Journal of Intelligent Decision Support System (IDSS) 8 (4) (2025), pp. 237-243

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Generative AI and multi-source intelligence for automated security triage

Fhatur Robby Tanzil Herris[1], Hondor Saragih[2], Aninditho[3]

[1,2,3] *Informatika, Universitas Pertahanan Republik Indonesia, Bogor, Indonesia*

Article Info

Article history:

Received Nov 3, 2025
Revised Nov 25, 2025
Accepted Dec 5, 2025

Keywords:

ChatOps
Cyber Security
Generative AI
Indicator of Compromise
SOAR

ABSTRACT

Security Operation Center (SOC) analysts encounter significant delays due to "Swivel Chair Analysis," a manual and fragmented process for triaging Indicators of Compromise (IoC). This study addresses this inefficiency by developing "CyberGuardianBot," an automated ChatOps assistant built using the Rapid Application Development (RAD) methodology and the Telegram Bot API. Applying Security Orchestration, Automation, and Response (SOAR) principles, the system asynchronously orchestrates multi-source intelligence from VirusTotal, AbuseIPDB, URLScan.io, AlienVault OTX, and MobSF. A key novelty is the integration of Google Gemini to perform cognitive synthesis, translating raw API data into actionable insights. Blackbox testing validated the system across 15 test cases, confirming the successful automation of URL, IP, and file triage. The bot generates natural language executive summaries and structured reports (.txt and .pdf), significantly enhancing the speed and accuracy of the triage process while reducing the cognitive load on analysts.

Corresponding Author:

Fhatur Robby Tanzil Herris,
Informatic, Faculty of Defense Engineering and Technology,
Indonesia Defense University,
IPSC Sentul Area, Sukahati, Citeureup District, Bogor Regency, West Java 16810, Indonesia
Email: fhatur.herris@tm.idu.ac.id

Introduction

In the modern digital landscape, Security Operation Center (SOC) analysts face a significant operational challenge known as "Swivel Chair Analysis"—a manual, fragmented workflow for triaging Indicators of Compromise (IoCs) that is increasingly insufficient against the overwhelming volume of attack data (Kazato et al., 2020). This operational inefficiency is exacerbated by the rise of Generative AI (GenAI). Dhoni & Kumar (2023) highlight that GenAI has significantly reduced the technical expertise required for malicious actors, enabling the automated creation of sophisticated threats, including polymorphic malware (Gupta et al., 2023). Consequently, the number of cyber-attacks utilizing GenAI is demonstrably rising (Oh & Shon, 2023), creating a technological arms race (Teo et al., 2024). Attackers now even utilize these tools to automate the evasion of traditional IoC blocking (Zheng et al., 2023). Since IoCs are time-sensitive and lose relevance over time as noted by Tostes et al. (2023), the traditional reactive posture fails to keep pace, necessitating a shift toward proactive, AI-driven defensive systems (Nadella et al., 2025).

To counter this, the state-of-the-art defensive strategy involves using GenAI for automating cyber defense (Gupta et al., 2023), allowing human analysts to focus on critical judgment while AI handles common threat scenarios (Sai et al., 2024). This is often achieved by integrating Large Language Models (LLMs) into ChatOps workflows to unify tool interfaces (Peci et al., 2025; Wang et al., 2024). However, existing research reveals a distinct gap in practical application. Prieto & Blakely (2024) focused primarily on theoretical GenAI defense agents, while Khai & Juremi (2023) developed simple validation tools that lack comprehensive integration. These solutions often fail to fully alleviate the analyst's cognitive load because they present raw, disjointed data rather than synthesized intelligence. There is a lack of empirical frameworks that fuse asynchronous multi-source orchestration with the cognitive synthesis capabilities of LLMs to not just gather, but interpret technical data in a unified workflow.

Addressing this gap, this research details the design and validation of CyberGuardianBot, a novel Security Orchestration, Automation, and Response (SOAR) tool. Unlike previous iterations, this system integrates Google Gemini as a super brain to synthesize complex JSON outputs into coherent, actionable insights. This architecture mitigates the risk of AI hallucinations (Rodger et al., 2025; Teo et al., 2024) by compelling the AI to synthesize provided, verified data rather than generating facts from scratch. The operational objective is to demonstrate a replicable, generative AI-enhanced framework that solves the "Swivel Chair Analysis" problem, building necessary user trust (Krishnamurthy & Vemulapalli, 2025; Salamun et al., 2023) and performance expectancy (Kim et al., 2024) for effective human-AI cooperation in cybersecurity (Nadella et al., 2025).

