A Framework for Embedding Generative and Agentic AI in Open Source Intelligence

Eduardo Almeida Palmieri, Mohamed Chahine Ghanem, Viktor Sowinski-Mydlarz * and Dipo Dunsin
Cyber Security Research Centre
London Metropolitan University
London, N78DB, UK
* Corresponding Author: eda0203@my.londonmet.ac.uk

***Abstract*—Open Source Intelligence (OSINT) plays a critical role in cybersecurity and threat intelligence. However, traditional methods are slow, manual, and difficult to scale. Although Large Language Models (LLMs) and Generative AI have been explored for OSINT, most existing approaches apply them to isolated tasks without developing an integrated and autonomous architecture. This paper proposes a novel Agentic AI-driven OSINT framework that enables autonomous information gathering, reasoning, and tool orchestration across heterogeneous open-source data streams. The system uses retrieval-augmented generation (RAG), chain-of-thought reasoning, and adaptive agent planning to determine which tools to invoke, how to process intermediate outputs, and when to escalate findings for human review. The proposed architecture includes modules for multi-source data ingestion, LLM-powered analysis, generative scenario simulation, and ethical safeguard enforcement. A proof-of-concept use case involving the location of missing persons from public data demonstrates how the framework improves coverage, accuracy, and decision speed when compared to conventional OSINT workflows. This work introduces the first unified and reproducible design for an Agentic AI OSINT system that incorporates transparency, accountability, and ethical compliance into its operational core.**

***Index Terms*—Large Language Models (LLMs), Generative AI, Open-Source Intelligence (OSINT), Cybersecurity, Chain-of-Thought Reasoning, Digital Investigations.**

1. Introduction

The rapid growth of unstructured, multilingual, and multimodal data from social media, online forums, dark web sources, and public repositories has transformed the field of cyber threat intelligence. OSINT is now essential for detecting threats, responding to incidents, and supporting humanitarian investigations such as the search for missing persons. Despite its importance, traditional OSINT workflows rely heavily on manual searches, keyword filtering, and disconnected tools. As a result, they are often slow, prone to errors, and unable to scale to meet the speed and complexity of modern information environments [1].

Recent advances in Large Language Models and Generative AI have introduced opportunities for automation in OSINT. However, existing applications typically depend on fixed pipelines or task-specific scripts. These approaches do not provide agentic capabilities, which means they cannot autonomously decide what information to seek, which analytical tools to use, and how to iteratively refine outputs based on evolving evidence. Without such capabilities, AI-enabled OSINT systems remain limited in adaptability, efficiency, and operational intelligence.

In response to these limitations, this paper proposes a unified Agentic AI OSINT framework that integrates LLM reasoning, generative simulation, and tool orchestration into a modular and scalable architecture. The system includes autonomous decision-making agents that dynamically select and execute OSINT tasks. It employs retrieval-augmented generation and chain-of-thought prompting for deeper contextual reasoning. It also applies generative AI for adversarial scenario simulation, synthetic training data creation, and automated report drafting. Furthermore, the framework incorporates built-in ethical safeguards, including bias detection, transparency logging, and human-in-the-loop validation [2].

The framework is validated through a proof-of-concept scenario that involves locating missing persons by autonomously monitoring and correlating signals from diverse online sources. Comparative results with baseline OSINT workflows show significant improvements in coverage, accuracy, and timeliness. The contributions of this work include the first end-to-end agentic AI architecture for OSINT, a practical method for integrating AI agents with traditional intelligence tools, and an ethical blueprint for deploying such systems responsibly in high-stakes environments.

1.1 Research Motivation

The growing complexity, velocity, and diversity of open-source information streams have transformed the operational demands placed on OSINT practitioners. Manual and semi-automated methods struggle to keep pace with the real-time nature of online events, leaving gaps in situational awareness and threat detection. While recent developments in Large Language Models (LLMs) and Generative AI have shown potential for automation, current OSINT systems generally lack the capacity for adaptive decision-making, multi-tool orchestration, and autonomous refinement of analytical strategies. This limitation creates an opportunity to design an OSINT architecture that not only processes vast amounts of heterogeneous data but also plans, prioritises, and executes investigative actions dynamically [3]. The motivation behind this research is to bridge the gap between isolated AI capabilities and a fully integrated, agentic intelligence-gathering system that can operate effectively in high-stakes contexts such as cyber incident response and humanitarian investigations [18].

