A Review of Ai-Based Cybersecurity Monitoring Systems: Contextualising The Securai Sentinel Intelligent Threat Detection And RESPONSE PLATFORM

Arikaran P[1], Yogendra Sai P[2], Kalluri Nagalakshmi[3], Yeruva Sai Hanumantha Reddy[4]
[1, 2, 3, 4] Dept of Cybers ecurity
[1, 2, 3, 4] Dhanalakshmi Srinivasan University Samayapuram, Tiruchirappalli - 621112, Tamil Nadu, India

Abstract- *Cybersecurity threats have grown exponentially in complexity and volume, rendering traditional rule-based monitoring tools increasingly inadequate for modern enterprise and institutional environments. Existing security platforms are often siloed, lack intelligent reasoning capabilities, and fail to integrate cross-domain threat intelligence with automated response workflows. This review paper surveys the evolution of AI-based cybersecurity monitoring systems, examining machine learning-based intrusion detection, threat intelligence platforms, dark web surveillance systems, MITRE ATT&CK-aligned detection frameworks, and large language model applications in security operations. Through systematic examination of eight significant research contributions, five persistent research gaps are identified: absence of unified multi-domain threat correlation, exclusion of AI-driven red team simulation, reliance on static signature-based rule databases, lack of natural-language explainability in threat analysis, and the absence of integrated zero-trust policy management. These gaps collectively justify the conceptual design of SecurAI Sentinel, a proposed full-stack AI-powered cybersecurity web application integrating eight intelligence modules—CVE Intelligence Hub, Dark Web Monitor, MITRE ATT&CK Mapper, Incident Response Playbook Generator, Forensics Timeline Builder, Packet Capture Analyzer, AI Red Team Agent, and Zero Trust Policy Builder—within a unified glassmorphic interface powered by Google Gemini AI, React, Node.js, and Express.js. The paper concludes with a discussion of the system's feasibility, societal impact, and directions for future research.*

Keywords: cybersecurity; artificial intelligence; threat detection; CVE; MITRE ATT&CK; dark web monitoring; incident response; Google Gemini; React; Node.js; zero trust architecture; red team simulation; packet analysis; large language models; security operations

I. INTRODUCTION

The global cybersecurity landscape has undergone a fundamental transformation over the past decade. Organisations of all scales face threats from sophisticated adversaries who employ automation, artificial intelligence, and zero-day exploits to penetrate defences designed for a previous era of attack sophistication. Traditional security tools—firewalls, signature-based antivirus engines, and manual vulnerability scanners—while necessary, are no longer sufficient as standalone defences against modern persistent threats.

The proliferation of connected devices, cloud-native workloads, and remote access architectures has dramatically expanded the attack surface available to malicious actors. According to data from the National Vulnerability Database (NVD), over 25,000 new Common Vulnerabilities and Exposures (CVEs) were recorded in 2023 alone, representing a sustained upward trend that exceeds the capacity of human security analysts to manually triage and remediate without intelligent automation.

The emergence of large language models (LLMs) such as Google Gemini, OpenAI GPT-4, and open-source alternatives including Mistral and LLaMA has introduced a new paradigm in AI-assisted security operations. These models demonstrate sophisticated natural language understanding, multi-step reasoning, and structured output generation capabilities directly applicable to threat analysis, vulnerability explanation, and incident response planning. When combined with real-time threat intelligence feeds and structured security frameworks such as MITRE ATT&CK, LLMs can produce analyses that are simultaneously contextual, actionable, and aligned with industry-standard taxonomies.

Yet the current landscape of AI security tools remains fragmented. Commercial Security Information and Event Management (SIEM) systems are expensive, require extensive configuration, and typically focus on log aggregation rather than intelligent threat reasoning. Open-source alternatives lack the integration breadth and AI-powered explainability that modern security operations require. No comprehensive, student-accessible platform exists

that unifies CVE intelligence, dark web monitoring, MITRE ATT&CK correlation, incident response planning, forensic analysis, packet inspection, red team simulation, and zero-trust policy generation under a single AI-augmented interface.

This review paper surveys and critically analyses the body of research underpinning AI-based cybersecurity monitoring systems, with the specific objective of identifying gaps that justify the design of SecurAI Sentinel. The paper is structured as follows: Section II provides background on cybersecurity monitoring systems; Section III presents the literature review; Section IV analyses strengths and limitations; Section V enumerates research gaps; Section VI introduces the proposed system; Section VII describes the system architecture; Section VIII discusses impact and feasibility; Sections IX through XI address advantages, future scope, and conclusions.