Method

The development of CyberGuardianBot adopted the Rapid Application Development (RAD) methodology. This approach was selected over traditional models due to the necessity for rapid iteration in prompt engineering for the Large Language Model (LLM) and the flexibility required to adapt to frequent specification changes in external threat intelligence feeds. The system is architected as an intelligent agent within a ChatOps service, implemented via the Telegram Bot API to create a unified, low-code interface (Peci et al., 2025; Wang et al., 2024). The technical environment relies on a backend capable of handling asynchronous requests to orchestrate the multi-source analyst team. This module integrates specialized REST API endpoints: VirusTotal (v3) for multi-vendor reputation, AbuseIPDB for confidence scoring, URLScan.io for DOM analysis, AlienVault OTX for pulse correlation, and the Mobile Security Framework (MobSF) for static analysis.

The core intelligence engine, referred to as the super brain, utilizes the Google Gemini Generative AI model. This design choice is critical; rather than asking the AI to generate facts from scratch, which risks hallucinations (Al Zaidy, 2024; Alqahtani & Kumar, 2025), the system acts as a cognitive synthesis engine (Sternberg, 2024). The AI processes verified, structured JSON payloads from the orchestrated APIs and transforms them into natural-language situational awareness. This data-grounded approach ensures the analyst remains the final decision-maker (Zhang & Yu, 2025).

To validate system reliability, a comprehensive Blackbox Testing methodology was employed. The validation process involved 15 distinct test scenarios covering benign and malicious inputs for URLs, IP addresses, and Files. The success criteria for these tests were defined as: (1) successful API connection (HTTP 200 OK), (2) correct parsing of complex JSON structures, (3) the generation of hallucination-free executive summaries by the AI, and (4) the automated production of accessible .txt and .pdf reports. This testing framework ensures the system meets the functional requirements of a high-stakes SOC environment.

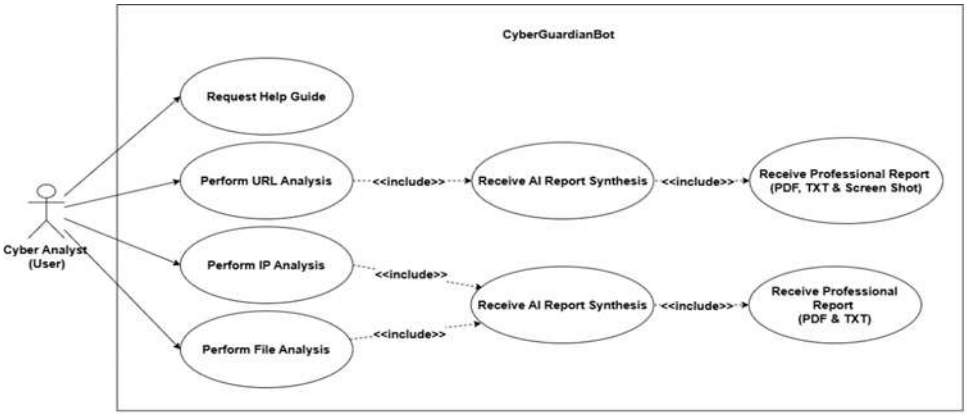

Figure 1. Use Case Diagram

Results and Discussions

The functionality of CyberGuardianBot was validated using the Blackbox Testing methodology. All fifteen defined test cases passed successfully, confirming that the system met the success criteria defined in the methodology: stable asynchronous API orchestration, accurate JSON parsing, and the generation of hallucination-free intelligence summaries.

The system's core capabilities are demonstrated through three primary workflows. First, in the URL analysis workflow, the system was tested against a known malicious domain. The AI-synthesized report (Figures 2-4) correctly assessed the target as "MALICIOUS" with a critical risk score of 90/100. Crucially, the system did not just flag the domain but identified specific threat vectors, synthesizing data to highlight potential phishing indicators and WordPress vulnerabilities.

Figure 2. Scanning the URL https://bagas-31.com/ with a bot.

Figure 3. The bot provides the scan results: a TXT report, a PDF report, and a website screenshot.