1.2 Research Context and Scope

The proposed Agentic AI OSINT framework arises from three converging trends in intelligence gathering. First, the rapid expansion of publicly available digital content, spanning social media, news media, and dark web marketplaces, has made OSINT indispensable yet increasingly complex. The scale, diversity, and velocity of these data streams necessitate automated, real-time analytical capabilities. Second, advances in large language models (LLMs) have introduced unprecedented capacities in language understanding, reasoning, and cross-lingual processing, making them potentially transformative for OSINT when embedded in adaptive, scalable architectures. Third, progress in agent-based AI and tool orchestration now enables systems that autonomously plan, monitor, and adapt investigative strategies rather than following static pipelines. Building on these developments, this research pursues two objectives. Technologically, it designs and validates a modular framework that integrates LLM-based reasoning, generative AI for scenario simulation, multi-tool orchestration, and embedded ethical safeguards. Operationally, it evaluates the framework's capacity to enhance the efficiency and reliability of OSINT workflows in high-stakes contexts. The proof-of-concept focuses on locating missing persons through open-source signals, a scenario selected for its complexity, urgency, and real-world applicability. However, the architecture is intended to be extensible to other domains, including cyber threat intelligence, crisis monitoring, and geopolitical analysis.

1.3 Research Aims

This research aims to design, implement, and validate a unified Agentic AI framework for Open Source Intelligence (OSINT) that operates autonomously, adapts to evolving intelligence requirements, and integrates with existing analytical tools. The framework will enhance OSINT workflows in cybersecurity and humanitarian contexts by increasing the speed, accuracy, and scalability of intelligence gathering and analysis. It will combine autonomous decision-making, retrieval-augmented generation, and chain-of-thought reasoning to enable dynamic processing, while leveraging Large Language Models with generative AI capabilities for adversarial simulation, synthetic data generation, and automated reporting. A modular orchestration layer will facilitate flexible interaction with established OSINT platforms and APIs, ensuring adaptability for future integrations. Ethical safeguards and human oversight will be embedded in the design to maintain transparency, accountability, and legal compliance. The framework's effectiveness will be demonstrated through a realistic use case, such as locating missing persons, to highlight its operational advantages over traditional OSINT methods.

1.4 Research Potential Contributions

This research aims to contribute to OSINT, cybersecurity, and AI-driven investigative systems by introducing a novel Agentic AI OSINT framework that resolves long-standing operational challenges. It will deliver a validated system architecture that unifies autonomous AI agents, advanced LLM-driven analysis, generative simulation, and multi-tool orchestration into an operationally ready framework. It will also embed transparency, bias mitigation, and human oversight into the design to ensure ethical compliance from the outset. In parallel, it will provide a practical blueprint for integrating AI tools with widely used OSINT platforms, improving interoperability and analyst efficiency. The research will further present empirical evidence of performance gains through a proof-of-concept comparison with current OSINT workflows in speed, coverage, and accuracy. Ultimately, it will offer a scalable, ethically compliant model that supports future academic study, policy development, and real-world deployment in high-stakes intelligence contexts.

1.5 Research Questions

This research is guided by the following questions:

1) How can the integration of agentic AI capabilities, such as autonomous decision-making and adaptive tool orchestration, directly address specific limita-

tions in current OSINT workflows?
2) How can Large Language Models (LLMs) be effectively incorporated into an agentic architecture to improve the accuracy, coverage, and efficiency of intelligence extraction from heterogeneous open-source data streams?
3) In what ways can Generative AI enhance OSINT operations through simulation, synthetic data generation, and automated report drafting while maintaining reliability and ethical compliance?
4) How does the proposed agentic AI OSINT framework perform in a real-world-inspired use case, such as locating missing persons, when compared to traditional and semi-automated OSINT workflows?
5) What ethical, legal, and operational safeguards are necessary to ensure that an agentic AI OSINT system operates responsibly and remains suitable for high-stakes environments?

