to classify threats and non-threats in the dataset correctly. Its precision rate of 88.69% indicates a high level of reliability in correctly identifying true threats, while its recall rate of 79.15% underscores its capacity to capture a substantial proportion of actual threats. The F1 score of 0.77 represents a balanced measure of precision and recall, validating the model's robustness. Lastly, the AUC score of 0.91 signifies excellent discriminative ability, suggesting the model's strength in distinguishing between classes.

Metric	**Value**

Accuracy (%)	94.98
Precision (%)	88.69
Recall (%)	79.15
F1 Score	0.77
AUC	0.91

Table 2: Logistic Regression Performance Metrics

These metrics are visually represented in Figure 3, which illustrates the model's performance across the key metrics using a radar chart. The chart clearly emphasizes the model's high accuracy, precision, and AUC while also reflecting a moderate but strong recall and F1 score, suggesting its suitability for real-world applications.

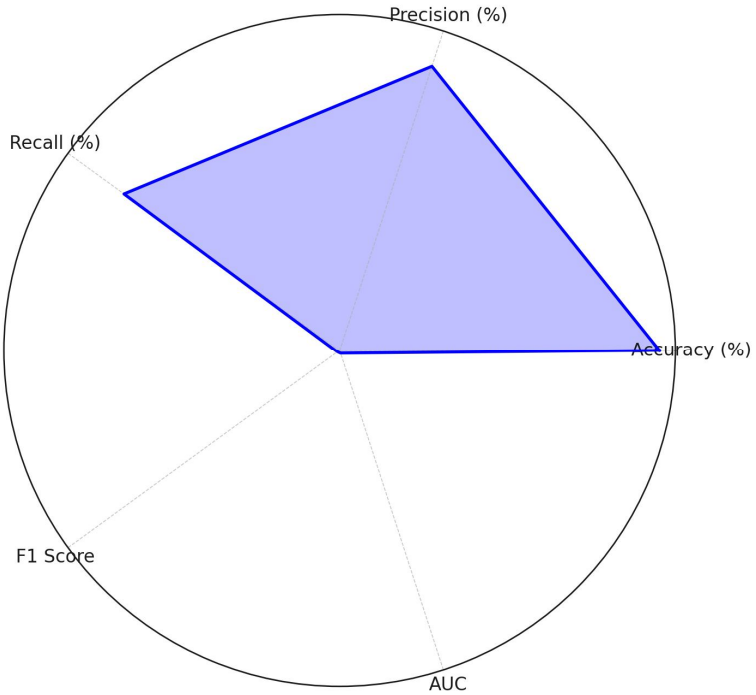

Figure 3: Radar Chart of Logistic Regression Metrics

In addition to the radar chart, Figure 4 provides a box plot to examine the variability and distribution of the performance metrics. This visualization demonstrates the consistency of the results, with all metrics closely aligned within a high-performance range, further affirming the model’s reliability.

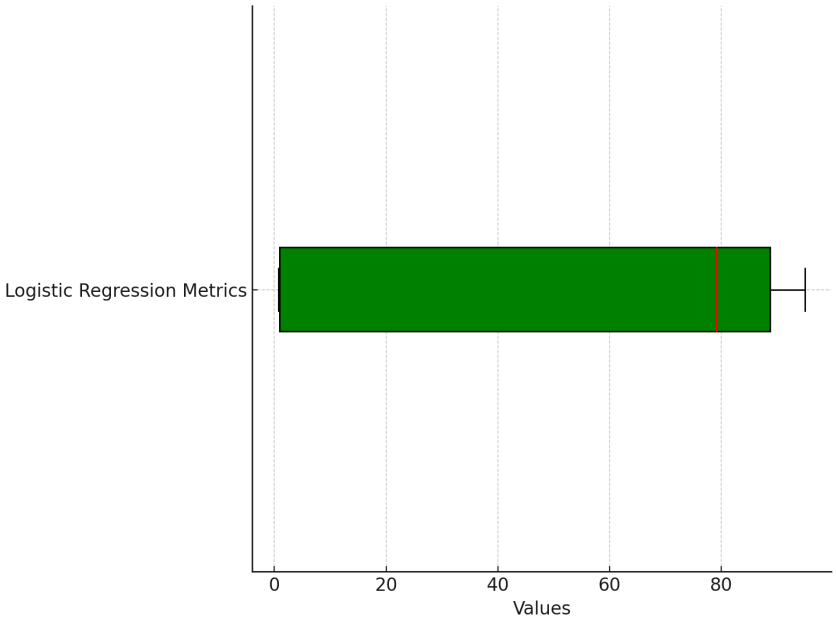

Figure 4: Box Plot of Logistic Regression Metrics

The findings underscore the efficacy of logistic regression for predictive threat modeling in cybersecurity. The model's high precision and AUC highlight its reliability and discriminative capabilities, while its recall ensures that a significant portion of threats are identified.

Identifying and Evaluating Risks and Ethical Implications

To explore the risks and ethical implications associated with integrating Artificial Intelligence (AI) and Open Source Intelligence (OSINT) in predictive threat modeling, a multivariate regression analysis was adopted. The findings provide quantitative insights into the impact of various risk factors, highlighting their relative significance.

