

JOURNAL OF DIGITAL LAW AND INNOVATION (JDLI)

LAW • TECHNOLOGY • INNOVATION • JUSTICE

 OPEN
ACCESS

ISSN
3093-9364



OUR FIELDS

- Digital law
- Cybersecurity law
- Artificial intelligence and robotics law
- Intellectual property law
- LegalTech and legal innovations
- E-government and digital public services
- Legal aspects of AgroTech and GreenTech
- Smart city and digital governance



OUR INDEXING

 Google Scholar

 Zenodo

 DOI

 OpenAIRE

 Open Access



AI

UNAUTHORIZED
ACCESS

SYSTEM
BREACH



CYBER ATTACK



DIGITAL LAW



 <https://unireview.uz/index.php/jdli>



Published by
UniLex Global



VOLUME 2

ISSUE 1



AUGUST 2026



“UNILEX GLOBAL” LLC

**Raqamli huquq va innovatsiya
jurnali
Journal of Digital Law and
Innovation**

**August, 2026
VOLUME 2, ISSUE 1**



EDITORIAL TEAM

Editor-in-Chief, Tukhtashev Hikmatilla Ismatillaevich - Head of the Department of Legal Sciences, Doctor of Philosophy in Legal Sciences (PhD), TIAME National Research University, Tashkent, Republic of Uzbekistan.

1. **Editorial Board Member Borotov Mirodiljon Xomudjonovich** — Doctor of Legal Sciences, Professor, Head of the Department of Civil Law, International Law and International Organizations, Institute of State and Law of the Academy of Sciences of the Republic of Uzbekistan, Tashkent, Republic of Uzbekistan.
2. **Editorial Board Member Maripova Sevarxon Aribjanovna** — Doctor of Philosophy in Legal Sciences (PhD), Associate Professor, Department of Social Sciences, Cyber University State University, Tashkent, Republic of Uzbekistan.
3. **Editorial Board Member Kaharova Munira Maxamajanovna** — Doctor of Political Sciences (DSc), Researcher, Institute of Criminology of the Republic of Uzbekistan, Tashkent, Republic of Uzbekistan.
4. **Editorial Board Member Yuldashev Mutalib Ibrohimovich** — Doctor of Economic Sciences, Professor, Department of Economics, Tashkent University of Science and Technology, Tashkent, Republic of Uzbekistan.
5. **Editorial Board Member Bafaev Shukhrat Gafurovich** — Doctor of Legal Sciences, Head of the Tashkent City Branch, Center for the Analysis of Democratic Processes, Tashkent, Republic of Uzbekistan.
6. **Editorial Board Member Khodjaeva Aziza Bakhtiyarovna** — Doctor of Philosophy in Legal Sciences (PhD), “Academy of Labor and Social Relations” under the Federation of Trade Unions of the Republic of Uzbekistan, Tashkent, Republic of Uzbekistan.
7. **Editorial Board Member Khudoyberdieva Viloyat Jabbrovanovna** — Candidate of Legal Sciences (PhD), Olmazor District Department of Public Education, Tashkent City, Republic of Uzbekistan.
8. **Editorial Board Member Boboumarova Munira Kurbonovna** — Doctor of Philosophy in Legal Sciences (PhD), Department for Organization of Investigative Activities, Investigation Division, Tashkent City Internal Affairs Department, Tashkent, Republic of Uzbekistan.
9. **Editorial Board Member Usmonov Vosid Muhammadievich** — Doctor of Philosophy in Legal Sciences (PhD), Diplomatic University, Tashkent, Republic of Uzbekistan.
10. **Editorial Board Member Mirzayeva Malika Baxadirovna** — Doctoral Researcher, PhD, Associate Professor, Department of Digital Technologies and Artificial Intelligence, Tashkent Institute of Irrigation and Agricultural Mechanization Engineers – National Research University, Tashkent, Republic of Uzbekistan.
11. **Editorial Board Member Allayorov Tulqin Ruziqulovich** — Doctor of Philosophy in Philology (PhD), Associate Professor, Nizami Named Tashkent State Pedagogical University, Tashkent, Republic of Uzbekistan.
12. **Editorial Board Member Tosheva Maftuna Anvar kizi** — Doctor of Philosophy in Legal Sciences (PhD), Tashkent Institute of Irrigation and Agricultural Mechanization Engineers – National Research University, Tashkent, Republic of Uzbekistan.
13. **Editorial Board Member Askarov Jaloliddin Toychievich** — Doctor of Philosophy in Legal Sciences (PhD), Deputy Head, Department of Business Law, Tashkent State University of Law, Tashkent, Republic of Uzbekistan.
14. **Editorial Board Member Abdixakimov Islombek Bahodir o'g'li** — Master of Law (LL.M), Head of the Artificial Intelligence and Legal Tech Laboratory, Tashkent State University of Law, Tashkent, Republic of Uzbekistan.
15. **Editorial Board Member Kalandarov Aminjon Amondullayevich** — Doctor of Philosophy in Legal Sciences (PhD), Acting Associate Professor, Department of Business Law, Tashkent State University of Law, Tashkent, Republic of Uzbekistan.
16. **Editorial Board Member Hakimov Komil Bakhtiyarovich** — Doctor of Legal Sciences (DSc), Professor, Head of the Department of Criminal Law, Criminology and Anti-Corruption, Tashkent State University of Law, Tashkent, Republic of Uzbekistan.