1.6 Research Approach

This research follows a structured, multi-phase approach to designing, developing, and validating an Agentic AI OSINT framework. It begins with a targeted literature review to identify specific limitations in current OSINT systems, particularly their limited adaptability, autonomy, and scalability across diverse data sources. Building on these findings, the second phase designs a system architecture that integrates autonomous decision-making agents, advanced LLM-driven analysis, generative simulation, and multi-tool orchestration into a cohesive operational model, ensuring seamless interaction between modules for end-to-end intelligence processing with minimal human input. The third phase develops and tests a proof-of-concept prototype in a controlled environment using a realistic case, such as locating missing persons, and evaluates its performance against traditional OSINT workflows to assess gains in accuracy, timeliness, and adaptability. Finally, ethical and legal safeguards are embedded within the framework to ensure regulatory compliance and uphold fairness, transparency, and accountability, resulting in a validated and operationally viable blueprint for deploying Agentic AI in OSINT contexts.

2. Literature Review

The integration of Large Language Models (LLMs) and related AI techniques into Open Source Intelligence (OSINT) has evolved rapidly, yet current solutions often lack fully autonomous, agentic capabilities. This section reviews prior work in four thematic areas that directly inform the proposed framework: LLM-enhanced OSINT and tool integration, reasoning and chain-of-thought techniques, generative AI applications, and ethical and operational considerations.

2.1 LLM-Enhanced OSINT and Tool Integration

Recent studies have demonstrated the potential of LLMs to enhance OSINT workflows through improved data retrieval, entity extraction, and analysis. Loumachi et al. [4] applied retrieval-augmented generation (RAG) to cyber incident timeline analysis, demonstrating gains in accuracy and contextual reasoning [17]. Cerny [1] examined the implications of LLMs for OSINT information acquisition, emphasising the role of prompt engineering in analyst workflows. Yuan et al. [2] proposed an open-source OSINT acquisition method that equips LLMs with external toolkits for improved search and structured data collection. Other work has explored comprehensive LLM-based OSINT enhancements [5], the evaluation of chatbot-driven OSINT for cyber threat awareness [10], and the co-design of collaborative OSINT systems for vulnerability assessment [11]. The importance of integrating LLMs into specialised OSINT platforms has been highlighted by Ghanem et al. [13], who developed protocols for forensic analysis of dark web activities. Hassanin and Moustafa [20] provided a broad survey of LLM opportunities in cyber defence, while Yang et al. [15] demonstrated LLM-driven automation for domain-specific threat modelling. Collectively, these works reveal significant potential but also confirm that most implementations remain task-specific or semi-automated, lacking adaptive orchestration across tools [19].

2.2 Reasoning and Chain-of-Thought Techniques

Effective OSINT often requires the synthesis of fragmented information into coherent intelligence. Brown et al. [3] introduced few-shot prompting for LLMs, while Wei et al. [9] demonstrated how chain-of-thought (CoT) prompting improves reasoning performance. Zhao et al. [7] proposed Verify-and-Edit, a knowledge-enhanced CoT framework for refining outputs. Ge et al. [25] applied CoT techniques to integrate LLM agents with concept generation, while Ferrag et al. [16] surveyed reasoning challenges and open problems in LLM research. Broader surveys [29] and targeted studies [23] underline the importance of combining reasoning capabilities with adaptive workflows. However, existing applications often limit CoT use to isolated tasks rather than embedding it within an autonomous, context-driven planning agent.

