International Journal of Innovative Research in Engineering
Volume 7, Issue 6 (November - December 2026), PP: 53-60.
https://www.doi.org/10.59256/ijire.20260706006
www.theijire.com
ISSN No: 2582-8746

Forensiq - Intelligent Security Operations & Investigation Platform

Ayush Shankar Kumbhar[1], Sharayu Nilesh Patil[2], Sakshi Chandrakant Patil[3*], Sakshi Anil Morti[4], Soham Vikas Jadhav[5], Gayatri Umesh Jadhav[6]
[1, 2,* 3, 4, 5, 6] *Department of Computer Science & Engineering, D. Y. Patil College of Engineering & Technology, Kolhapur, Maharashtra, India.*
[*]*Corresponding Author.*

How to cite this paper:
Ayush Shankar Kumbhar[1], Sharayu Nilesh Patil[2], Sakshi Chandrakant Patil[3*], Sakshi Anil Morti[4], Soham Vikas Jadhav[5], Gayatri Umesh Jadhav[6], "Forensiq - Intelligent Security Operations & Investigation Platform", IJIRE-V7I6-53-60.

Copyright © 2026 by author(s) and Fifth Dimension Research Publication. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

Abstract: *Security Operations Centers process security alerts from many different systems. Analysts must match events gather evidence add data and evaluate risk before deciding. Forensiq is a platform that uses AI agents to help with these tasks. It works inside Splunk. Handles alert ingestion, event exploration, alert triage, threat-intelligence enrichment of indicators of compromise, mapping to MITRE ATT&CK correlating events calculating risk reasoning through investigations checking claims with facts and producing reports for analysts. A multi-agent design splits the work into tasks but the original data stays for analysts to check. Tests will look at how alerts are sorted how well ATT&CK mapping works how much threat data is added, how well evidence supports claims how many unsupported claims appear, how good event correlation is, how efficient investigations become, how good reports are and how the system scales. The final test will compare the Forensiq workflow to a baseline under the same data. Forensiq is meant to help analysts; the system gives reasoning support. Does not replace human checking and response decisions.*

Key *Words*: *security operations center, multi-agent systems, Splunk, MITRE ATT&CK, threat intelligence.*

I.INTRODUCTION

Security Operations Centers are tasked with watching security data all the time spotting suspicious activity looking into alerts and aiding incident response. Modern setups produce different kinds of evidence from SIEM platforms end points, login systems, network gear, applications and outside threat intelligence feeds. A good investigation needs more than finding alerts. Analysts have to gather the evidence set the context understand how attackers act and record why a conclusion is true.

Current incident-response guidance stresses putting incident response into risk management and making detection, response and recovery faster and better [5]. In work analysts often jump between several tools and manually link events before reaching a conclusion. This adds work and delays especially when many alerts must be handled in a time.

New progress in Large Language Models and agentic systems gives chances to automate log summaries, alert triage, context reasoning, threat-intelligence enrichment and report writing. Cognitive SOC looks at evidence-backed security stories using agents [1] while CORTEX studies collaborative agents for important alert triage [2]. AgentSOC offers an agentic plan for automating security operations [4]. These studies show that having agents specialize and use tools can make the analysis structure better.

A main problem is trust. An LLM can produce an explanation that sounds right but has no proof. In a security investigation such a statement should not be treated automatically as fact. A recent review of LLM-based cyber defense points out reliability tool-augmented reasoning, multi-agent coordination, safety and governance, as worries [3].

Forensiq solves this by dividing duties. LLM agents handle interpretation orchestration, context reasoning and hypothesis building while fixed components get data validate claims, calculate risk, control access and keep a record of origin. The key idea is that important facts should only be shown as verified if source evidence can be shown. Analysts still decide the checks and responses.

Objectives and problem statement

Security analysts must look at alerts while linking evidence from many places. Manual work raises the load and time needed while free LLM automation can bring unsupported or made-up claims. The research question Forensiq tackles is: How can a multi-agent LLM system run security-alert investigation while keeping the results grounded in real source evidence and showing the outcome through a human-in-the-loop process?