II. OVERVIEW OF CYBERSECURITY MONITORING SYSTEMS

A. Traditional Cybersecurity Tools

Traditional cybersecurity monitoring has historically relied upon network perimeter defences, signature-based intrusion detection systems (IDS), and manual vulnerability scanning. Tools such as Snort, Suricata, and Nessus represent the first generation of automated security monitoring. While these provide foundational detection capabilities, they suffer from high false-positive rates, inability to detect zero-day threats, and the requirement for continuously updated signature databases. The operational overhead of maintaining these systems at scale is substantial, particularly for resource-constrained organisations.

B. SIEM and Threat Intelligence Platforms

The second generation of security monitoring platforms introduced Security Information and Event Management (SIEM) systems that aggregate logs across the enterprise and apply correlation rules to detect anomalous behaviour patterns. Splunk, IBM QRadar, and Microsoft Sentinel represent this paradigm. These platforms improved detection coverage but retained critical limitations: they require significant tuning effort, generate alert fatigue through high false-positive rates, and provide minimal contextual intelligence about the meaning or severity of detected events beyond rule-matched signatures.

C. AI-Driven Security Operations

The third generation incorporates machine learning, behavioural analytics, and large language models to automate the detection, triage, and response workflow. AI-driven platforms can identify anomalous patterns without predefined rules, correlate events across disparate data sources, and generate natural-language explanations of threat activity. The integration of LLMs specifically enables the translation of complex technical threat data into analyst-accessible summaries, the automatic generation of incident response procedures, and the simulation of adversarial attack paths for proactive defence posture improvement.

D. Google Gemini in Security Applications

Google Gemini's multimodal reasoning and structured output generation capabilities make it particularly suited for cybersecurity applications. Its ability to process structured JSON threat data, generate MITRE ATT&CK-aligned analyses, produce contextual CVE summaries, and simulate adversarial reasoning chains provides a comprehensive AI substrate for a unified security platform. The availability of the Gemini API through Google AI Studio enables academic and student projects to access production-grade LLM capabilities without custom model training or significant infrastructure investment.

III. LITERATURE REVIEW

This section presents a systematic analysis of eight significant research contributions spanning AI-based cybersecurity, intrusion detection, MITRE ATT&CK integration, dark web intelligence, LLM applications in security, zero-trust architecture, red team simulation, and network traffic analysis.

A. ML Methods for Cyber Security Intrusion Detection (Buczak & Guven, 2016)

Buczak and Guven published a comprehensive survey in IEEE Communications Surveys & Tutorials examining data mining and machine learning methods applied to cybersecurity intrusion detection [1]. The authors systematically catalogue classification algorithms including decision trees, neural networks, support vector machines, and ensemble methods across benchmark datasets. The survey establishes foundational findings: ensemble methods consistently outperform single-classifier approaches, feature selection is critical for reducing false-positive rates, and no single algorithm dominates across all attack categories. The work provides the theoretical basis for ML-based anomaly detection modules.

The survey is significant for establishing the theoretical foundations of ML-based threat detection. However, it predates the LLM era and does not address

explainability, natural-language threat analysis, or integration with standardised threat taxonomies such as MITRE ATT&CK, which are core requirements for a modern unified security platform.

B. MITRE ATT&CK: Design and Philosophy (Strom et al., 2018)

Strom and colleagues introduced the MITRE ATT&CK framework in a foundational technical report describing its design principles, taxonomy structure, and intended applications [2]. ATT&CK organises adversary techniques and sub-techniques across fourteen tactical categories from Initial Access through Exfiltration and Impact, providing a structured common language for describing threat actor behaviour. The framework has since become the de facto standard for threat intelligence sharing and detection engineering across the security industry.

The framework itself is a classification system rather than an automated detection tool. No existing implementation in the reviewed literature provides real-time automated classification of detected threats from multiple security data sources into ATT&CK techniques through LLM reasoning, which represents a significant implementation gap that SecurAI Sentinel's MITRE ATT&CK Mapper module directly addresses.