TABLE I: Comprehensive Tool Integration Framework for AI-Driven OSINT

Tool/API	Integration Approach	Features Enabled	Example Use Case	Limitations Constraints	Typical Response Time	Integration Complexity
Shodan	Python/REST API, agent framework	Search for exposed devices, ports, vulnerabilities	Locate open RDP servers in a given region	API rate limits; some data behind paywall	Seconds to minutes	Easy (well-documented API)
Censys	API integration	Internet-wide search for hosts, certificates	Asset discovery; vulnerability mapping	API key required; rate limits	Seconds to minutes	Easy
Twitter/X API	API integration (rate limits apply)	Real-time social media monitoring	Monitor hashtags; extract threat chatter	Strict rate limits; approval required; data access may change	Real time to minutes	Moderate (API setup and approval)
Telegram Bots	Telethon, python-telegram-bot, API	Group/channel monitoring; message extraction	Summarise threat actor discussions	Bot must be added to groups/channels; API subject to change	Seconds to minutes	Moderate (bot setup and API handling)
Pastebin Scraper	Custom scripts/APIs	Leak detection; credential dumps	Monitor for new pastes; extract IOCs	May be blocked; scraping restrictions apply	Seconds to minutes	Moderate (custom code)
Custom Web Scrapers	Selenium, BeautifulSoup, Scrapy	Data extraction from forums, news sites, or the dark web	Extract posts from dark web forums or security blogs	Anti-bot measures; legal and ethical concerns	Seconds to minutes	Hard custom code
Threat Intel Feeds	API integration	Threat intelligence data; IOC ingestion	Fetch the latest IOCs linked to a threat actor	Depends on provider; some feeds are paid	Seconds to minutes	Easy to moderate
VirusTotal	API integration, callable function	File, URL, and hash reputation; malware analysis	Query file hashes; generate risk summaries	API rate limits; privacy implications	Seconds to minutes	Easy
Hybrid Analysis	API integration	Malware sandbox reports; IOC extraction	Submit files/URLs and parse sandbox results	API key required; subject to rate limits	Minutes	Easy
MISP	PyMISP library, REST API	Threat sharing; IOC management; event correlation	Automate IOC ingestion, deduplication, and enrichment	Requires setup; data sharing governed by policy	Seconds to minutes	Moderate (initial setup and API handling)
CVE/NVD	REST API, data dump parsing	CVE details; CVSS scores; exploit references	Retrieve the latest CVEs and assess product exposure	Data may lag; requires parsing effort	Seconds to minutes	Moderate
GeoIP Services	API (MaxMind, ipinfo.io)	IP geolocation; ASN lookup	Enrich IP addresses with geolocation and ASN data	API key required; rate limits apply	Seconds	Easy
WHOIS Lookup	API or command-line integration	Domain registration data; ownership history	Automate domain lookups; flag suspicious changes	Rate limits; privacy masking by registrars	Seconds	Easy
OpenAI/LLM APIs	Direct integration	Summarisation; translation; named entity recognition; reasoning	Orchestrate tool use; interpret results; generate reports	API cost; privacy concerns; prompt length limits	Seconds	Easy

2.3 Generative AI Applications in OSINT and Cyber Threat Intelligence

Generative AI has been explored for data synthesis, scenario simulation, and intelligence augmentation. Zhou et al. [6] analysed the application of generative AI in OSINT, noting its potential for automated content generation and adversarial scenario modelling. Ranade et al. [8] demonstrated the generation of fake cyber threat intelligence using transformers, while Saddi et al. [12] examined generative AI for enhancing threat intelligence operations. Broader overviews of generative AI advances [14] emphasise emerging opportunities and challenges. Domain-specific innovations include the integration of LLMs with speech recognition [21], intent mining in dialogue systems [22], domain model extraction from textual requirements [24], LLM-enabled IIoT fault diagnosis [26], and sentiment classification for AI-related social media content [27], [28]. While these works illustrate technical diversity, few integrate generative AI outputs directly into real-time OSINT decision-making loops.

2.4 Ethical, Operational, and Analyst-Centric Considerations

Ethical and operational concerns in AI-driven OSINT include privacy, bias, and transparency. Mukhopadhyay and Luther [11] explored co-design approaches that involve analysts in AI-augmented investigations, ad-

dressing trust and usability. Shafee et al. [10] evaluated chatbot-based OSINT tools for accuracy and reliability, while Cerny [1] highlighted risks related to over-reliance on AI outputs. Ghanem et al. [13] stressed the need for secure handling of sensitive intelligence, particularly from deep and dark web sources. These studies converge on the principle that ethical safeguards should be embedded in system architecture rather than applied as afterthoughts. This proactive integration of ethics boosts user trust and guarantees the design of AI systems with accountability in mind.