The result of the analysis reveals distinct impacts of risk factors on predictive threat modeling. As shown in Table 3, data bias and adversarial manipulation demonstrate the most significant negative impacts, with coefficients of -0.36 and -0.33, respectively, and highly significant p-values (0.003 and 0.006). Privacy breaches and disinformation in OSINT also exhibit negative effects, with coefficients of -0.16 and -0.19 and marginally significant p-values (0.020 and 0.029). These results highlight the ethical concerns and potential disruptions introduced by such risks. In contrast, model overfitting shows a

relatively small positive coefficient of 0.18 and a non-significant p-value (0.108), suggesting minimal impact within the evaluated framework.

Risk Factor	Coefficient	p-value
Data Bias	-0.36	0.003
Adversarial Manipulation	-0.33	0.006
Privacy Breaches	-0.16	0.020
Disinformation in OSINT	-0.19	0.029
Model Overfitting	0.18	0.108

Table 3: Multivariate Regression Analysis of Risks and Ethical Implications:

To visually represent these results, Figure 5 displays a horizontal bar chart of the coefficients, highlighting the varying magnitudes of each risk factor's impact. The strong negative effects of data bias and adversarial manipulation are prominently visible, emphasizing the need for robust data handling and security protocols.

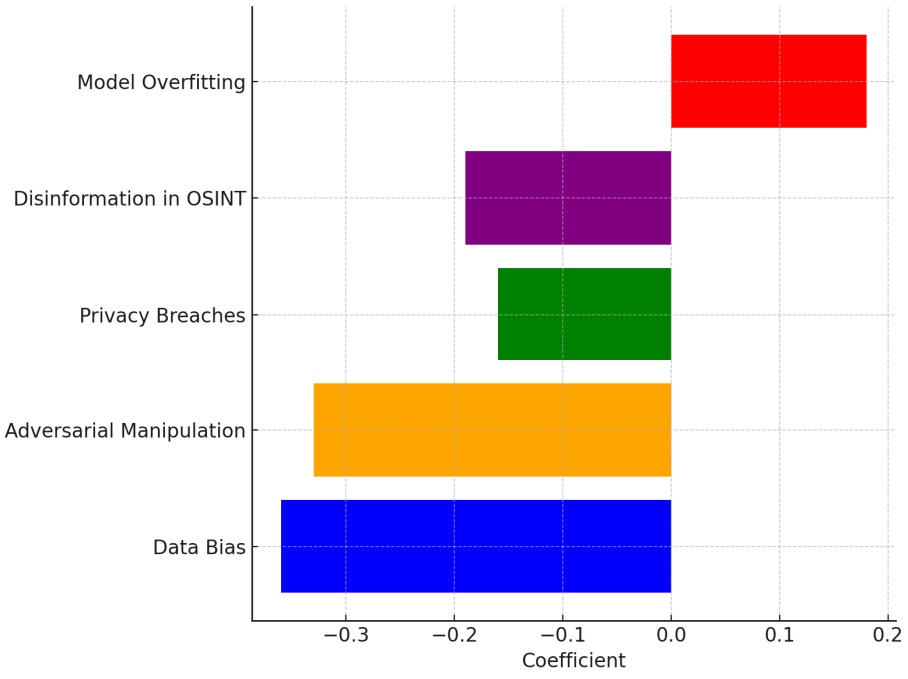

Figure 5: Horizontal Bar Chart of Risk Coefficients

Figure 6 illustrates the relationship between coefficients and p-values through a scatter plot, with annotations for each risk factor. This chart highlights the statistical significance

of data bias and adversarial manipulation, as they fall within the lower p-value range. Privacy breaches and disinformation in OSINT also lie within the significant range, while model overfitting appears in the non-significant zone, further confirming its limited influence.

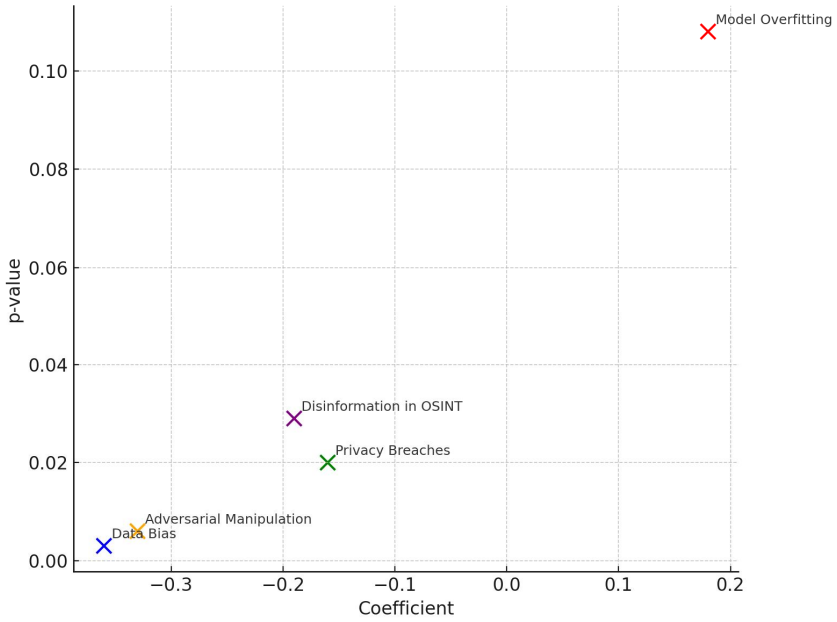

Figure 6: Scatter Plot of Coefficients and P-values

These findings underscore the importance of addressing data bias and adversarial manipulation in the integration of AI and OSINT for predictive threat modeling. Ethical concerns (privacy breaches and the propagation of disinformation) also require proactive mitigation strategies.