REDUCING SOC ANALYST ALERT FATIGUE THROUGH LLM-ASSISTED TRIAGE: A PROVIDER-AGNOSTIC SUMMARIZATION PIPELINE FOR WAZUH

To‘ychiboyev Shahriyor

“Cyber University” davlat universiteti 1-kurs talabasi

Samarov Sherzod Hamrayevich

“Cyber University” davlat universiteti

“Huquq va biznes” fakulteti uslubchisi

ARTICLE INFO

Keywords:

security operations centre,
alert fatigue,
large language model,
Wazuh,
triage automation,
prompt injection,
structured output,
human-in-the-loop

ABSTRACT

The volume of alerts processed by Security Operations Centres (SOCs) often exceeds analysts' capacity for sustained attention. Even low-priority notifications require analysts to examine raw logs, interpret rule metadata, and make an initial determination as to whether an alert constitutes a true positive or a false positive. This contributes to alert fatigue and increases the risk that significant security events may be overlooked. This article presents an open-source, provider-agnostic system that uses a large language model to perform the initial triage of alerts generated by Wazuh 4.9.2. The system combines a severity-prioritized summarization service with a read-only conversational assistant operating within the analyst's existing access permissions. Importantly, the LLM component is isolated from the primary alerting pipeline; therefore, a model or provider failure does not affect Wazuh's ability to detect and record threats. Engineering verification demonstrated that the system successfully passed all quality-control tests. However, because the accuracy of its conclusions was not evaluated against a human-labelled ground-truth dataset, the results confirm the system's technical and operational feasibility rather than its absolute reliability in making final security decisions.

СНИЖЕНИЕ УСТАЛОСТИ СОС-АНАЛИТИКОВ ОТ ПОТОКА ОПОВЕЩЕНИЙ С ПОМОЩЬЮ LLM-АССИСТИРОВАННОЙ ПЕРВИЧНОЙ СОРТИРОВКИ: НЕЗАВИСИМЫЙ ОТ ПРОВАЙДЕРА КОНВЕЙЕР СУММАРИЗАЦИИ ДЛЯ WAZUH

ИНФОРМАЦИЯ СТАТЬИ

Ключевые слова:

центр управления
информационной
безопасностью, усталость
от предупреждений,
большая языковая модель,
Wazuh,
автоматизация триажа,
prompt-инъекция,
структурированный вывод,
контроль человеком