C. Hacker Assets in Underground Forums (Samtani et al., 2020)

Samtani and colleagues examined dark web and underground hacker forums as sources of early-warning threat intelligence [3]. Their research demonstrates that cyber threat indicators—including newly developed exploits, credential dumps, and malware-as-a-service offerings—consistently appear in underground marketplaces before formal CVE disclosure and vendor patches. The study employs natural language processing and graph analytics to extract structured threat intelligence from unstructured dark web content, providing empirical validation of dark web monitoring as a proactive threat intelligence mechanism.

The proposed system requires significant infrastructure for Tor network access and NLP processing that is not accessible in lightweight web application deployments. No integration with user-facing monitoring interfaces or breach notification workflows is described. The gap between research capability and accessible implementation directly motivates the Dark Web Monitor module in SecurAI Sentinel.

D. LLMs for Cybersecurity Applications (Ferrag et al., 2023)

Ferrag and colleagues published a comprehensive survey examining the application of large language models across the cybersecurity domain [4]. The authors review LLM applications across vulnerability analysis, malware analysis, threat intelligence, intrusion detection, and security code generation. They find that instruction-tuned models demonstrate strong performance on structured cybersecurity tasks but exhibit hallucination risks when producing specific technical claims about vulnerability details, CVSS scores, or affected software versions without retrieval augmentation.

The survey directly motivates the use of retrieval-augmented generation in any LLM-based cybersecurity tool making specific factual claims. These findings justify the NVD API integration in SecurAI Sentinel's CVE Intelligence Hub, which grounds LLM summaries in verified CVE records rather than relying on model memory, eliminating the hallucination risk identified by Ferrag et al.

E. Zero Trust Architecture (Rose et al., 2020)

Rose and colleagues authored NIST Special Publication 800-207, establishing the formal definition and architectural principles of Zero Trust Architecture (ZTA) [5]. The publication defines zero trust as a collection of concepts and ideas designed to minimise uncertainty in enforcing accurate, least-privilege per-request access decisions. Seven tenets of zero trust are enumerated, covering network segmentation, continuous verification, device health assessment, and data-centric security policy.

While NIST SP 800-207 provides comprehensive policy guidance, it does not provide automated policy generation tools or intelligent assistance for translating organisational requirements into specific zero-trust policy configurations. The gap between the policy framework and practical implementation guidance represents a significant usability barrier that the Zero Trust Policy Builder module in SecurAI Sentinel directly addresses.

F. Machine Learning in Cybersecurity (Apruzzese et al., 2023)

Apruzzese and colleagues conducted an empirical study published in IEEE Transactions on Information Forensics and Security examining the role of machine learning in operational cybersecurity [6]. Through structured interviews with security practitioners, the authors identify a critical gap between academic ML security research and operational

reality: security analysts require not only detection outputs but contextual explanations of why a given activity is suspicious, what attacker objective it serves, and what response actions are appropriate.

The study's findings directly motivate SecurAI Sentinel's emphasis on natural-language explainability as a first-class design requirement. The operational insight that unexplained detection outputs generate alert fatigue and analyst disengagement informs the platform's design philosophy: every module produces both structured data outputs and human-readable AI-generated explanations.

G. Automated Cyber Threat Intelligence (Al-Shaer et al., 2022)

Al-Shaer and colleagues examined automated approaches to cyber threat intelligence generation, focusing on the extraction of structured threat indicators from unstructured security reports and their automated mapping to STIX/TAXII formats and ATT&CK techniques [7]. Their system demonstrates that transformer-based models fine-tuned on security corpora can accurately classify threat descriptions into ATT&CK technique categories with statistically significant improvement over keyword-matching baselines.

The work establishes technical feasibility for automated ATT&CK mapping but requires fine-tuned domain-specific models trained on labelled security corpora, which are unavailable in standard student project deployments. SecurAI Sentinel pursues a zero-shot approach using instruction-tuned general-purpose LLMs, representing a more accessible implementation path that achieves comparable classification utility without specialised training data.

H. LLM-Based Vulnerability Analysis (arXiv, 2024)

Recent work published on arXiv examines the use of LLMs for automated vulnerability triage and security report generation in continuous integration pipelines [8]. The authors demonstrate that instruction-tuned LLMs can parse SAST and DAST scanner outputs, classify findings by severity and exploitability, generate natural-language summaries for non-technical stakeholders, and produce prioritised remediation guidance. Evaluation against manually produced analyst reports shows strong semantic similarity, suggesting LLM-generated security analysis is approaching the quality of experienced analyst output for structured vulnerability types.
The work highlights the transformative potential of LLM integration in security workflows while noting the persistent risk of confident-sounding but factually incorrect analysis on novel vulnerability classes. This reinforces the requirement for grounding mechanisms and API-verified data integration in any production security application, a design principle SecurAI Sentinel operationalises through its NVD API integration.