2.5 Summary of Gaps

The reviewed literature confirms that while LLMs, reasoning frameworks, and generative AI have each advanced OSINT capabilities, most implementations remain static, task-bound, or loosely integrated. Few systems exhibit the agentic qualities necessary for autonomous, adaptive, and ethically compliant intelligence gathering. The proposed Agentic AI OSINT framework addresses these gaps by combining autonomous task planning, embedded reasoning, integrated generative capabilities, and built-in ethical safeguards into a unified architecture.

3. Methodology and System Design

This research introduces a complete Agentic AI OSINT framework designed to overcome the limitations of existing manual and semi-automated intelligence workflows. The methodology focuses on creating a modular autonomous system that can gather, analyse, and synthesise open-source intelligence while operating within ethical and legal boundaries. The framework consists of six integrated components, each contributing to the system's ability to operate with minimal human intervention and high adaptability.

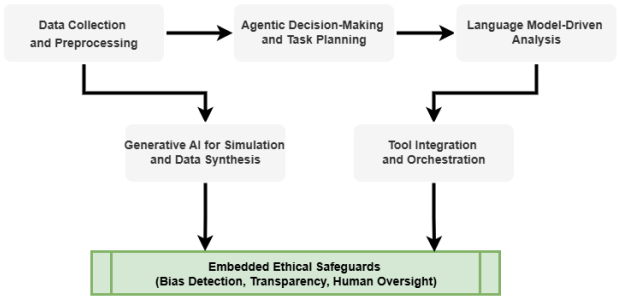

Fig. 1: High-Level Agentic AI OSINT System Architecture

3.1 Data Collection and Preprocessing

The first stage involves the continuous collection of publicly available information from multiple OSINT sources, including social media platforms, dark web forums, security blogs, vulnerability databases, and breach repositories. The system uses APIs, web crawlers, and bot frameworks to access these data streams. Data preprocessing includes noise filtering, language normalisation, duplicate removal, and content tagging. A lightweight stream processing layer ensures low-latency ingestion, which is essential for time-sensitive scenarios such as cyber incident response or urgent humanitarian investigations.

3.2 Agentic Decision-Making and Task Planning

At the core of the framework is an agentic decision-making module that autonomously determines which data sources to query, which tools to invoke, and how to sequence analysis tasks. This planning process is guided by predefined objectives, such as detecting indicators of compromise or locating individuals based on public activity. The decision-making module evaluates the relevance and quality of incoming data, adapts its strategy to evolving situations, and escalates high-priority findings for human review when necessary.

3.3 Language Model-Driven Analysis

The analytical component uses advanced Large Language Models such as GPT, BERT, and LLaMA to extract threat-relevant entities, detect sentiment, and generate structured summaries. Prompt engineering techniques, including few-shot prompting and chain-of-thought reasoning, are employed to improve contextual accuracy. Retrieval-augmented generation is used to integrate up-to-date information from live data sources, thereby overcoming static knowledge cut-off limitations. This combination enables precise and timely analysis of diverse intelligence inputs.

3.4 Generative AI for Simulation and Data Synthesis

Generative AI methods, including transformer-based text generators and generative adversarial networks (GANs), are used to simulate potential adversarial scenarios and enrich training datasets. These capabilities support red-teaming exercises, model robustness testing, and synthetic report generation in structured formats such as STIX or MITRE ATT&CK. The simulation module also provides scenario-based forecasting, which helps analysts anticipate future developments and evaluate operational risks.