The objectives are: 1) design a multi-agent architecture that can ingest security alerts from Splunk SIEM; 2) automate the collection of contextual evidence enrichment of threat intelligence, correlation of events and mapping to MITRE ATT&CK; 3) implement a deterministic verification layer for findings generated by large language models; 4) produce structured investigation reports that are backed by evidence include risk information and recommendations; 5) provide a responsive dashboard that allows analysts to review, validate, give feedback and view investigation history; and 6) define a reproducible evaluation methodology that measures investigation quality, evidence coverage, latency and analyst workload.

II. LITERATURE SURVEY AND RESEARCH GAP

The proposed work is motivated by the gap between the power of artificial intelligence systems and the day-to-day needs of a reliable security operations centre workflow. Four key challenges are tackled. First security findings must be linked to evidence so that an analyst can trace a conclusion back to the event its timestamp, the indicator of compromise or the source record that supports it. Second automation of investigations must work with security infrastructure instead of just isolated text samples. Third investigation tasks can be split into activities such as triage, evidence collection, enrichment of indicators of compromise, mapping to ATT&CK building a timeline and creating a report. Fourth response actions that have impact must involve human oversight.

The literature shows that multi-agent reasoning and evidence-aware security analysis are useful [1] [2] [4]. Forensiq however concentrates on combining investigation agents with a deterministic claim-verification step and an audit trail that analysts can view. Forensiq therefore offers a framework that grounds SOC investigations in evidence rather than giving a fully autonomous replacement for security staff.

A. Related Work

Cognitive SOC offers an evidence-backed multi-agent approach where specialised agents work together to create security stories that include evidence and citation chains [1]. This work shows how important it is to keep links between conclusions and the evidence that backs them. Forensiq builds on this idea for a Splunk-based investigation workflow. Puts a deterministic claim-verification stage ahead of the final report to analysts.

CORTEX studies large language model agents for high-stakes alert triage and separates behaviour analysis, evidence gathering and reasoning [2]. This backs the Forensiq choice to use agents. Forensiq adds a validation step that checks factual claims produced by the agents against the source records that were retrieved.

AgentSOC proposes an agentic process for security operations that covers normalisation, contextual enrichment, reasoning, validation and risk-based planning [4]. Forensiq also separates processing stages. Puts a stronger focus on where evidence comes from deterministic verification and the need for human approval on decisions that have big consequences.

The systematic review by Alqahtani and Ahuja points out that autonomy, tool-augmented reasoning, coordination among agents and governance are key concerns for large language model based cyber defence [3]. These concerns drive Forensiq to include controlled tools, storage of evidence audit records and a human-in-the-loop review process.

MITRE ATT&CK supplies a knowledge base that represents the tactics and techniques used by adversaries [6]. Forensiq uses ATT&CK mapping to give a shared vocabulary of behaviour while still keeping references to evidence and a confidence score that analysts can examine.

Work	Approach and key distinction
Cognitive SOC [1]	Multi-agent LLM; evidence-backed narratives. Forensiq adds deterministic claim verification within a SIEM-oriented workflow.
Work	Approach and key distinction
CORTEX [2]	agents for high-stakes alert triage. Forensiq adds source-evidence validation before reporting.
AgentSOC [4]	agents for normalisation, reasoning, validation and risk planning. Forensiq emphasises provenance and analyst approval.
Systematic review [3]	Identifies autonomy, tool use, coordination, safety and governance concerns. Forensiq translates these concerns into a concrete verification architecture.

Table I. Comparison of Related Approaches

III. METHODOLOGY

Forensiq follows a staged investigation pipeline. The main design idea is to use language models for tasks that need semantic interpretation and flexible reasoning and to use deterministic software for retrieving data validating it scoring it controlling access and keeping a trace of evidence. Figure 1 shows the proposed architecture.

The data-source layer holds Splunk SIEM, network devices, endpoints and firewall logs. The ingestion layer pulls alert and event information, from the Splunk API. Puts normalised data into the investigation workflow. PostgreSQL stores data permanently while Redis supplies a cache and temporary state. A vector store can aid retrieval and retrieval-augmented generation when the investigation context is indexed.

The AI Agent Layer is orchestrated using LangGraph and Python services. The AI Agent Layer design includes an Orchestrator Agent, Triage Agent, IOC Enrichment Agent, Correlation Agent MITRE Mapping Agent, Timeline Agent, Risk Assessment Agent and Investigation Reasoning Agent.

Each agent in the AI Agent Layer has a responsibility and each agent in the AI Agent Layer communicates through a shared investigation state.