IV. ANALYSIS OF EXISTING SYSTEMS

A. Comparative Summary

Table 1 presents a structured comparative analysis of the reviewed works across dimensions critical to a unified AI cybersecurity platform.

Table 1: Comparative Analysis of Reviewed Cybersecurity Systems

Study	Unified	MITRE	Dark Web	Red Team	LLM Expl.
Buczak & Guven (2016)	No	No	No	No	No
Strom et al. (2018)	Part.	Yes	No	No	No
Samtani et al. (2020)	No	No	Yes	No	No
Ferrag et al. (2023)	Part.	Part.	No	No	Part.
Rose et al. (2020)	No	Part.	No	No	No
Apruzzese et al. (2023)	Part.	Part.	No	No	Part.
Al-Shaer et al. (2022)	Part.	Yes	No	Part.	No
arXiv (2024)	Part.	Part.	No	No	Yes
SecurAI Sentinel	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

B. Strengths of Existing Literature

The reviewed body of work demonstrates several commendable advances. The maturation of ML-based anomaly detection has substantially reduced false-negative rates for known attack families across benchmark datasets. The MITRE ATT&CK framework has provided the security

community with a shared taxonomy enabling structured threat intelligence sharing and detection engineering at scale. The demonstrated applicability of transformer models to security text analysis establishes the technical foundation for LLM-powered security operations. The growing literature on dark web intelligence confirms that proactive monitoring of underground ecosystems provides genuine early-warning capability beyond traditional reactive detection.

C. Limitations of Existing Literature

Notwithstanding these strengths, five persistent limitations define the research gap:

- Fragmented Tool Ecosystems: No reviewed system integrates CVE intelligence, dark web monitoring, MITRE ATT&CK mapping, incident response, forensics, packet analysis, red team simulation, and zero-trust policy in a single platform.
- Absence of LLM-Powered Explainability: The majority of reviewed systems produce detection outputs without natural-language reasoning enabling analyst understanding and stakeholder communication.
- Static Signature Dependence: Many reviewed systems rely on static rule databases or pre-trained models that cannot incorporate new threat intelligence without retraining or manual rule updates.
- No AI Red Team Integration: None of the reviewed systems include an AI-driven adversarial simulation capability for proactive attack surface assessment and kill-chain modelling.
- Zero Trust Policy Gap: No reviewed system provides AI-assisted policy generation aligned with NIST SP 800-207 ZTA principles for organisational security configuration.

V. RESEARCH GAPS IDENTIFIED

Based on the systematic analysis in Sections III and IV, five research gaps are formally identified.

A. Absence of Unified Multi-Domain Threat Correlation

Every reviewed AI-based cybersecurity system addresses a subset of the threat intelligence domain in isolation. A security analyst investigating a potential incident must context-switch across CVE databases, OSINT tools, dark web monitors, SIEM dashboards, and incident response playbooks. No unified platform correlates intelligence across all these dimensions through a single AI reasoning engine. The operational cost of this fragmentation is measurable in mean-time-to-respond (MTTR) and analyst cognitive load, both of which directly impact organisational security posture.

B. Exclusion of AI Red Team Simulation

Proactive security posture assessment through red team exercises is universally acknowledged as a critical component of mature security programmes. Yet no reviewed AI system provides an automated red team agent capable of simulating realistic attack chains against a described target environment and generating MITRE ATT&CK-mapped adversarial paths. The availability of instruction-tuned LLMs with broad cybersecurity knowledge makes this capability technically feasible without requiring purpose-built offensive security tooling or certified penetration testing infrastructure.

C. Static Knowledge and Hallucination Risk

LLM-based security analysis systems operating without real-time API grounding are vulnerable to producing plausible but factually incorrect information about CVE details, CVSS scores, affected software versions, and patch availability. Ferrag et al.'s survey quantifies this hallucination risk empirically. In operational security contexts, an analyst acting on incorrect vulnerability information faces direct organisational risk. Real-time NVD API grounding is therefore non-negotiable for any LLM-based vulnerability analysis tool deployed in a professional or educational context.