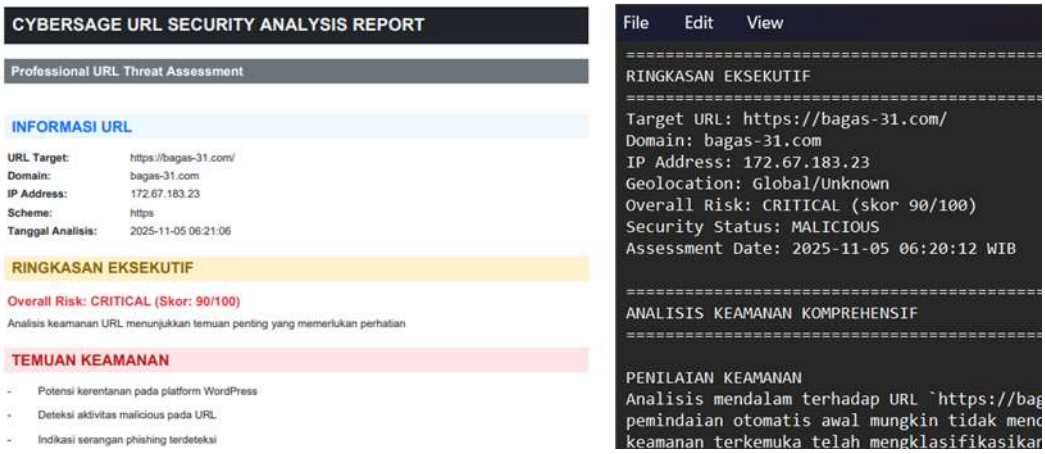

Figure 4. The PDF and TXT reports, both showing a 'CRITICAL' risk status for the URL.

Second, for the IP address analysis, the system processed a benign public IP. The resulting report (Figures 5-7) accurately assessed the IP as "Low Risk" (0/100). This result validates the system's operational utility in quickly filtering benign traffic, effectively reducing the time analysts spend on false positives .

Figure 5. Checking the IP address 139.255.79.117.

Figure 6. The bot provides the IP analysis reports.

Figure 7. The PDF and TXT reports, both showing a 'Low' risk status for the IP address.

The most critical findings stem from the file analysis workflow using a known Android malware sample (*undangan_rapat_pokja.apk*). The system successfully triggered the orchestration of VirusTotal hash lookups and MobSF static analysis. The AI-generated executive summary (Figures 8-10) demonstrated the system's primary novelty: cognitive synthesis. Instead of merely reporting raw statistics (e.g., "detected by 22/67 vendors"), the Google Gemini engine analyzed the technical MobSF output—specifically the request for sensitive permissions like RECEIVE_SMS. It synthesized this with reputation data to correctly conclude that the file was a high-risk Spyware/Trojan designed specifically to steal SMS messages .

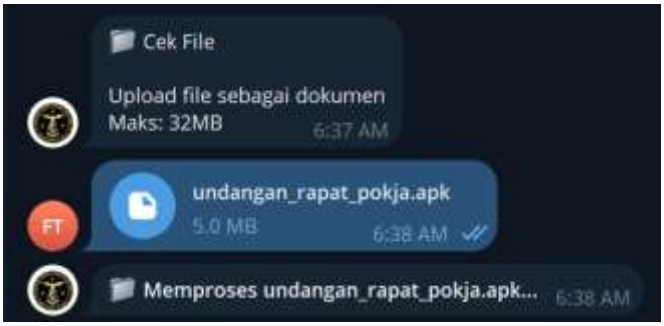

Figure 8. Uploading the undangan_rapat_pokja.apk file for processing.

Figure 9. The bot provides analysis results as .txt and .pdf report files.

Figure 10. The analysis report, showing its 'CRITICAL' status and detailed Indicators of Compromise (IOCs).

This capability confirms the research hypothesis. CyberGuardianBot moves beyond simple automation to act as a cognitive assistant, directly solving the "cognitive overload" problem in manual

analysis (Prieto & Blakely, 2024). The system aligns with contemporary frameworks that advocate for using AI to handle general threat situations, freeing human analysts for strategic decision-making (Gupta et al., 2023; Sai et al., 2024). Furthermore, the architecture addresses the risks of GenAI hallucinations and trust erosion (Dua & Patel, 2024; Ishtaiwi et al., 2025). By constraining the model to synthesize and recombine verified API data rather than generating net-new facts, the system operates within the AI's core strength (Sternberg, 2024). This data-grounded approach builds the necessary performance expectancy and user trust required for AI adoption in high-risk SOC environments (Ali et al., 2024; Shaila Rana & Ronda Chicone, 2025; Zhang & Yu, 2025).