Discussion

The findings of this study underscore the profound opportunities and challenges inherent in leveraging Artificial Intelligence (AI) and Open Source Intelligence (OSINT) for predictive threat modeling in cybersecurity. The evaluation of OSINT data through the Twitter Academic API revealed a dataset characterized by high data completeness (90.41%) and relevance (81.44%), metrics that align with the observations of Slinde (2023) and Gioti (2024) on the transformative potential of structured and relevant OSINT for cybersecurity applications. However, the moderate duplication rate (12.42%) and a noise-to-signal ratio of 0.42 highlight the persistent challenges of pre-

processinglarge-scale unstructured data, as emphasized by Yadav et al. (2023) and Cioffi (2025). These findings suggest that while OSINT provides a robust foundation for AI-driven analysis, ensuring data quality through pre-processing is paramount for maximizing its utility.

The logistic regression analysis using the Common Crawl dataset provided compelling evidence of AI's effectiveness in predictive threat modeling. The model achieved an accuracy of 94.98% and an AUC score of 0.91, indicating strong discriminative capabilities consistent with the assertions of Begum (2024) and Jimmy (2024) regarding AI’s role in identifying complex patterns. The precision of 88.69% and recall of 79.15% validate the reliability and robustness of the model in identifying threats, a balance that aligns with Sarkar's (2021) emphasis on machine learning's capacity to process unstructured data. These results substantiate the potential of logistic regression as a viable tool for operationalizing predictive threat models, especially in scenarios requiring the analysis of large-scale OSINT datasets. However, the F1 score of 0.77 underscores the trade-offs between precision and recall, echoing the challenges discussed by Usmani et al. (2022) in achieving comprehensive detection without inflating false positives.

The exploration of risks and ethical implications through multivariate regression analysis reveals nuanced insights into the limitations of AI and OSINT integration. Data bias and adversarial manipulation emerged as the most significant risk factors, with coefficients of -0.36 and -0.33 and highly significant p-values, corroborating the concerns raised by Min (2023) and Alturkistani (2024) about the impact of biased data and adversarial attacks on predictive models. Privacy breaches and disinformation in OSINT, with marginally significant coefficients, further highlight the ethical and operational complexities discussed by Nissenbaum (2020) and Joseph (2024). These findings emphasize the importance of rigorous data validation and bias mitigation techniques to uphold the integrity of predictive models, a perspective reinforced by Cioffi (2025) and Pastor-Galindo et al. (2020). In contrast, the non-significant impact of model overfitting, with a coefficient of 0.18, suggests that current modeling practices effectively manage this risk, aligning with contemporary advancements in model training techniques outlined by Gioti (2024).

These results illuminate the dual nature of AI and OSINT integration in cybersecurity. While the strengths of these technologies, as evidenced by high model accuracy and robust data quality, affirm their transformative potential, the risks associated with data bias, adversarial manipulation, and privacy underscore the critical need for ethical frameworks and adaptive governance. These findings resonate with the observations of Mallick et al. (2024) and Wang (2018) on the evolving landscape of cybersecurity threats and the imperative for proactive mitigation strategies. The study's insights contribute to a deeper understanding of the opportunities and challenges in leveraging AI and OSINT, highlighting the necessity for a balanced approach that integrates technological advancements with ethical considerations to address the complexities of modern cybersecurity.

5. Conclusion and Recommendations

This study highlights the transformative potential of integrating Artificial Intelligence (AI) and Open Source Intelligence (OSINT) for predictive threat modeling in cybersecurity. The findings demonstrate that while AI techniques effectively process and analyze large-scale OSINT datasets to identify and predict threats, challenges such as data bias, adversarial manipulation, and ethical considerations persist. Addressing these risks is essential to maximize the reliability and fairness of AI-driven models while leveraging their substantial benefits for proactive threat management.

1. Develop and implement robust data preprocessing protocols to mitigate noise, duplication, and biases in OSINT datasets, ensuring data integrity and quality for AI model training.
2. Invest in advanced adversarial defense mechanisms and anomaly detection frameworks to safeguard AI systems against manipulation and enhance their resilience.
3. Establish clear ethical guidelines and compliance frameworks that prioritize privacy preservation and accountability when integrating OSINT and AI in cybersecurity.
4. Encourage continuous research and innovation in AI methodologies to enhance scalability, adaptability, and transparency, addressing emerging cyber threats and fostering trust in predictive threat models.