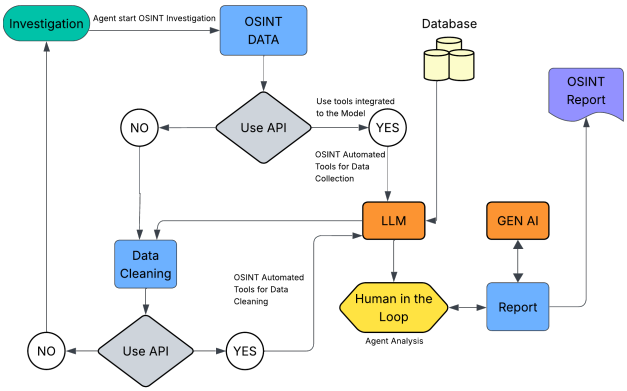

Fig. 2: Proposed Agentic AI OSINT Workflow Loop

3.5 Proposed Integration and Orchestration

The framework incorporates an orchestration layer that connects the AI agents to established OSINT tools such as Maltego, Shodan, VirusTotal, and MISP. The orchestration engine determines the appropriate sequence of tool calls based on task requirements and integrates their outputs into the ongoing analysis. This enables smooth collaboration between the agentic AI system and existing cybersecurity infrastructures while maintaining flexibility for future tool integration, as illustrated in Figure 2

4. Proof-of-Concept Use Case: Locating Missing Persons

To demonstrate the system's practical value, this research presents a proof-of-concept scenario in which the framework autonomously monitors and correlates data to assist in locating missing persons. In this use case, the system simulates the real-time collection of social media posts, open databases, and dark web chatter, followed by entity extraction, credibility assessment, and generative scenario simulation. Finally, the process ends with passing the credible findings to a human analyst for review.

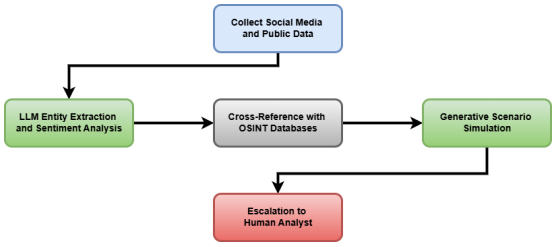

Fig. 3: Missing Persons Use Case Flow

5. Conclusion and Future Work

This paper presents a unified Agentic AI OSINT framework that integrates autonomous decision-making, advanced language model analysis, generative simulation, and tool orchestration into a single operational architecture. The system overcomes key limitations of traditional OSINT by enabling continuous, adaptive, and ethically compliant intelligence gathering and analysis. Unlike existing AI-based approaches, it offers a fully integrated environment where AI agents can plan, act, and refine strategies in response to evolving intelligence needs. Incorporating retrieval-augmented generation, chain-of-thought reasoning, and embedded ethical safeguards, the framework delivers improved coverage, accuracy, and speed. A proof-of-concept for locating missing persons demonstrates measurable gains over conventional methods. Contributions include the first reproducible agentic AI architecture for OSINT, a practical blueprint for integrating autonomous agents with existing tools, and an embedded ethical model suited to sensitive operations.

Future developments for the proposed Agentic AI OSINT framework will focus on three main areas: expanded deployment, broader evaluation, and increased accessibility for the research and practitioner communities. First, the system will be deployed in real-world operational environments across both cybersecurity and humanitarian domains. This will involve integration with live OSINT feeds, threat intelligence sharing platforms, and investigative collaboration tools used by law enforcement and emergency response agencies. Second, the system will be benchmarked extensively against existing OSINT tools and platforms using both publicly available and domain-specific datasets. These benchmarks will evaluate processing speed, coverage of relevant intelligence, accuracy of entity recognition and event correlation, and analyst workload reduction. Third, to encourage transparency and foster collaborative improvement, a version of the framework will be prepared for open-source release. This release will include a modular codebase, documentation for integration with common OSINT tools, and a sample dataset for testing. In the long term, the framework will be extended to handle multimodal intelligence streams, including geospatial imagery, audio signals, and sensor data, thereby broadening its applicability to national security, disaster response, and large-scale threat monitoring.

References

[1] Cerny, J. (2024) 'Implications of Large Language Models for OSINT: Assessing the Impact on Information Acquisition and Analyst Expertise in Prompt Engineering', European Conference on Cyber Warfare and Security, 23(1), pp. 116–124. https://doi.org/10.34190/eccws.23.1.2261.

[2] Yuan, X., Wang, J., Zhao, H., Yan, T. and Qi, F. (2024) 'Empowering LLMs With Toolkits: An Open-Source Intelligence Acquisition Method'. Future Internet, vol. 16, no. 12, Dec. 2024, p. 461. DOI.org (Crossref). https://doi.org/10.3390/fi16120461.