АННОТАЦИЯ

Центры управления информационной безопасностью обрабатывают объёмы предупреждений, которые зачастую превышают устойчивые возможности внимания аналитиков. Даже малозначимые уведомления требуют изучения исходных журналов, интерпретации метаданных правил и вынесения предварительного суждения об истинном или ложном срабатывании. В исследовании представлена открытая, независимая от поставщика система, использующая большую языковую модель для первичного анализа предупреждений Wazuh 4.9.2. Система объединяет сервис приоритетного суммирования по уровню критичности и диалогового помощника с доступом только для чтения, действующего в пределах существующих прав аналитика. Компонент LLM изолирован от штатного контура оповещения, поэтому отказ модели или провайдера не задерживает обнаружение событий. Инженерные проверки подтвердили прохождение всех предусмотренных тестов качества и безопасности. Вместе с тем точность вердиктов не оценивалась на размеченном экспертами эталонном наборе данных; следовательно, результаты подтверждают техническую осуществимость и эксплуатационную интеграцию, но не автономную надёжность принятия решений.

WAZUH UCHUN PROVAYDERGA BOG'LIQ BO'LMAGAN UMUMLASHTIRISH KONVEYERI ASOSIDA LLM YORDAMIDA HODISALARNI SARALASH ORQALI SOC TAHLILCHILARINING OGOHLANTIRISHLARDAN CHARCHASHINI KAMAYTIRISH

MAQOLA MA'LUMOTI

Kalit so'zlar:

xavfsizlik operatsiyalari
markazi,
ogohlantirishlardan
charchash,
katta til modeli,
Wazuh,
avtomatlashtirilgan saralash,
prompt-inyeksiya,
tuzilmaviy natija,
inson nazorati

ANNOTATSIYA

Xavfsizlik operatsiyalari markazlarida qayta ishlanadigan ogohlantirishlar hajmi ko'pincha tahlilchilarning barqaror diqqat imkoniyatidan oshib ketadi. Hatto past ahamiyatli xabarlar ham xom loglarni o'rganish, qoida metama'lumotlarini talqin qilish va dastlabki haqiqiy yoki yolg'on ijobiy xulosa chiqarishni talab etadi. Bu esa ogohlantirishlardan charchashni kuchaytirib, muhim hodisalarning e'tibordan chetda qolish xavfini oshiradi. Mazkur tadqiqot Wazuh 4.9.2 tizimida shakllantirilgan ogohlantirishlarni dastlabki saralashda katta til modelidan foydalanadigan ochiq manbali va provayderga bog'liq bo'lmagan tizimni taqdim etadi. Tizim jiddiylik darajasiga ko'ra ustuvorlashtirilgan umumlashtirish xizmati hamda tahlilchining mavjud kirish huquqlari doirasida ishlaydigan faqat o'qishga mo'ljallangan suhbat yordamchisini birlashtiradi. Muhimi, LLM komponenti asosiy ogohlantirish jarayonidan ajratilgan bo'lib, model yoki provayderdagi nosozlik Wazuh tomonidan tahdidlarni aniqlash va qayd etishga ta'sir qilmaydi. Muhandislik tekshiruvlari barcha sifat nazorati sinovlaridan muvaffaqiyatli o'tilganini ko'rsatdi. Biroq xulosalarning aniqligi inson tomonidan belgilangan etalon ma'lumotlar bilan baholanmaganligi sababli natijalar tizimning texnik va operatsion imkoniyatlarini tasdiqlaydi, uning yakuniy qaror qabul qilishdagi mutlaq ishonchligini emas.

INTRODUCTION

Security operations centres (SOCs) function as the operational core of organizational cyber defence. Their analysts continuously review endpoint, network, identity and application telemetry in order to identify malicious activity, determine its scope and coordinate an appropriate response. The usefulness of this model, however, depends on the capacity to distinguish a relatively small number of consequential incidents from a much larger volume of benign, duplicated or context-poor alerts. Alert fatigue is therefore not merely a matter of employee discomfort; it is an operational risk capable of delaying investigation, increasing inconsistency and concealing genuine threats within routine noise [1].