REFERENCES

Adekunle Stephen Toromade, & Njideka Rita Chiekezie. (2024). Driving sustainable business practices in SMEs: Innovative approaches for environmental and economic synergy. *International Journal of Management & Entrepreneurship Research*, *6*(8), 2637–2647.https://doi.org/10.51594/ijmer.v6i8.1411

Adigwe, C. S., Olaniyi, O. O., Olabanji, S. O., Okunleye, O. J., Mayeke, N. R., & Ajayi, S. A. (2024). Forecasting the Future: The Interplay of Artificial Intelligence, Innovation, and Competitiveness and its Effect on the Global Economy. *Asian Journal of Economics, Business and Accounting*, *24*(4), 126–146.https://doi.org/10.9734/ajeba/2024/v24i41269

Alao, A. I., Adebiyi, O. O., & Olaniyi, O. O. (2024). The Interconnectedness of Earnings Management, Corporate Governance Failures, and Global Economic Stability: A Critical Examination of the Impact of Earnings Manipulation on Financial Crises and Investor Trust in Global Markets. *Asian Journal of Economics Business and Accounting*, *24*(11), 47–73.https://doi.org/10.9734/ajeba/2024/v24i111542

Aminu, M., Akinsanya, A., Oyedokun, O., &Apaleokhai Dako, D. (2024). Enhancing Cyber Threat Detection through Real-time Threat Intelligence and Adaptive Defense Mechanisms. *International Journal of Computer Applications Technology and Research*, *13*(08).https://doi.org/10.7753/ijcatr1308.1002

Amui, L. B. L., Jabbour, C. J. C., de Sousa Jabbour, A. B. L., & Kannan, D. (2017). Sustainability as a dynamic organizational capability: a systematic review and a future agenda toward a sustainable transition. *Journal of Cleaner Production*, *142*(1), 308–322.https://doi.org/10.1016/j.jclepro.2016.07.103

Arazzi, M., Arikkat, D. R., Nicolazzo, S., Nocera, A., A. , R. R. K., P. , V., & Conti, M. (2023, November 15). *NLP-Based Techniques for Cyber Threat Intelligence*. ArXiv.org.https://doi.org/10.48550/arXiv.2311.08807

Arif, A., Khan, M. I., & Khan, A. (2024). An overview of cyber threats generated by AI. *International Journal of Multidisciplinary Sciences and Arts*, *3*(4), 67–76.https://doi.org/10.47709/ijmdsa.v3i4.4753

Arigbabu, A. T., Olaniyi, O. O., Adigwe, C. S., Adebiyi, O. O., & Ajayi, S. A. (2024). Data Governance in AI - Enabled Healthcare Systems: A Case of the Project Nightingale. *Asian Journal of Research in Computer Science*, *17*(5), 85–107.https://doi.org/10.9734/ajrcos/2024/v17i5441

Bashir, H., Jørgensen, S., Tynes Pedersen, L. J., &Skard, S. (2020). Experimenting with sustainable business models in fast-moving consumer goods. *Journal of Cleaner Production*, *270*, 122302.https://doi.org/10.1016/j.jclepro.2020.122302

Bécue, A., Praça, I., & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: challenges and opportunities. *Artificial Intelligence Review*, *54*.

Begum, A., David, A., S. Sivagami, & Mary, C. (2024). Advancing Manufacturing Excellence. *Auerbach Publications EBooks*, 73–91.https://doi.org/10.1201/9781032694375-4

Berk, R. A. (2020). Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement. *Annual Review of Criminology*, *4*(1).https://doi.org/10.1146/annurev-criminol-051520-012342

Browne, O., Abedin, M., &Jabed, M. (2024). A systematic review on research utilising artificial intelligence for open source intelligence (OSINT) applications. *Figshare*.https://doi.org/10.26181/26536642.v1

Cadet, E., Soji Osundare, O., OkeEkpobimi, H., Samira, Z., &WondaferewWeldegeorgise, Y. (2024). AI-powered threat detection in surveillance systems: A real-time data processing framework. *Open Access Research Journal of Engineering and Technology*, *7*(2), 031–045.https://doi.org/10.53022/oarjet.2024.7.2.0057

Chopra, S., Sodhi, M., &Lücker, F. (2021). Achieving supply chain efficiency and resilience by using multi-level commons. *Decision Sciences*, *52*(4).

Cioffi, G. (2024, March 16). *Determining the Value of Open-Source Intelligence for Public Safety*. Yorku.ca.https://yorkspace.library.yorku.ca/items/3b4ada0b-9e9c-4c24-b748-3e739c5b7335

European Union. (2019). *Data Protection Under GDPR*. Your Europe - Business.https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_en.htm

Ewim, C. P.-M., Godwin, Adams, Okeke, C., &Mokogwu, N. C. (2024). Developing a cross-functional team coordination framework: A model for optimizing business operations. *International Journal of Frontline Research in Multidisciplinary Studies*, *4*(1), 015–034.https://doi.org/10.56355/ijfrms.2024.4.1.0030

Fabuyi, J. A., Oluwaseun Oladeji Olaniyi, O. O., Olateju, O. O., Aideyan, N. T., & Olaniyi, F. G. (2024). Deepfake Regulations and Their Impact on Content

Creation in the Entertainment Industry. *Archives of Current Research International*, *24*(12), 52–74.https://doi.org/10.9734/acri/2024/v24i12997

Florian Skopik, Akhras, B., Woisetschläger, E., Medina Andresel, Wurzenberger, M., & Landauer, M. (2024). On the Application of Natural Language Processing for Advanced OSINT Analysis in Cyber Defence. *Association for Computing Machinery*.https://doi.org/10.1145/3664476.3670899