The Investigation and Evidence Layer maintains investigation records, evidence metadata, threat-intelligence information and ATT&CK context.

The Investigation and Evidence Layer integrates services such as VirusTotal and AbuseIPDB to provide IOC enrichment while MITRE ATT&CK provides a standardized behavior taxonomy.

The Investigation and Evidence Layer presents investigation reports, evidence, alerts, recommendations and analyst feedback.

Technology	Role
Python / Django REST / FastAPI	Backend services, REST integration, investigation APIs
LangChain / LangGraph	Multi-agent orchestration and stateful workflow
PostgreSQL	Persistent investigation, evidence, and audit data
Redis	Cache and transient investigation state
Vector Store / pgvector	Semantic retrieval and RAG support
Splunk Enterprise	SIEM alerts and security-event evidence
VirusTotal / AbuseIPDB / OTX	IOC enrichment and reputation context
MITRE ATT&CK	Standardized adversary behavior mapping
Next.js / React / TypeScript	SOC analyst dashboard
Ollama / API Models	LLM inference and model integration

Table II. proposed technology stack and responsibilities.

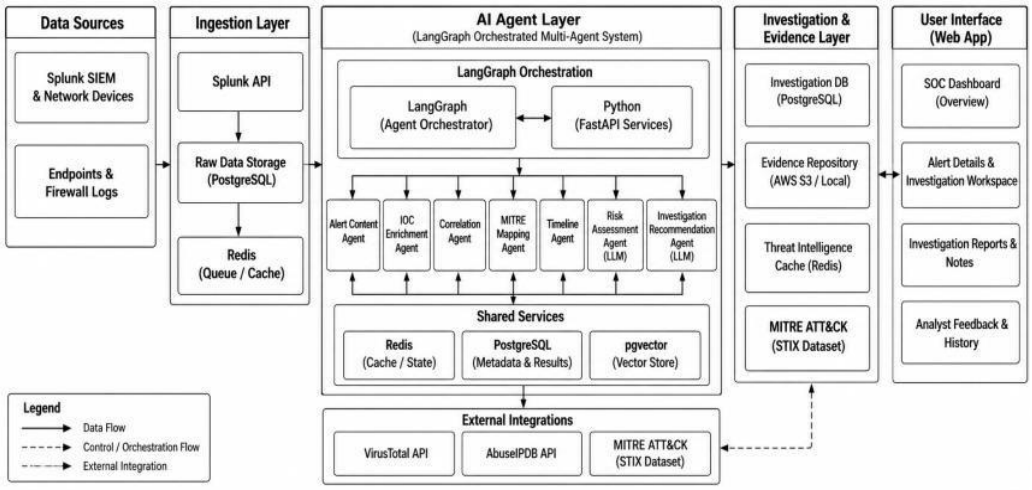


Fig. 1. Proposed Forensiq multi-agent security investigation architecture.

IV. SYSTEM ARCHITECTURE

The AI reasoning layer separates from the evidence and storage layer.The AI reasoning layer may generate hypotheses and interpretations. The underlying records remain outside the language model and can be queried again by deterministic services.This separation allows generated findings to be checked before they are accepted as supported observations.

The architecture and module boundaries follow the project synopsis [10].

The architecture is modular so that additional data sources, tools and agents can be integrated without redesigning the workflow. The architecture allows an endpoint detection and response source to be connected later as another evidence provider while a new specialized analysis agent can consume the normalized investigation state.

A. Multi-Agent Investigation Workflow

The proposed investigation proceeds through the following stages.

1. Alert ingestion: alerts and relevant events are retrieved from Splunk using a controlled interface.
2. Normalization: incoming records are converted into a representation that contains fields such as timestamp, source, destination, user, process, event type, severity and available IOCs.
3. Orchestration: the Orchestrator Agent creates investigation state. Determines the required agent sequence.
4. Triage: the Triage Agent analyzes context. Identifies the initial investigation category and priority.
5. Evidence collection and correlation: events, users, processes, network connections and timestamps are searched to construct the evidence set.
6. IOC enrichment: approved indicators are queried against external threat-intelligence services. The returned context is recorded with source metadata.