Traditional security information and event management workflows rely heavily on rule-

based detection. Such rules provide reproducibility and transparent triggering conditions, but their outputs are frequently optimized for machine processing rather than human comprehension. Analysts must reconstruct context from rule identifiers, event fields, endpoint details, MITRE ATT&CK mappings and unstructured log fragments. The repeated cognitive cost of this preliminary interpretation can become significant when thousands of events are generated during a single shift.

Research on automated provenance triage has demonstrated that alert reduction is possible when causal and contextual relationships are reconstructed before events reach human investigators [2]. More recently, large language models have introduced a complementary capability: they can translate heterogeneous log records into concise natural-language explanations and can support flexible questions over security data [3, 4]. Multi-agent frameworks such as CORTEX explore collaborative reasoning for high-stakes alert triage [5], while LogPrompt demonstrates that prompt engineering can improve zero-shot and interpretable log analysis [6]. Field research has also begun to examine how LLM tools can be introduced into a live SOC without disrupting established operations [7].

These capabilities do not remove the fundamental risks of applying generative models to cybersecurity. LLMs may produce plausible but incorrect interpretations, follow malicious instructions embedded in log content, disclose sensitive telemetry to an external provider or return inconsistent output formats. OWASP identifies prompt injection, sensitive-information disclosure and excessive agency among the principal risks of LLM-enabled applications [8]. A defensible design must therefore limit model authority, separate untrusted evidence from instructions, constrain outputs and preserve human responsibility for consequential decisions.

Against this background, the present study asks whether first-pass Wazuh alert triage can be delegated to an LLM without placing the model in the critical detection path or granting it authority to modify the security environment. The study evaluates an open-source architecture consisting of two bounded components: a severity-prioritized summarization pipeline and a read-only conversational assistant. Its contribution is principally architectural and operational. It demonstrates a provider-agnostic pattern for integrating generative reasoning with a production SIEM while preserving native alert delivery and existing access controls.

The study pursues four objectives: to define a fail-safe LLM integration pattern for Wazuh; to describe controls for structured output, prompt injection and least privilege; to

report reproducible engineering verification results; and to distinguish demonstrated technical feasibility from unmeasured analytical accuracy. This distinction is essential because passing software tests cannot substitute for validation against expert-labelled security decisions.

METHODOLOGY

The research adopted a design-science and case-study methodology. The artifact was developed as an operational software system and evaluated within a real Wazuh 4.9.2 deployment protected by Transport Layer Security and authentication and connected to genuine Windows and Linux agents [9]. The evaluation focused on architectural safety, functional correctness, persistence, encryption and tool authorization. It did not treat LLM-generated verdicts as verified detection labels.

The analytical unit was an individual Wazuh alert. The pipeline extracts the rule description and severity, agent identity, timestamp, MITRE ATT&CK mapping and original log record. Alerts are then placed in a priority queue so that critical and high-severity events are processed before lower-value noise. Rotation-safe log tailing preserves continuity when the alerts file is rotated. The service is independent of Wazuh's native alerting subsystem and writes its summaries as supplementary records.

The prompt design treats the raw log block as untrusted evidence rather than executable instruction. The model is required to produce four structured fields: what happened, why the event matters, a suggested analyst action and a constrained verdict. The verdict is limited to three states—likely true positive, likely false positive and needs review—so that uncertain cases remain visible to a human analyst. Provider abstraction permits use of Gemini Flash or an OpenAI-compatible endpoint without changing the core pipeline.

A separate dashboard plugin implements the conversational component. It exposes 29 read-only tools covering findings, vulnerability data, file-integrity monitoring, MITRE mappings and agent inventory. No mutating or remediation tool is available. Queries execute under the identity of the authenticated dashboard user, use an allowlist of indices and enforce result limits. This design applies least privilege at both the tool and data layers.