Conclusions

This research successfully designed, validated, and implemented CyberGuardianBot, an intelligent analyst assistant that resolves the "Swivel Chair Analysis" operational bottleneck in Security Operation Center (SOC) environments. The primary scientific contribution of this work lies in the empirical validation of cognitive synthesis; unlike prior theoretical models or simple look-up tools, this system effectively integrates a Generative AI "super brain" (Google Gemini) to interpret and recombine asynchronous multi-source intelligence into actionable, hallucination-free insights. Practically, the system enhances incident response efficiency by automating the triage of URLs, IPs, and files, and standardizing reporting via chat and formal documents (PDF/TXT), thereby significantly reducing the cognitive load on analysts. However, the study bears limitations, primarily its dependence on external public API quotas which constrain high-volume throughput, and the lack of direct integration with internal network logs. Consequently, future research should focus on integrating the bot with SIEM platforms for holistic alert correlation, implementing a local database cache to optimize API usage, and expanding the analysis capabilities to support a broader range of threat indicators.

Reference

Al Zaidy, A. (2024). The Impact of Generative AI on Student Engagement and Ethics in Higher Education. *Journal of Information Technology, Cybersecurity, and Artificial Intelligence*, *1*(1), 30–38. https://doi.org/10.70715/jitcai.2024.v1.i1.004

Ali, M. S. M., Wasel, K. Z. A., & Abdelhamid, A. M. M. (2024). Generative AI and Media Content Creation: Investigating the Factors Shaping User Acceptance in the Arab Gulf States. *Journalism and Media*, *5*(4), 1624–1645. https://doi.org/10.3390/journalmedia5040101

Alqahtani, H., & Kumar, G. (2025). A comprehensive review of generative AI techniques and their impact on cybersecurity. *Soft Computing*, *29*(13), 4945–4982. https://doi.org/10.1007/s00500-025-10702-z

Dhoni, P., & Kumar, R. (2023). *Synergizing Generative AI and Cybersecurity: Roles of Generative AI Entities, Companies, Agencies, and Government in Enhancing Cybersecurity*. https://doi.org/10.36227/techrxiv.23968809.v1

Dua, I. K., & Patel, P. G. (2024). Software Optimization for Generative AI. In I. K. Dua & P. G. Patel (Eds.), *Optimizing Generative AI Workloads for Sustainability: Balancing Performance and Environmental Impact in Generative AI* (pp. 85–122). Apress. https://doi.org/10.1007/979-8-8688-0917-0_4

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. *IEEE Access*, *11*, 80218–80245. https://doi.org/10.1109/ACCESS.2023.3300381

Ishtaiwi, A., al-Qerem, A., Aldweesh, A., & Alkasassbeh, M. (2025). *A Framework for Addressing Cybersecurity Risks in the Metaverse Safeguarding Against Generative AI Threats* (pp. 381–400). https://doi.org/10.4018/979-8-3373-0832-6.ch016

Kazato, Y., Nakagawa, Y., & Nakatani, Y. (2020). Improving Maliciousness Estimation of Indicator of Compromise Using Graph Convolutional Networks. *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*, 1–7. https://doi.org/10.1109/CCNC46108.2020.9045113

Khai, C. D., & Juremi, J. (2023). BEsafe - Validating URLs and Domains with the aid of Indicator of Compromise. *2023 15th International Conference on Developments in ESystems Engineering (DeSE)*, 309–313. https://doi.org/10.1109/DeSE58274.2023.10099965

Kim, G., Lee, J., Kang, M., Go, W., & Hou, J.-U. (2024). *Highlights Enhancing Optical Character Recognition Performance in the Cybersecurity Domain for Indicator of Compromise Analysis Enhancing Optical Character Recognition*

Performance in the Cybersecurity Domain for Indicator of Compromise Analysis. https://github.com/GangsuKim/CAP

Krishnamurthy, O., & Vemulapalli, G. (2025). Advancing Sustainable Cybersecurity: Exploring Trends and Overcoming Challenges with Generative AI. In P. Whig, N. Silva, A. A. Elngar, N. Aneja, & P. Sharma (Eds.), *Sustainable Development through Machine Learning, AI and IoT* (pp. 16–25). Springer Nature Switzerland.