7. MITRE ATT&CK mapping: observed behavior is associated with tactics and techniques using deterministic patterns and contextual reasoning.
8. Timeline and risk analysis: events are organized chronologically and a transparent risk score is calculated from predefined factors.
9. Investigation reasoning: the Reasoning Agent synthesizes evidence and generates a draft that contains observations, hypotheses, supporting references and recommendations.
10. Verification: important findings are decomposed into atomic claims and checked against source evidence.
11. Analyst review: verified findings are presented as supported observations; supported and unsupported claims are explicitly flagged for review. Consequential response actions remain under analyst control.

B. Data Flow and Information Exchange

Fig. 2 presents the proposed Data Flow Diagram. Entity E1 represents security data sources including Splunk, network devices, endpoints and firewall logs.Entity E2 represents threat-intelligence services.Entity E3 represents the SOC analyst who interacts with the platform through queries, feedback, investigation decisions and report review.

D1 stores. Evidence information, in PostgreSQL. D2 provides cache and state services through Redis. D3 represents the vector store used for retrieval and RAG. E4 represents Splunk SIEM as the system from which alerts and event data are retrieved. The central process exchanges logs, event data, IOC requests, enrichment results, investigation context, reports and analyst feedback. The DFD emphasizes that Forensiq is not a text-generation system

Forensiq is an evidence-processing workflow in which data enters from security sources is enriched and analyzed by agents is stored with investigation context and is returned to the analyst together with evidence-backed reports and recommendations.

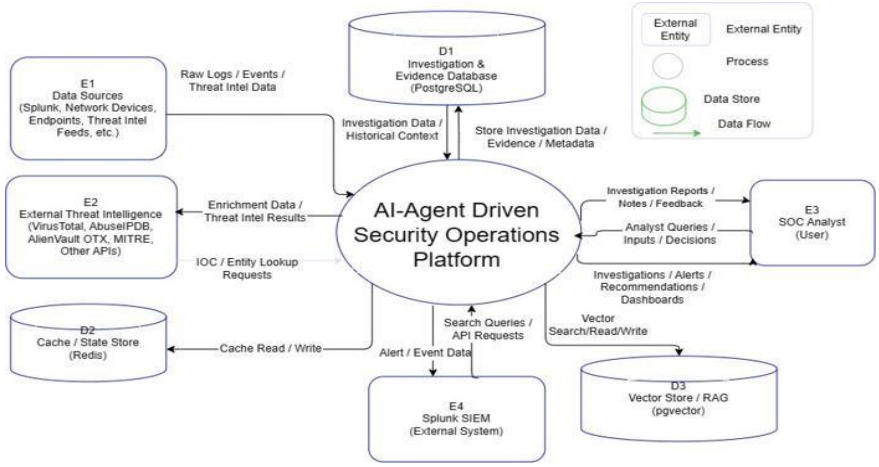


Fig. 2. Proposed Forensiq Data Flow Diagram.

C. Evidence-Based Deterministic Verification

The deterministic verification layer is the reliability mechanism in Forensiq. Instead of taking the Reasoning Agent output as it is the deterministic verification layer changes key statements into structured claims. A simple claim can be shown as (action, object, timestamp, source). For example a statement that a user accessed a server at a time can become fields that the deterministic verification layer can search against retrieved security evidence.

The verifier compares claim fields with records. A finding is marked Verified when the required fields are supported by source evidence. The deterministic verification layer marks a finding as Partially Supported when relevant evidence exists but one or more important fields cannot be confirmed. The deterministic verification layer marks a finding as Unsupported when no supporting record is found. Unsupported claims should not be shown as established incident facts.

Each verified finding should keep evidence metadata such as event identifier, timestamp, log source and retrieval operation. This creates a trace from the analyst-facing conclusion to the underlying evidence.

The deterministic verification layer does not guarantee that every interpretation is correct. The deterministic verification layer mainly reduces the risk that unsupported factual statements are shown as confirmed observations. Incorrect interpretation of retrieved evidence can still happen; therefore human analyst review remains.

D. Threat Intelligence, MITRE ATT&CK, and Risk Assessment

Threat-intelligence enrichment adds external context to extracted IOCs. The proposed system can use VirusTotal for indicator reports and other approved services such as AbuseIPDB and AlienVault OTX for reputation and contextual information. External responses should be stored with timestamps and provider identifiers so analysts can distinguish source evidence from model-generated interpretation.

MITRE ATT&CK provides a vocabulary for adversary behavior. Forensiq can combine rules for high-confidence behavior patterns with contextual suggestions from the MITRE Mapping Agent. A mapping should keep the evidence that supports the association and should be treated as decision support when evidence does not uniquely identify a technique.