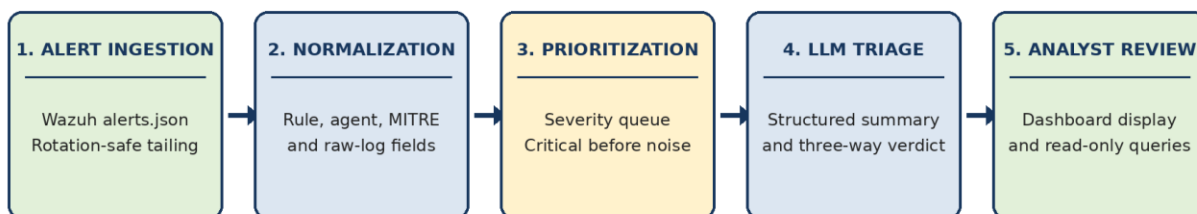
Evaluation was divided into two tiers. Tier 1 did not require a live stack and comprised static type checking, unit testing, linting and an adversarial guardrail corpus. Tier 2 used a live virtual machine to test alert delivery from ingestion to persisted summary, conversation storage and ownership isolation, encryption round-trips and the complete read-only tool catalogue. Operational dashboard figures were recorded as a descriptive snapshot rather than

as a controlled experiment.

Validity was protected by separating directly observed outcomes from claims requiring additional evidence. Test passage established that the evaluated build behaved as specified; it did not establish that the LLM's security judgments were correct. A human-labelled measurement set and a freeze-gate rubric were defined, but no completed accuracy or honesty score was available. Accordingly, no claim of reduced incident rate, improved mean time to respond or analyst-hours saved is made.

Figure 1. Provider-agnostic LLM-assisted SOC triage pipeline

PROVIDER-AGNOSTIC LLM-ASSISTED SOC TRIAGE PIPELINE



Native Wazuh alerting remains operational even when the model provider is unavailable.

Source: Developed by the author based on the implemented ThreatLens architecture and the evaluated Wazuh 4.9.2 deployment.

RESULTS

The evaluation produced two categories of results: an operational snapshot of the deployed pipeline and reproducible engineering verification outcomes. The dashboard snapshot showed that 966 alerts had been summarized and 4,442 remained pending. The monitored distribution comprised no critical alerts, 46 high-severity alerts, 130 medium-severity alerts and 5,232 low-severity alerts. Year-to-date model usage amounted to 300,041 tokens, including 146,758 tokens on the day of capture, with a measured cost of USD 0.0000 under the provider's free tier. The pipeline processed approximately 1,000 events per day in the observed configuration.

Sample outputs complied with the prescribed schema. A registry event associated with the OpenSSH agent was classified as a likely false positive, while an unquoted service path mapped to MITRE technique T1574.009 and a netsh firewall-rule event were escalated as needs-review. These examples demonstrate the system's ability to translate raw telemetry into

concise triage records. They do not independently prove that the verdicts were correct.

Table 1. Operational snapshot of the evaluated deployment

Indicator	Observed value	Interpretation
Alerts summarized	966	Completed LLM summaries available to analysts
Alerts pending	4,442	Backlog awaiting model processing
Severity distribution	0 critical; 46 high; 130 medium; 5,232 low	Workload dominated by low-severity events
Approximate throughput	≈1,000 events per day	Observed with Gemini Flash in the evaluated configuration
Year-to-date tokens	300,041	Cumulative provider usage at capture
Measured provider cost	USD 0.0000	Free-tier observation; not a general cost estimate

***Source:** Compiled by the author from the dated dashboard snapshot of the evaluated deployment.*

Engineering-quality gates passed completely on the build verified on 26 July 2026. Type checking returned zero errors; all 362 unit tests and all 28 adversarial guardrail-lint cases passed; and code linting was clean. Live-stack verification also passed all 44 pipeline checks, all nine persistence and owner-isolation checks, all four encryption round-trips and all 29 catalogue-tool exercises. Table 2 separates each result from the conclusion that can legitimately be drawn from it.