Nadella, G. S., Addula, S. R., Yadulla, A. R., Sajja, G. S., Meesala, M., Maturi, M. H., Meduri, K., & Gonaygunta, H. (2025). Generative AI-Enhanced Cybersecurity Framework for Enterprise Data Privacy Management. *Computers*, *14*(2), 55. https://doi.org/10.3390/computers14020055

Oh, S., & Shon, T. (2023). Cybersecurity Issues in Generative AI. *2023 International Conference on Platform Technology and Service (PlatCon)*, 97–100. https://doi.org/10.1109/PlatCon60102.2023.10255179

Peci, F., Hamiti, E., & Khan, I. (2025). Agentic AI with Chatops for Large Scale Network Operations. *2025 IEEE Conference on Artificial Intelligence (CAI)*, 1617–1626. https://doi.org/10.1109/CAI64502.2025.00242

Prieto, I., & Blakely, B. (2024). Proposed Uses of Generative AI in a Cybersecurity-Focused Soar Agent. *Proceedings of the AAAI Symposium Series*, *2*(1), 386–390. https://doi.org/10.1609/aaaiss.v2i1.27704

Rodger, D., Mann, S. P., Earp, B., Savulescu, J., Bobier, C., & Blackshaw, B. P. (2025). Generative AI in healthcare education: How AI literacy gaps could compromise learning and patient safety. *Nurse Education in Practice*, *87*, 104461. https://doi.org/10.1016/j.nepr.2025.104461

Sai, S., Yashvardhan, U., Chamola, V., & Sikdar, B. (2024). Generative AI for Cyber Security: Analyzing the Potential of ChatGPT, DALL-E, and Other Models for Enhancing the Security Space. *IEEE Access*, *12*, 53497–53516. https://doi.org/10.1109/ACCESS.2024.3385107

Salamun, M. A., Muttaqin, F. Z., & Rosyid, N. R. (2023). *Design and implementation of honeypot indicator of compromise (IoC) profiling using malware information sharing platform (MISP)*. 110003. https://doi.org/10.1063/5.0164216

Shaila Rana, & Ronda Chicone. (2025). Generative AI in Cybersecurity. In *Generative AI Security* (pp. 1–24). https://doi.org/https://doi.org/10.1002/9781394368532.ch1

Sternberg, R. J. (2024). Do Not Worry That Generative AI May Compromise Human Creativity or Intelligence in the Future: It Already Has. *Journal of Intelligence*, *12*(7), 69. https://doi.org/10.3390/jintelligence12070069

Teo, Z. L., Quek, C. W. N., Wong, J. L. Y., & Ting, D. S. W. (2024). Cybersecurity in the generative artificial intelligence era. *Asia-Pacific Journal of Ophthalmology*, *13*(4), 100091. https://doi.org/10.1016/J.APJO.2024.100091

Tostes, B., Ventura, L., Lovat, E., Martins, M., & Menasché, D. (2023). Learning When to Say Goodbye: What Should be the Shelf Life of an Indicator of Compromise? *2023 IEEE International Conference on Cyber Security and Resilience (CSR)*, 503–510. https://doi.org/10.1109/CSR57506.2023.10224937

Wang, S. K., Ma, S. P., Lai, G. H., & Chao, C. H. (2024). ChatOps for microservice systems: A low-code approach using service composition and large language models. *Future Generation Computer Systems*, *161*, 518–530. https://doi.org/10.1016/J.FUTURE.2024.07.029

Zhang, G., & Yu, T. (2025). Association between Generative AI self-efficacy and Generative AI acceptance: The mediating role of Generative AI trust and the moderating role of Generative AI risk perception. *Acta Psychologica*, *261*, 105791. https://doi.org/10.1016/J.ACTPSY.2025.105791

Zheng, D. Y., Tong, K. K. S., Lim, M. T. A., Chan, W. J., & Goh, W. (2023). AfterImage: Evading Traditional Indicator of Compromise (IOC) Blocking. *2023 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, 1–6. https://doi.org/10.1109/SOLI60636.2023.10425081