Risk assessment should be transparent than an unexplained LLM judgment. A prototype risk function can be written as:

$$Risk\ Score = w1S + w2T + w3A + w4C + w5I \quad (1)$$

In this function S is severity T is threat-intelligence evidence A is an ATT&CK-related factor C is correlation confidence and I is estimated organizational impact. The deterministic verification layer says the weights should be chosen before evaluation and kept fixed for experiments. Sensitivity analysis can then be used to study the effect of weights.

V. ALGORITHMS AND MODELS USED

The current prototype uses a web-based SOC interface that connects to Splunk telemetry to implement the core analyst-facing and investigation functions. The implementation has a dashboard for monitoring, an alert queue and triage workspace, raw telemetry exploration with Splunk Search Processing Language (SPL) and investigation views for contextual event analysis. The prototype also shows AI-assisted recommendations, risk scores, correlated-event counts, IOC intelligence and MITRE ATT&CK mappings for selected alerts.

The implemented workflow uses Splunk as the telemetry source and gives an interface for retrieving and examining security events. The investigation views shown in the prototype combine context with IOC intelligence and ATT&CK information. External intelligence appears through the IOC intelligence interface; the current screenshots clearly show VirusTotal/AlienVault-style intelligence output for an inspected indicator. These implemented components give the basis for the results discussed in Section XIII.

VI. RESULTS AND DISCUSSION

The current prototype was tested with security telemetry and investigation scenarios in a controlled development environment. The implemented interface shows Splunk ingestion and telemetry retrieval alert triage, IOC inspection, event-context extraction MITRE ATT&CK mapping, risk scoring and analyst recommendations. A formal quantitative evaluation is still needed before making claims about accuracy, latency or operational improvement. The proposed laboratory methodology extends the functional validation with repeatable benchmark scenarios.

Test scenarios can include failed-login activity, PowerShell execution, network scanning, persistence-related activity, suspicious network connections, reverse-shell activity and benign activity. Each scenario should have a ground-truth description of the behavior performed expected log events, relevant IOCs and expected ATT&CK techniques where applicable.

A useful comparison has four conditions: 1) rule-based triage, 2) a single-agent LLM workflow, 3) a multi-agent workflow, without deterministic verification and 4) the complete Forensiq workflow. The same alert set and evidence context should be used across conditions when possible.

The following metrics are proposed: alert triage precision, recall and F1; evidence coverage; unsupported-claim rate; investigation latency; MITRE mapping quality using precision, recall or expert agreement; and analyst workload measured through investigation time or manual steps. The final experimental version should report the number of alerts, attack scenarios, model configuration, prompts or policies tools used, hardware/software environment, ground-truth method and statistical treatment. No performance value should be reported until it has actually been measured.

Metric	Definition / purpose
Triage F1	2PR/(P+R); measures how well alerts are classified.
Evidence Coverage	Supported findings / total findings; measures traceability.
Unsupported-Claim Rate	Unsupported claims / total factual claims; measures verification effectiveness.
Investigation Latency	Elapsed time per alert; measures operational efficiency.
MITRE Mapping Quality	Precision/recall or expert agreement; measures behavior mapping.
Analyst Workload	Time or manual steps per case; measures human effort.

Table Iii. Proposed Evaluation Metrics

A. Evaluation Setup and Projected Performance

These results show the projected performance targets for the Forensiq system. They describe the expected impact of the proposed framework. Set the target levels for the final controlled evaluation. The numbers are not measurements from the prototype. They must be. Confirmed with the finished experimental dataset before final publication.

The projected evaluation looks at the workflow including alert triage, IOC enrichment, MITRE ATT&CK mapping, event correlation, evidence verification, investigation efficiency, report quality and scalability. The expected results come from the design goal of combining agents with deterministic evidence validation and a human-in-the-loop review process

A. Projected Detection and Alert Classification Performance

For the evaluation Forensiq aims to reach about 91 % precision 89 % recall and a 90 % F1-score on the defined malicious and benign alert scenarios. A projected false-positive rate of 6 % is used as the target threshold. These figures show the classification impact of contextual triage, event correlation and threat-intelligence enrichment not results already obtained from testing