Table 2. Engineering verification results and evidentiary meaning

Verification area	Result	What the result establishes	What it does not establish
Static type checking	0 errors	Internal type consistency of the evaluated build	Correctness of security verdicts
Unit tests	362/362 passed	Expected component behaviour under tested conditions	Performance on unseen attacks
Guardrail lint	28/28 passed	Compliance with the tested adversarial prompt rules	Immunity from all prompt-injection techniques
Pipeline checks	44/44 passed	Alert-to-summary delivery in the live environment	Reduction in analyst workload
Persistence checks	9/9 passed	Conversation storage and owner isolation	Security of every deployment configuration
Encryption checks	4/4 passed	Correct tested encryption/decryption round-trips	End-to-end elimination of disclosure risk
Read-only tools	29/29 exercised	Functional coverage of the tool catalogue	Accuracy of model interpretation of tool results

Source: Compiled by the author from the CI and live-stack verification records dated 26 July 2026.

No completed human-judged accuracy or honesty score was recorded. The defined freeze gate required at least 85% correct classifications and zero confidently wrong answers, but only the rubric and threshold were available. This unresolved measurement gap is the most important empirical result because it limits the scope of all effectiveness claims. The study therefore establishes a functioning and bounded integration, not autonomous alert adjudication.

DISCUSSION

Architectural contribution. The principal contribution is the separation of generative analysis from the SIEM's native detection path. Many security copilots are tightly coupled to

proprietary cloud backends. In contrast, the evaluated system is open source, deployable on premises and capable of switching providers through a common abstraction. If the model, network connection or external provider becomes unavailable, Wazuh continues to detect and record events. The failure affects only the convenience layer that produces summaries.

Operational relevance. The severity-prioritized queue addresses a practical weakness of simple first-in, first-out summarization. A large volume of low-severity events should not prevent a newly generated critical alert from receiving prompt analysis. The design therefore improves the ordering of analyst support without modifying the underlying Wazuh severity determination. The backlog shown in the operational snapshot also demonstrates that prioritization is necessary when model throughput is lower than alert generation.

Human–AI allocation of responsibility. The three-way verdict deliberately reserves an explicit needs-review category. This is preferable to a forced binary classification because ambiguous evidence is common in SOC work. The model’s role is to compress and organize evidence, not to close incidents. Final determination should remain with a qualified analyst who can review endpoint context, correlated events and organizational baselines that may not be present in the model prompt.

Structured output and prompt injection. A schema improves machine validation, dashboard rendering and downstream consistency. It cannot guarantee semantic truth. Likewise, instructing the model to treat logs as data and validating prompts against an adversarial corpus reduces known injection pathways but does not eliminate the possibility that specially crafted telemetry will influence model behaviour. Defence should therefore be layered through content delimitation, output validation, least privilege, provider isolation and monitoring for abnormal responses [8].

Access control and excessive agency. The conversational assistant inherits the authenticated user’s permissions and exposes only read-only tools. This creates a narrower blast radius than a security orchestration agent capable of disabling accounts, blocking network traffic or changing endpoint configuration. Tool allowlisting and result caps further reduce the risk of unrestricted data extraction. Nevertheless, read access to security telemetry remains sensitive and should be logged, monitored and periodically reviewed.

Privacy and provider dependence. The current implementation sends full log content to the configured model without field-level redaction. Logs may contain usernames, hostnames, internal addresses, command lines or other operationally sensitive data.

Deployment therefore requires a documented data-processing decision, appropriate contractual safeguards and a clear retention policy. A self-hosted model or pre-processing redaction layer would reduce external disclosure, although local deployment introduces its own maintenance and model-quality requirements.

Comparison with prior work. NoDoze demonstrates noise suppression through provenance relationships [2], while CORTEX explores multi-agent collaboration for high-stakes triage [5]. The present system adopts a more constrained objective: it translates and preliminarily categorizes existing Wazuh alerts without modifying their generation or authorizing remediation. Its value lies less in novel model reasoning than in operational containment, provider independence and explicit failure isolation.