[3] Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. and Amodei, D. (2020) 'Language Models are Few-Shot Learners'. arXiv. https://doi.org/10.48550/ARXIV.2005.14165.

[4] Loumachi, F.Y., Ghanem, M.C. and Ferrag, M.A., (2025). Advancing Cyber Incident Timeline Analysis Through Retrieval-Augmented Generation and Large Language Models. Computers, 14(2), p.67. https://doi.org/10.3390/computers14020067.

[5] Rajendran, G., Kumar, A.A., Sridhar, P.K., Perumalsamy, K.K. and Srinivasan, N., A Comprehensive Approach for Enhancing OSINT through Leveraging LLMs. https://mail.irjes.com/Papers/vol13-issue2/H13026166.pdf.

[6] Zhou, L., Qin, Y., Yan, S., Zhang, G. and Hu, W. (2024) '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). 2024 39th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Dalian, China: IEEE, pp. 1603–1608. https://doi.org/10.1109/YAC63405.2024.10598766.

[7] Zhao, R., Li, X., Joty, S., Qin, C. and Bing, L. (2023) 'Verify-And-Edit: A Knowledge-Enhanced Chain-Of-Thought Framework'. arXiv. https://doi.org/10.48550/ARXIV.2305.03268.

[8] Ranade, P., Piplai, A., Mittal, S., Joshi, A. and Finin, T. (2021) 'Generating Fake Cyber Threat Intelligence Using Transformer-Based Models'. arXiv. https://doi.org/10.48550/arXiv.2102.04351

[9] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. and Zhou, D. (2023) 'Chain-of-Thought Prompting Elicits Reasoning in Large Language Models'. arXiv. https://doi.org/10.48550/arXiv.2201.11903.

[10] Shafee, S., Bessani, A. and Ferreira, P.M. (2024) 'Evaluation of LLM Chatbots for OSINT-based Cyber Threat Awareness'. arXiv. https://doi.org/10.48550/arXiv.2401.15127.

[11] Mukhopadhyay, A. and Luther, K. (2024) 'OSINT Clinic: Co-Designing AI-Augmented Collaborative OSINT Investigations for Vulnerability Assessment'. arXiv. https://doi.org/10.48550/arXiv.2409.11672.

[12] Saddi, V.R., Gopal, S.K., Mohammed, A.S., Dhanasekaran, S. and Naruka, M.S. (2024) 'Examine the Role of Generative AI in Enhancing Threat Intelligence and Cyber Security Measures', in 2024 2nd International Conference on Disruptive Technologies (ICDT). Greater Noida, India: IEEE, pp. 537–542. https://doi.org/10.1109/ICDT61202.2024.10489766.

[13] Ghanem, M.C., Mulvihill, P., Ouazzane, K., Djemai, R. and Dunsin, D., 2023. D2WFP: a novel protocol for forensically identifying, extracting, and analysing deep and dark web browsing activities. Journal of Cybersecurity and Privacy, 3(4), pp.808-829. https://doi.org/10.3390/jcp3040036.

[14] Hagos, D.H., Battle, R. and Rawat, D.B. (2024) 'Recent Advances in Generative AI and Large Language Models: Current Status, Challenges, and Perspectives', IEEE Transactions on Artificial Intelligence, 5(12), pp. 5873–5893. https://doi.org/10.1109/TAI.2024.3444742.

[15] Yang, S., Wu, T., Liu, S., Nguyen, D., Jang, S. and Abuadbba, A. (2024) 'ThreatModeling-LLM: Automating Threat Modeling using Large Language Models for Banking System'. arXiv. https://doi.org/10.48550/arXiv.2411.17058.

[16] Ghanem, M.C., (2022). Towards an efficient automation of network penetration testing using model-based reinforcement learning (Doctoral dissertation, City, University of London). https://openaccess.city.ac.uk/id/eprint/29885/.

[17] Alshinwan, M., Memon, A.G., Ghanem, M.C. and Almaayah, M., 2025. Unsupervised text feature selection approach based on improved Prairie dog algorithm for the text clustering. Jordanian Journal of Informatics and Computing, 2025(1), pp.27-36. https://jjic.thestap.com/archives/volume-2025-1/680e7f2bcd605702a281648b.