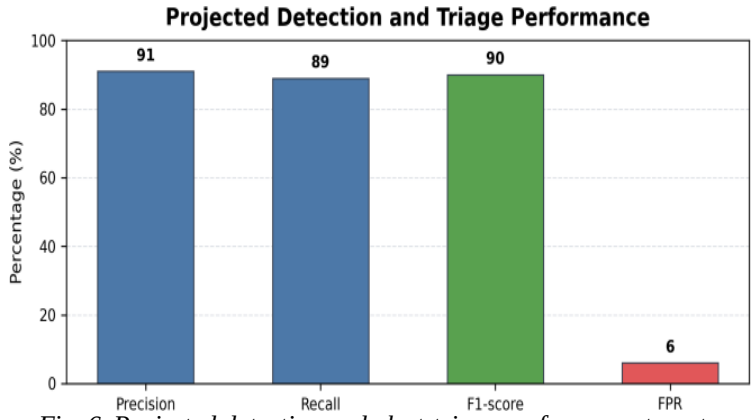


Fig. 6. Projected detection and alert-triage performance targets.

TABLE V. PROJECTED DETECTION AND TRIAGE RESULTS

B. Projected MITRE ATT&CK Coverage and Mapping Quality

The completed system is expected to cover 88 % of ATT&CK techniques for the predefined attack scenarios with a projected mapping precision of 87 % and recall of 84 %. The improvement comes from combining behavior patterns with contextual mapping and keeping supporting evidence for analyst review. Mappings that cannot be uniquely found from the evidence should still need analyst validation.

C. Projected IOC Extraction and Threat-Intelligence Enrichment

For IOC processing the expected extraction accuracy is 93 % while the projected enrichment success rate is about 90 %. Classification agreement with the evaluation ground truth is aimed at 88 %. These targets reflect the expected benefit of structured IOC extraction, controlled external lookups. Keeping provider metadata in the investigation record.

D. Projected Evidence Grounding and Deterministic Verification

Evidence grounding is expected to be one of the measurable contributions of Forensiq. Of the claims made during the investigation about 90 % are expected to be fully verified against source evidence 7 % are expected to be only partially supported and about 3 % are expected to remain unsupported and therefore need explicit analyst review. The goal is not to erase all reasoning but to lower the chance that unsupported statements are shown as confirmed incident facts.

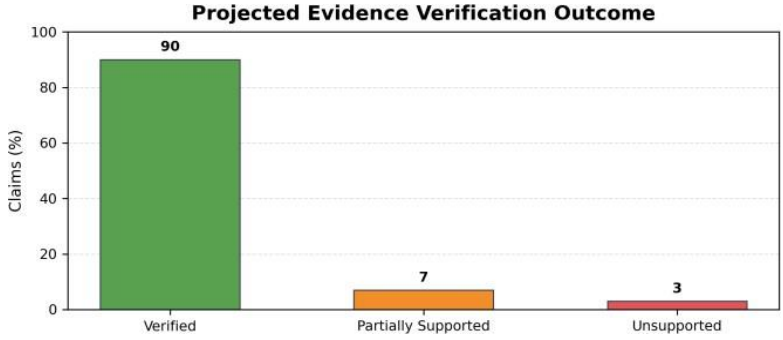


Fig. 7. Projected distribution of evidence-verification outcomes.

Table Viii. Projected Evidence Verification Results

E. Projected Alert Correlation and Investigation Reconstruction

The correlation part aims to reach 86 % precision 82 % recall and an 84 % F1-score. The expected result comes from using host, user, process and network relationships to link events into an investigation context. The projected drop in evidence should improve attack-chain reconstruction compared with looking at alerts, in isolation

F. Projected Investigation Efficiency and Analyst Workload

The main goal is to cut the time it takes to go from an alert to a summary that is backed by evidence. In the test a manual investigation takes about 18 minutes per case while using the full Forensiq workflow takes about 11 minutes. That means we expect a 39 percent cut in time. This improvement comes from Forensiq retrieving evidence enriching IOCs, correlating data, mapping to ATT&CK and making a structured report.