D. Lack of Natural-Language Explainability

Security teams increasingly include analysts at varying levels of technical expertise, and security findings must be communicated to non-technical stakeholders including management, legal counsel, and compliance officers. No reviewed system produces natural-language explanations of threat detections, vulnerability impacts, or incident timelines accessible to non-specialists. This gap limits the organisational value of technical security findings and creates barriers to security awareness development within institutions.

E. Absence of Zero Trust Policy Generation

The zero-trust architecture paradigm has been formally endorsed by NIST, government agencies across multiple jurisdictions, and enterprise security leadership globally. Yet the translation from ZTA principles to specific implementable policy configurations remains a manual, expert-dependent process with no AI-assisted tooling in the reviewed literature. This gap particularly impacts smaller organisations and educational institutions lacking dedicated security architecture teams.

VI. PROPOSED SYSTEM: SECURAI SENTINEL

A. Conceptual Overview

SecurAI Sentinel is conceived as a third-generation AI cybersecurity monitoring platform directly addressing each of the five research gaps identified in Section V. The system is built on three foundational principles derived from the reviewed literature: (i) LLM-powered unified threat intelligence, where Google Gemini serves as the central reasoning engine across all eight security modules; (ii) real-time API grounding in verified data sources including the NVD CVE database and breach intelligence APIs, motivated by the hallucination risk identified by Ferrag et al.; and (iii) MITRE ATT&CK alignment as a first-class design constraint, ensuring all threat analyses are expressed in industry-standard adversarial taxonomy.

The system accepts varied inputs across its eight modules—CVE identifiers, email addresses, domain names, network packet captures, incident descriptions, target environment specifications, and organisational security requirements—and produces structured outputs including vulnerability summaries, breach reports, ATT&CK-mapped threat analyses, incident response playbooks, forensic timelines, traffic anomaly reports, adversarial attack paths, and zero-trust policy documents.

B. Addressing Identified Gaps

The design of SecurAI Sentinel maps directly to each identified gap:

- Gap 1 (Unified Correlation): Addressed through a single React frontend unifying all eight modules under one interface, with shared AI session context enabling cross-module threat correlation through the Gemini reasoning engine.
- Gap 2 (Red Team Simulation): Addressed through the AI Red Team Agent module, which accepts target environment descriptions and generates MITRE ATT&CK-mapped adversarial kill chains with specific technique references and remediation guidance.
- Gap 3 (Hallucination Risk): Addressed through mandatory NVD API integration in the CVE Intelligence Hub, grounding all vulnerability claims in verified NVD records before LLM analysis is applied.
- Gap 4 (Explainability): Addressed through Gemini-generated natural-language analysis in every module output, producing analyst-accessible summaries and non-technical stakeholder reports alongside structured technical data.
- Gap 5 (Zero Trust Policy): Addressed through the Zero Trust Policy Builder, which uses prompt-embedded NIST SP 800-207 tenets as structural constraints to generate implementable policy documents from organisational input.

C. Technical Innovation

The principal technical innovation of SecurAI Sentinel relative to prior work is the integration of eight AI-powered security intelligence modules under a unified React frontend with a consistent glassmorphic cyberpunk interface, a Node.js/Express.js backend providing API proxying and data aggregation, and Google Gemini as the shared AI reasoning engine. This integration, combined with real-time NVD API grounding and the MITRE ATT&CK taxonomy as a structural output constraint, represents a combination not previously implemented in an accessible, open-architecture student-level security platform.

VII. SYSTEM ARCHITECTURE

A. Architectural Overview

SecurAI Sentinel is implemented as a three-tier web application. The presentation layer is a React single-page application providing the unified glassmorphic cyberpunk interface. The application layer is a Node.js/Express.js server handling API requests, external service proxying, and data transformation. The intelligence layer integrates the Google Gemini API as the AI reasoning engine alongside external data APIs including the NVD CVE API, breach intelligence APIs, and the MITRE ATT&CK STIX data feed.

B. Frontend Layer (React)

The frontend is built with React 18 and Vite, implementing a consistent glassmorphic cyberpunk aesthetic characterised by translucent glass-effect panels, neon accent colours on a dark background, and animated status indicators. Eight module components each provide a dedicated interface for their respective security intelligence function. A global navigation sidebar enables rapid module switching. All module results are rendered as structured cards with expandable detail sections, MITRE ATT&CK technique badges, and severity indicators colour-coded by CVSS score or threat level.