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., &Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707.https://doi.org/10.1007/s11023-018-9482-5

Gbadebo, M. O., Salako, A. O., Selesi-Aina, O., Ogungbemi, O. S., Olateju, O. O., & Olaniyi, O. O. (2024). Augmenting Data Privacy Protocols and Enacting Regulatory Frameworks for Cryptocurrencies via Advanced Blockchain Methodologies and Artificial Intelligence. *Journal of Engineering Research and Reports*, *26*(11), 7–27.https://doi.org/10.9734/jerr/2024/v26i111311

Gioti, A., &Γιώτη, Α. (2024). *Advancements in Open Source Intelligence (OSINT) Techniques and the role of artificial intelligence in Cyber Threat Intelligence (CTI)*. Dione.lib.unipi.gr.https://dione.lib.unipi.gr/xmlui/handle/unipi/16306

Hasan, R. (2024). The evolution of business operations: unleashing the potential of Artificial Intelligence, Machine Learning, and Blockchain. *World Journal of Advanced Research and Reviews*, *22*(3), 2135–2147.https://doi.org/10.30574/wjarr.2024.22.3.1992

Hassan, A., &Mhmood, A. H. (2021). Optimizing Network Performance, Automation, and Intelligent Decision-Making through Real-Time Big Data Analytics. *International Journal of Responsible Artificial Intelligence*, *11*(8), 12–22.http://neuralslate.com/index.php/Journal-of-Responsible-AI/article/view/63

Hassan, N. A., & Hijazi, R. (2018). *Open Source Intelligence Methods and Tools*. Apress.https://doi.org/10.1007/978-1-4842-3213-2

Hilalah Alturkistani, &SuriayatiChuprat. (2024). Artificial Intelligence and Large Language Models in Advancing Cyber Threat Intelligence: A Systematic Literature Review. *Research Square (Research Square)*.https://doi.org/10.21203/rs.3.rs-5423193/v1

Hribar, G., Podbregar, I., & Ivanuša, T. (2014). OSINT: A “Grey Zone”?.*International Journal of Intelligence and CounterIntelligence*, *27*(3), 529–549.https://doi.org/10.1080/08850607.2014.900295

Ijiga, O. M., Idoko, I. P., Ebiega, G. I., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing adversarial machine learning for advanced threat detection: AI-driven strategies in cybersecurity risk assessment and fraud prevention. *Open Access Research Journal of Science and Technology*, *11*(1), 001–004.https://doi.org/10.53022/oarjst.2024.11.1.0060

Isabel. (2024). *Evaluating Circular Economy Strategies for Plastics Reduction in the Fast-Moving Consumer Goods Industry*. DIVA.https://www.diva-portal.org/smash/record.jsf?pid=diva2:1899394

Jimmy, F. (2024). Emerging Threats: The Latest Cybersecurity Risks and the Role of Artificial Intelligence in Enhancing Cybersecurity Defenses. *Valley International Journal Digital Library*, 564–574.https://doi.org/10.18535/ijsrm/v9i2.ec01

Joeaneke, P. C., Kolade, T. M., Val, O. O., Olisa, A. O., Joseph, S. A., & Olaniyi, O. O. (2024). Enhancing Security and Traceability in Aerospace Supply Chains through Block Chain Technology. *Journal of Engineering Research and Reports*, *26*(10), 114–135.https://doi.org/10.9734/jerr/2024/v26i101294

Joeaneke, P. C., Val, O. O., Olaniyi, O. O., Ogungbemi, O. S., Olisa, A. O., & Akinola, O. I. (2024). Protecting Autonomous UAVs from GPS Spoofing and Jamming: A Comparative Analysis of Detection and Mitigation Techniques. *Journal of Engineering Research and Reports*, *26*(10), 71–92.https://doi.org/10.9734/jerr/2024/v26i101291

John-Otumu, A. M., Ikerionwu, C., Olaniyi, O. O., Dokun, O., Eze, U. F., &Nwokonkwo, O. C. (2024). Advancing COVID-19 Prediction with Deep Learning Models: A Review. *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2024*, 1–5.https://doi.org/10.1109/seb4sdg60871.2024.10630186

Joseph, S. A. (2024). Balancing Data Privacy and Compliance in Blockchain-Based Financial Systems. *Journal of Engineering Research and Reports*, *26*(9), 169–189.https://doi.org/10.9734/jerr/2024/v26i91271

Kavitha, D., & Thejas, S. (2024). AI Enabled Threat Detection: Leveraging Artificial Intelligence for Advanced Security and Cyber Threat Mitigation. *IEEE Access*, 1–1.https://doi.org/10.1109/access.2024.3493957

Kolade, T. M., Aideyan, N. T., Oyekunle, S. M., Ogungbemi, O. S., & Olaniyi, O. O. (2024). Artificial Intelligence and Information Governance: Strengthening Global Security, through Compliance Frameworks, and Data Security. *Asian Journal of Research in Computer Science*, *17*(12), 36–57.https://doi.org/10.9734/ajrcos/2024/v17i12528

Kuchinka, D., Balazs, S., Gavriletea, M., & Djokic, B.-B. (2018). Consumer Attitudes toward Sustainable Development and Risk to Brand Loyalty. *Sustainability*, *10*(4), 997.https://doi.org/10.3390/su10040997

Kupa, E., Adanma, U. M., Ogunbiyi, E. O., & Solomon, N. O. (2024). Cultivating a culture of safety and innovation in the FMCG sector through leadership and organizational change. *International Journal of Management & Entrepreneurship Research*, *6*(6), 1787–1803.https://doi.org/10.51594/ijmer.v6i6.1165

Lloret, A. (2016). Modeling corporate sustainability strategy. *Journal of Business Research*, *69*(2), 418–425.