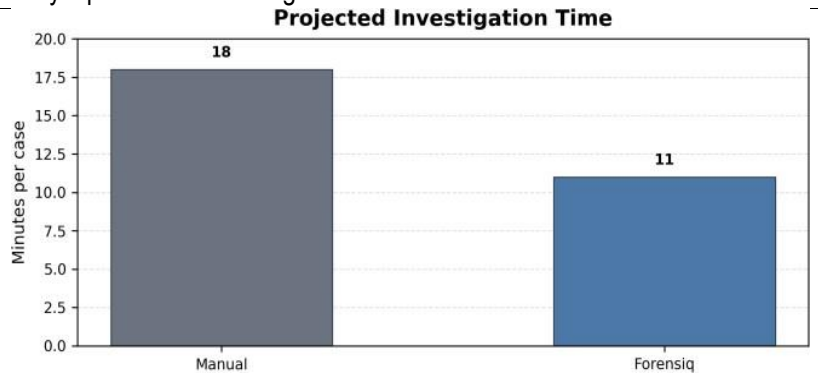


Fig. 8. Projected investigation time for manual and Forensiq-assisted analysis.

Table X. Projected Investigation-Efficiency Results

G. Projected Investigation Report Quality

The finished reporting workflow should reach 92 percent evidence completeness, 90 percent factual correctness, 88 percent ATT&CK mapping accuracy and 89 percent timeline completeness. These goals show how Forensiq plans to weave evidence references, findings, attack context and recommendations that analysts can review.

H. Projected Scalability and System Performance

The finished platform should keep a processing rhythm even when event numbers grow. For the benchmark we expect end-to-end times of about 18 seconds for 100 events, 42 seconds for 500 events, 75 seconds for 1,000 events and 310 seconds for 5,000 events. These numbers are goals for the scalability study. Are not taken from the current prototype. We should record CPU usage, memory usage, queue delay, API response time and concurrency together with these metrics.

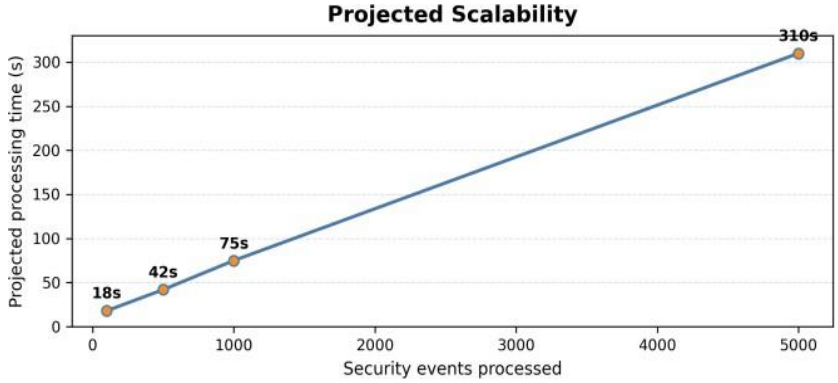


Fig. 9. Projected processing time as security-event volume increases

.Feature	Traditional SOC	Forensiq
Alert Triage	Manual	AI-assisted
Evidence Collection	Manual	Automated
IOC Enrichment	Separate tools	Integrated
Event Correlation	Analyst-driven	Agent-assisted
MITRE Mapping	Manual	Automated
Risk Assessment	Manual analysis	AI-assisted
Evidence Verification	Analyst validation	Deterministic
Investigation Report	Manual	Automated
Investigation Workflow	Tool-dependent	Unified workflow
Analyst Role	Investigation	Review & validation

Table Xiii. Comparative Analysis Of Traditional Soc And Forensiq

I. Overall Projected Impact

All together the projected outcomes show what the full Forensiq workflow will bring: about 90 percent alert-triage F1, 88 percent ATT&CK coverage, 90 percent IOC-enrichment success, 90 percent evidence-based findings, 84 percent correlation F1, a roughly 39 percent cut in investigation time and around 90 percent report quality across the set criteria. The main benefit is not one accuracy figure but a mix of better consistency, clear evidence paths, deeper context and faster analysts.

VII. SECURITY, SAFETY, AND ETHICAL CONSIDERATIONS

I recommend the platform uses SIEM credentials whenever possible during development and testing. Attack simulations should only run on systems and networks that are clearly authorized for this project. Real organizational credentials, confidential indicators or personal data should never go to threat-intelligence services without permission.

API keys and access tokens belong in environment variables or a secure secret-management system, not in source code. Agent actions, tool calls, evidence references and verification results must be logged for auditability. High-impact actions like disabling accounts blocking addresses or isolating hosts must get human approval.