[18] Ali, A. and Ghanem, M.C., 2025. Beyond detection: large language models and next-generation cybersecurity. SHIFRA, 2025, pp.81-97. https://doi.org/10.70470/SHIFRA/2025/005.

[19] Ghanem, M.C., Almeida Palmieri, E., Sowinski-Mydlarz, W., Al-Sudani, S. and Dunsin, D., 2025. Weaponized IoT: A Comprehensive Comparative Forensic Analysis of Hacker Raspberry Pi and PC Kali Linux Machine. IoT, 6(1), p.18. https://doi.org/10.3390/iot6010018.

[20] Hassanin, M. and Moustafa, N. (2024) 'A Comprehensive Overview of Large Language Models (Llms) For Cyber Defences: Opportunities and Directions'. arXiv. https://doi.org/10.48550/ARXIV.2405.14487.

[21] Guan, B., Cao, J., Wang, X., Wang, Zhuoyue, Sui, M. and Wang, Zixiang (2024) 'Integrated Method of Deep Learning and Large Language Model in Speech Recognition', in 2024 IEEE 7th International Conference on Electronic Information and Communication Technology (ICEICT). Xi'an, China: IEEE, pp. 487–490. https://doi.org/10.1109/ICEICT61637.2024.10671048.

[22] Shukhman, A.E., Badikov, V.R. and Legashev, L.V. (2024) 'Application of Large Language Models For Intent Mining in Goal-oriented Dialogue Systems', in 2024 V International Conference on Neural Networks and Neurotechnologies (NeuroNT). Saint Petersburg, Russian Federation: IEEE, pp. 28–31. https://doi.org/10.1109/NeuroNT62606.2024.10585408.

[23] Zhuang, L., Wang, Q., Song, L. and Wu, P. (2024) 'Research and Application of Artificial Intelligence Large Language Models Based on Feature Enhancement', in 2024 4th International Conference on Consumer Electronics and Computer Engineering (ICCECE). 2024 4th International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China: IEEE, pp. 344–347. https://doi.org/10.1109/ICCECE61317.2024.10504180.

[24] Arulmohan, S., Meurs, M.-J. and Mosser, S. (2023) 'Extracting Domain Models from Textual Requirements in the Era of Large Language Models', in 2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C). 2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), Västerås, Sweden: IEEE, pp. 580–587. https://doi.org/10.1109/MODELS-C59198.2023.00096.

[25] Ge, S., Sun, Y., Cui, Y. and Wei, D. (2025) 'An Innovative Solution to Design Problems: Applying the Chain-of-Thought Technique to Integrate LLM-Based Agents With Concept Generation Methods', IEEE Access, 13, pp. 10499–10512. https://doi.org/10.1109/ACCESS.2024.3494054.

[26] Liu, Y., Zhang, W., Bao, Z., Chai, X., Gu, M., Jiang, W., Zhang, Z., Tian, Y. and Wang, F.-Y. (2024) 'Brain-like Cognition-Driven Model Factory for IIoT Fault Diagnosis by Combining LLMs With Small Models', IEEE Internet of Things Journal, pp. 1–1. https://doi.org/10.1109/JIOT.2024.3503274.

[27] Farzaan, M.A., Ghanem, M.C., El-Hajjar, A. and Ratnayake, D.N., 2025. AI-powered system for an efficient and effective cyber incidents detection and response in cloud environments. IEEE Transactions on Machine Learning in Communications and Networking. https://doi.org/10.1109/TMLCN.2025.3564912

[28] Riad, M.J.A., Debnath, R., Shuvo, M.R., Ayrin, F.J., Hasan, N., Tamanna, A.A. and Roy, P. (2024) 'Fine-Tuning Large Language Models for Sentiment Classification of AI-Related Tweets', in 2024 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). 2024 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Chennai, India: IEEE, pp. 186–191. https://doi.org/10.1109/WIECON-ECE64149.2024.10914746.

[29] Naveed, H., Khan, A.U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N. and Mian, A. (2024) 'A Comprehensive Overview of Large Language Models'. arXiv. https://doi.org/10.48550/arXiv.2307.06435.