Maddireddy, B. R., &Maddireddy, B. R. (2020). Proactive Cyber Defense: Utilizing AI for Early Threat Detection and Risk Assessment. *International Journal of Advanced Engineering Technologies and Innovations*, *1*(2), 64–83.https://ijaeti.com/index.php/Journal/article/view/321

Mallick, A., & Nath, R. (2024). *Navigating the Cyber security Landscape: A Comprehensive Review of Cyber-Attacks, Emerging Trends, and Recent Developments*.https://worldscientificnews.com/wp-content/uploads/2024/01/WSN-1901-2024-1-69-1.pdf

Marelli, M. (2022). The SolarWinds hack: Lessons for international humanitarian organizations. *International Review of the Red Cross*, *104*(919), 1–18.https://doi.org/10.1017/s1816383122000194

Min, A. (2023). Artifical Intelligence and Bias: Challenges, Implications, and Remedies. *Journal of Social Research*, *2*(11), 3808–3817.https://doi.org/10.55324/josr.v2i11.1477

Mirza, A., & Iqbal, R. (2024). Harnessing AI in IT Operations: Transforming Automation and Efficiency. *Asian American Research Letters Journal*, *1*(9), 22–34.https://doi.org/10.5281/cdv3hw65

Mirza, S., Begum, L., Niu, L., Pardo, S., Abouzied, A., Papotti, P., &Pöpper, C. (2023). Tactics, Threats & Targets: Modeling Disinformation and its Mitigation. *Proceedings 2023 Network and Distributed System Security Symposium*.https://doi.org/10.14722/ndss.2023.23657

Mkhize, T., Carter, W., Khumalo, N., Thompson, J., Dube, S., Smith, A., & Bennett, O. (2022, March 24). *Natural Language Processing Frameworks for Detecting Phishing Attacks and Text-Based Intrusions*. ResearchGate International Journal of Information and Cybersecurity; DLpress.

Nissenbaum, H. (2020). Protecting Privacy in an Information Age: The Problem of Privacy in Public. *Routledge EBooks*, 141–178.https://doi.org/10.4324/9781003075011-12

Okon, S. U., Olateju, O. O., Ogungbemi, O. S., Joseph, S. A., Olisa, A. O., & Olaniyi, O. O. (2024). Incorporating Privacy by Design Principles in the Modification of AI Systems in Preventing Breaches across Multiple Environments, Including Public

Cloud, Private Cloud, and On-prem. *Journal of Engineering Research and Reports*, *26*(9), 136–158.https://doi.org/10.9734/jerr/2024/v26i91269

Olabanji, S. O., Marquis, Y. A., Adigwe, C. S., Abidemi, A. S., Oladoyinbo, T. O., & Olaniyi, O. O. (2024). AI-Driven Cloud Security: Examining the Impact of User Behavior Analysis on Threat Detection. *Asian Journal of Research in Computer Science*, *17*(3), 57–74.https://doi.org/10.9734/ajrcos/2024/v17i3424

Olabanji, S. O., Olaniyi, O. O., & Olagbaju, O. O. (2024). Leveraging Artificial Intelligence (AI) and Blockchain for Enhanced Tax Compliance and Revenue Generation in Public Finance. *Asian Journal of Economics, Business and Accounting*, *24*(11), 577–587.https://doi.org/10.9734/ajeba/2024/v24i111577

Olabanji, S. O., Oluwaseun Oladeji Olaniyi, O. O., &Olaoye, O. O. (2024). Transforming Tax Compliance with Machine Learning: Reducing Fraud and Enhancing Revenue Collection. *Asian Journal of Economics Business and Accounting*, *24*(11), 503–513.https://doi.org/10.9734/ajeba/2024/v24i111572

Oladoyinbo, T. O., Olabanji, S. O., Olaniyi, O. O., Adebiyi, O. O., Okunleye, O. J., & Alao, A. I. (2024). Exploring the Challenges of Artificial Intelligence in Data Integrity and its Influence on Social Dynamics. *Asian Journal of Advanced Research and Reports*, *18*(2), 1–23.https://doi.org/10.9734/ajarr/2024/v18i2601

Olaniyi, O. O. (2024). Ballots and Padlocks: Building Digital Trust and Security in Democracy through Information Governance Strategies and Blockchain Technologies. *Asian Journal of Research in Computer Science*, *17*(5), 172–189.https://doi.org/10.9734/ajrcos/2024/v17i5447

Olaniyi, O. O., Olaoye, O. O., & Okunleye, O. J. (2023). Effects of Information Governance (IG) on Profitability in the Nigerian Banking Sector. *Asian Journal of Economics, Business and Accounting*, *23*(18), 22–35.https://doi.org/10.9734/ajeba/2023/v23i181055