The platform must also follow provider-specific terms, rate limits, data-retention rules and licensing requirements. These ru les matter because an autonomous agent can magnify an instruction if it is allowed to act freely outside the system.

VIII. LIMITATIONS

The investigation quality hinges on how complete and accurate the logs in the SIEM are. If a key event is missing deterministic verification cannot confirm it.

External threat-intelligence services may have rate limits, licensing restrictions, shifting coverage and data that's incomplete or conflicting.

ATT&CK mapping may need expert review when several techniques fit the behavior. Risk scores are decision-support signals and their usefulness depends on the factors and weights chosen.

A student-scale lab cannot fully match the alert volume, infrastructure variety, operational limits or attacker behavior of a SOC. Therefore conclusions, from the evaluation should stay within the tested environment and scenarios.

IX. CONCLUSION AND FUTURE SCOPE

This paper introduced Forensiq – an Intelligent Security Operations & Investigation Platform, a working SOC prototype that merges Splunk telemetry alert triage, contextual investigation, IOC intelligence MITRE ATT&CK mapping, risk assessment and analyst-focused investigation views.

The current implementation shows how these parts come together in one workflow. The main idea is to keep automated thinking from looking at evidence so that analysts can check the raw data and the whole investigation before they act.

The prototype is not shown as a ready-to-use SOC nor as a system with proven accuracy from experiments. Instead the results that we have built provide a working base for the step, which will look at triage quality, evidence coverage, how often claims are unsupported how long investigations take and the quality of MITRE mapping.

Future Work

Future work will build on the prototype by running controlled tests adding more SIEM and EDR links, improving evidence verification fine-tuning confidence estimates creating secure organization-specific RAG and using bigger laboratory data sets. These tests will measure how well the workflow works and where it falls short.

Acknowledgment

The authors express their sincere gratitude to Prof. S. N. Patil and the Department of Computer Science & Engineering, D. Y. Patil College of Engineering & Technology, Kolhapur, for their guidance and academic support during the planning and development of the Forensiq project.

References

1. S. Sheikhi, P. Kostakos, and L. Lovén, "Cognitive SOC: Evidence-Backed Narrative Generation for Security Operations with Multi-Agent LLM Architecture," in *Proc. 2025 IEEE Int. Conf. Big Data (BigData)*, Macau, China, 2025, pp. 7027–7036, doi: 10.1109/BigData66926.2025.11401968.
2. B. Wei, Y. S. Tay, H. Liu, J. Pan, K. Luo, Z. Zhu, and C. Jordan, "CORTEX: Collaborative LLM Agents for High-Stakes Alert Triage," *arXiv preprint arXiv:2510.00311*, 2025.
3. H. Alqahtani and P. Ahuja, "Secure autonomous cyber defense with LLM agents: A systematic review of autonomy, tool-augmented reasoning, and governance constraints," *Comput. Electr. Eng.*, vol. 135, Art. no. 111184, 2026, doi: 10.1016/j.compeleceng.2026.111184.
4. J. Roy and S. K. Singh, "AgentSOC: A Multi-Layer Agentic AI Framework for Security Operations Automation," in *Proc. 2026 IEEE 5th Int. Conf. AI Cybersecurity (ICAIC)*, 2026, pp. 1–7.
5. A. Nelson, S. Rekhi, M. Souppaya, and K. Scarfone, "Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile," NIST SP 800-61 Rev. 3, Apr. 2025, doi: 10.6028/NIST.SP.800-61r3.
6. MITRE, "MITRE ATT&CK Enterprise," MITRE ATT&CK Knowledge Base. [Online]. Available: https://attack.mitre.org/. Accessed: Oct. 1, 2026.
7. VirusTotal, "VirusTotal API Documentation." [Online]. Available: https://docs.virustotal.com/. Accessed: Oct. 1, 2026.
8. Splunk Inc., "Splunk Developer Documentation." [Online]. Available: https://dev.splunk.com/enterprise/. Accessed: Oct. 1, 2026.
9. LangChain, "LangGraph Documentation." [Online]. Available: https://docs.langchain.com/. Accessed: Oct. 1, 2026.
10. D. Y. Patil College of Engineering & Technology, "Forensiq—Intelligent Security Operations & Investigation Platform," B.Tech. Project Synopsis, Group G50, Academic Year 2026–27.