Limitations. The study is based on one production deployment and a dated operational snapshot. It does not include a randomized analyst study, a human-labelled confusion matrix, precision and recall, inter-rater agreement, mean time to triage or a longitudinal measure of alert-volume reduction. Full-log prompts are not yet redacted, historical backfill is absent and two guardrail gaps remain open. Free-tier cost cannot be generalized to commercial workloads, and approximately 1,000 events per day should not be interpreted as a universal throughput benchmark.

Future research. The next evaluation should freeze a representative, de-identified and human-labelled alert corpus and report precision, recall, false-negative rate, calibration and confident-error frequency for each verdict class. A controlled analyst study should compare time-to-triage, correction rate and cognitive workload with and without LLM assistance. Engineering work should add configurable field-level redaction, resumable summaries, local-model support, model-output provenance, drift monitoring and compatibility with a unified Wazuh 5.x release.

CONCLUSION

This study presented an open-source, provider-agnostic framework for reducing the repetitive cognitive burden associated with first-pass Wazuh alert triage. The system combines a severity-aware summarization pipeline with a read-only conversational assistant and constrains each model response to a defined explanatory schema and three-way verdict. Its most important safety property is architectural: the model remains outside the native alerting path and therefore cannot prevent Wazuh from recording or delivering security events.

The evaluated build passed all reported static, unit, guardrail, pipeline, persistence,

encryption and tool-catalogue checks. The deployed instance had summarized 966 real alerts at the time of observation and demonstrated practical integration with production Windows and Linux telemetry. These findings establish software maturity and operational feasibility within the tested environment.

At the same time, engineering success must not be conflated with analytical accuracy. No completed human-labelled evaluation was available, and the study does not claim measured reductions in false positives, analyst-hours or incident-response time. Until such evidence is produced, LLM output should be treated as an auditable recommendation subject to human review rather than as an autonomous security decision.

The broader implication is that generative AI can support SOC operations when its authority is deliberately bounded. Reliable adoption requires failure isolation, least-privilege tools, access-control inheritance, structured outputs, prompt-injection safeguards, data minimization and continuous empirical validation. Future development should prioritize redaction, self-hosted model options, resumable processing and rigorous comparison against expert-labelled ground truth.

REFERENCES

1. Alahmadi, M., et al. (2025). Alert fatigue in security operations centres: Research challenges and opportunities. *ACM Computing Surveys*. <https://doi.org/10.1145/3723158>
2. Hassan, W. U., Guo, S., Li, D., Chen, Z., Jee, K., Li, Z., & Bates, A. (2019). NoDoze: Combatting threat alert fatigue with automated provenance triage. *Proceedings of the Network and Distributed System Security Symposium*. <https://doi.org/10.14722/ndss.2019.23349>
3. Habibzadeh, A., et al. (2025). Large language models for security operations centers: A comprehensive survey.
4. Hassanin, M., & Moustafa, N. (2024). A comprehensive overview of large language models for cyber defences: Opportunities and directions.
5. Wei, B., Tay, Y. S., Liu, H., Pan, J., Luo, K., Zhu, Z., & Jordan, C. (2025). CORTEX: Collaborative LLM agents for high-stakes alert triage.
6. Liu, Y., Tao, S., Meng, W., Yao, F., Zhao, X., & Yang, H. (2024). LogPrompt: Prompt engineering towards zero-shot and interpretable log analysis. *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering: Companion*, 364–365.
7. Hahn, F., Mamoon, M., Bardas, A. G., Collins, M., Dudek, J. L., Lende, D., & Ou, X. (2026). Non-disruptive disruption: An empirical experience of introducing LLMs in the SOC. *Workshop on Security Operations Center, co-located with NDSS*.

8. OWASP Foundation. (2025). OWASP Top 10 for Large Language Model Applications. <https://owasp.org/www-project-top-10-for-large-language-model-applications/>
9. Wazuh, Inc. (2026). Wazuh: The open-source security platform. <https://wazuh.com/>