Olaniyi, O. O., Omogoroye, O. O., Olaniyi, F. G., Alao, A. I., & Oladoyinbo, T. O. (2024). CyberFusion Protocols: Strategic Integration of Enterprise Risk Management, ISO 27001, and Mobile Forensics for Advanced Digital Security in the Modern Business Ecosystem. *Journal of Engineering Research and Reports*, *26*(6), 32.https://doi.org/10.9734/JERR/2024/v26i61160

Olateju, O. O., Okon, S. U., Igwenagu, U. T. I., Salami, A. A., Oladoyinbo, T. O., & Olaniyi, O. O. (2024). Combating the Challenges of False Positives in AI-Driven Anomaly Detection Systems and Enhancing Data Security in the Cloud. *Asian Journal of Research in Computer Science*, *17*(6), 264–292.https://doi.org/10.9734/ajrcos/2024/v17i6472

Olateju, O. O., Okon, S. U., Olaniyi, O. O., Samuel-Okon, A. D., & Asonze, C. U. (2024). Exploring the Concept of Explainable AI and Developing Information Governance Standards for Enhancing Trust and Transparency in Handling Customer Data. *Journal of Engineering Research and Reports*, *26*(7), 244–268.https://doi.org/10.9734/jerr/2024/v26i71206

Ozkan-Ozay, M., Akin, E., Aslan, Ö., Kosunalp, S., Iliev, T., Stoyanov, I., &Beloev, I. (2024). A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions. *IEEE Access*, *12*, 12229–12256.https://doi.org/10.1109/access.2024.3355547

Paladini, T., Ferro, L., Polino, M., Stefano Zanero, & Carminati, M. (2024). You Might Have Known It Earlier: Analyzing the Role of Underground Forums in Threat Intelligence. *Association for Computing Machinery*, 368–383.https://doi.org/10.1145/3678890.3678930

Pasaribu, H. (2024). *Welcome To Zscaler Directory Authentication*. Thejoas.com.http://thejoas.com/index.php/thejoas/article/view/72

Pastor-Galindo, J., Nespoli, P., GomezMarmol, F., &MartinezPerez, G. (2020). The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends. *IEEE Access*, *8*(8), 10282–10304.https://doi.org/10.1109/access.2020.2965257

Patel, K. R. (2023). Harmonizing Sustainability, Functionality, and Cost: Navigating Responsible Packaging Innovations in Modern Supply Chains. *American Journal of Economic and Management Business (AJEMB)*, *2*(8), 287–300.http://ajembjournal.us/index.php/gp/article/view/51

Prall , C. (2025, January 7). *AI-Native XDR | CrowdStrike*. Crowdstrike.com.https://www.crowdstrike.com/en-us/cybersecurity-101/endpoint-security/ai-native-xdr/

Qiu, S., Liu, Q., Zhou, S., & Wu, C. (2019). Review of Artificial Intelligence Adversarial Attack and Defense Technologies. *Applied Sciences*, *9*(5), 909.https://doi.org/10.3390/app9050909

RecordedFuture. (2024). *Recorded Future Launches Enterprise AI for Intelligence*. Recordedfuture.com.https://www.recordedfuture.com/press-releases/recorded-future-launches-enterprise-ai-for-intelligence

Rodriguez, A., & Okamura, K. (2020). Enhancing data quality in real-time threat intelligence systems using machine learning. *Social Network Analysis and Mining*, *10*(1).https://doi.org/10.1007/s13278-020-00707-x

Saeed, S., Suayyid, S. A., Al-Ghamdi, M. S., Al-Muhaisen, H., &Almuhaideb, A. M. (2023). A Systematic Literature Review on Cyber Threat Intelligence for Organizational Cybersecurity Resilience. *Sensors*, *23*(16), 7273.https://doi.org/10.3390/s23167273

Salako, A. O., Fabuyi, J. A., Aideyan, N. T., Selesi-Aina, O., Dapo-Oyewole, D. L., & Olaniyi, O. O. (2024). Advancing Information Governance in AI-Driven Cloud Ecosystem: Strategies for Enhancing Data Security and Meeting Regulatory Compliance. *Asian Journal of Research in Computer Science*, *17*(12), 66–88.https://doi.org/10.9734/ajrcos/2024/v17i12530

Samuel-Okon, A. D., Akinola, O. I., Olaniyi, O. O., Olateju, O. O., & Ajayi, S. A. (2024). Assessing the Effectiveness of Network Security Tools in Mitigating the Impact of Deepfakes AI on Public Trust in Media. *Archives of Current Research International*, *24*(6), 355–375.https://doi.org/10.9734/acri/2024/v24i6794

Sarker, I. H. (2021a). Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. *SN Computer Science*, *2*(3).https://doi.org/10.1007/s42979-021-00535-6

Sarker, I. H. (2021b). Deep Learning: a Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, *2*(6). Springer.https://doi.org/10.1007/s42979-021-00815-1

Sarker, I. H. (2024). *AI-Driven Cybersecurity and Threat Intelligence*. Springer.https://doi.org/10.1007/978-3-031-54497-2

Selesi-Aina, O., Obot, N. E., Olisa, A. O., Gbadebo, M. O., Olateju, O. O., & Olaniyi, O. O. (2024). The Future of Work: A Human-centric Approach to AI, Robotics, and Cloud Computing. *Journal of Engineering Research and Reports*, *26*(11), 62–87.https://doi.org/10.9734/jerr/2024/v26i111315

Shah, A. (2024, December 13). *AI Attacks on the Rise - Amazon Fights 1 Billion+ Threats a Day*. AccuKnox.https://www.accuknox.com/blog/ai-attacks-on-the-rise

Shah, V. (2022, December 4). *Machine Learning Algorithms for Cybersecurity: Detecting and Preventing Threats*. ResearchGate.https://www.researchgate.net/profile/Varun-Shah-27/publication/378396020_Machine_Learning_Algorithms_for_Cybersecurity_Detecting_and_Preventing_Threats/links/65e8d7eaadf2362b637d02ad/Machine-Learning-Algorithms-for-Cybersecurity-Detecting-and-Preventing-Threats.pdf

Sharma, S., & Arjunan, T. (2023). Natural Language Processing for Detecting Anomalies and Intrusions in Unstructured Cybersecurity Data. *International Journal of Information and Cybersecurity*, *7*(12), 1–24.https://publications.dlpress.org/index.php/ijic/article/view/69

Shneiderman, B. (2020). Bridging the Gap between Ethics and Practice. *ACM Transactions on Interactive Intelligent Systems*, *10*(4), 1–31.https://doi.org/10.1145/3419764

Singh, S. (2021). Energy Crisis and Climate Change. *Energy*, 1–17.https://doi.org/10.1002/9781119741503.ch1

Slinde, J. S. (2023). *Unveiling the Potential of Open-Source Intelligence (OSINT) for Enhanced Cybersecurity Posture*. Uia.brage.unit.no.https://uia.brage.unit.no/uia-xmlui/handle/11250/3084458

Soyege, O. O., Makinde, G. O., & Akinlabi, B. H. (2023). Green Supply Chain Management and Organizational Performance of Fast-Moving Consumer Goods Firms in Lagos Nigeria. *International Journal of Entrepreneurship*, *6*(2), 1–20.https://doi.org/10.47672/ije.1517

Szymoniak, S., & Kacper Foks. (2024). Open Source Intelligence Opportunities and Challenges: a Review. *Advances in Sciences and Technology/PostępyNauki I Techniki*, *18*(3), 123–139.https://doi.org/10.12913/22998624/186036

Tapscott, B. (2024). *Trivergence*. Google Books.https://books.google.com/books?hl=fr&lr=&id=qpTuEAAAQBAJ&oi=fnd&pg=PT4&dq=intelligence+generated+by+AI-processed+OSINT++systems+provide+timely

Tripathi, A., Waqas, A., Venkatesan, K., Yilmaz, Y., & Rasool, G. (2024). Building Flexible, Scalable, and Machine Learning-Ready Multimodal Oncology Datasets. *Sensors*, *24*(5), 1634.https://doi.org/10.3390/s24051634

U. A. Usmani, A. Happonen, & J. Watada. (2022). A Review of Unsupervised Machine Learning Frameworks for Anomaly Detection in Industrial Applications. *Science and Information Conference*, *2*, 158–189.https://doi.org/10.1007/978-3-031-10464-0_11

Val, O. O., Kolade, T. M., Gbadebo, M. O., Selesi-Aina, O., Olateju, O. O., & Olaniyi, O. O. (2024). Strengthening Cybersecurity Measures for the Defense of Critical

Infrastructure in the United States. *Asian Journal of Research in Computer Science*, *17*(11), 25–45.https://doi.org/10.9734/ajrcos/2024/v17i11517

Val, O. O., Olaniyi, O. O., Selesi-Aina, O., Gbadebo, M. O., &Kolade, T. M. (2024). Machine Learning-enabled Smart Sensors for Real-time Industrial Monitoring: Revolutionizing Predictive Analytics and Decision-making in Diverse Sector. *Asian Journal of Research in Computer Science*, *17*(11), 92–113.https://doi.org/10.9734/ajrcos/2024/v17i11522

Wang, L., Ma, Y., Yan, J., Chang, V., Zomaya, A., Wang, L., Ma, Y., Yan, J., Chang, V., &Zomaya, A. (2018). *To appear in: Future Generation Computer Systems*.https://doi.org/10.1016/j.future.2016.06.009

Wang, X., Wu, Y. C., Zhou, M., & Fu, H. (2024). Beyond surveillance: Privacy, ethics, and regulations in face recognition technology. *Frontiers in Big Data*, *7*.https://doi.org/10.3389/fdata.2024.1337465

Watters, P. A. (2023). Counterintelligence in a Cyber World. In *Springer eBooks*. Springer Nature.https://doi.org/10.1007/978-3-031-35287-4

Williamson, S. M., &Prybutok, V. (2024). The Era of Artificial Intelligence Deception: Unraveling the Complexities of False Realities and Emerging Threats of Misinformation. *Information*, *15*(6), 299.https://doi.org/10.3390/info15060299

Yadav, A., Kumar, A., & Singh, V. (2023). Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber security. *Artificial Intelligence Review*.https://doi.org/10.1007/s10462-023-10454-y