C. Backend Layer (Node.js/Express.js)

The Express.js backend serves three primary functions: it proxies requests to the NVD CVE API and breach intelligence APIs, preventing CORS issues and protecting API keys from client-side exposure; it preprocesses and normalises

API responses into structured formats suitable for Gemini prompt construction; and it provides the PCAP file upload and parsing endpoint using the node-pcap library for network traffic analysis. The backend is stateless, enabling horizontal scaling without session management complexity.

D. AI Intelligence Layer (Google Gemini)

The Gemini API serves as the unified AI reasoning engine across all eight modules. Each module constructs a module-specific system prompt embedding relevant domain constraints: the CVE module embeds CVSS scoring criteria and NVD data schemas; the ATT&CK mapper embeds the complete tactic-technique taxonomy; the red team agent embeds adversarial reasoning chains and kill-chain stages; the zero-trust builder embeds NIST SP 800-207 tenets as invariant output constraints. Structured JSON output is enforced through prompt engineering, ensuring consistent, parseable responses the frontend renders as typed component data.

E. Architecture Diagram Description

The architecture diagram is drawn as a horizontal left-to-right flow beginning with the React frontend module components (CVE Hub, Dark Web Monitor, ATT&CK Mapper, IR Playbook, Forensics, Packet Analyzer, Red Team Agent, Zero Trust Builder), progressing through the Express.js API gateway to the parallel intelligence sources: the Gemini AI engine with module-specific prompt templates; the NVD CVE API for real-time vulnerability grounding; breach intelligence APIs for dark web monitoring; and the MITRE ATT&CK STIX feed for taxonomy alignment. Response data flows back through the API gateway where it is validated and normalised before rendering in the React interface.

VIII. DISCUSSION

A. Significance of the Proposed System

The primary significance of SecurAI Sentinel lies in its potential to democratise access to sophisticated, multi-domain security intelligence for students, early-career security analysts, small organisations, and educational institutions that cannot afford enterprise SIEM deployments. The system's AI-powered explainability dimension is particularly significant: by generating natural-language threat analyses, incident response procedures, and policy documents, the platform makes advanced security concepts accessible to audiences without deep technical backgrounds, accelerating security awareness and capability development across India's rapidly growing cybersecurity workforce.

The MITRE ATT&CK alignment implemented as a first-class design constraint across all modules ensures that the platform's outputs are immediately useful in professional security contexts, as ATT&CK-mapped findings are directly actionable within modern SOC workflows and can be incorporated into threat reports, detection engineering tasks, and purple team exercises without requiring post-processing or reformatting.

B. Feasibility Assessment

The technical feasibility of SecurAI Sentinel is supported by three observable facts. First, the Google Gemini API, NVD CVE API, and HaveIBeenPwned API are production-ready services with documented interfaces requiring no custom model training or proprietary data infrastructure. Second, the React and Node.js/Express.js technology stack is mature, widely adopted, and supported by extensive open-source tooling. Third, the NVD CVE API is a free government-operated service, making the system's core intelligence grounding accessible without commercial API budget constraints that would preclude student project deployment.

C. Educational and Societal Impact

At the educational level, SecurAI Sentinel provides a hands-on learning environment in which cybersecurity students can explore real CVE data, understand MITRE ATT&CK techniques in context, simulate adversarial reasoning, and develop practical incident response skills. The platform's integration of industry-standard frameworks—NVD, MITRE ATT&CK, NIST ZTA, and breach intelligence—exposes students to the tools and taxonomies they will encounter in professional security roles, bridging the well-documented gap between academic cybersecurity education and operational security practice in India.

IX. ADVANTAGES OF THE PROPOSED APPROACH

- Fills documented research gaps by unifying eight security intelligence domains absent from any single reviewed system as a first-class architectural design requirement.
- Delivers AI-powered natural-language explainability across all modules, enabling analysts at all expertise levels to understand and act on security findings without deep specialist knowledge.
- Eliminates CVE hallucination risk through mandatory NVD API grounding, ensuring vulnerability details, CVSS scores, and patch availability information are factually accurate and current.

- Provides the first AI Red Team Agent implementation in the reviewed literature, enabling proactive adversarial simulation and kill-chain modelling without offensive security tooling or budget.
- MITRE ATT&CK alignment as a structural output constraint ensures all threat analyses are immediately actionable in professional SOC environments and compatible with existing threat intelligence workflows.
- Glassmorphic cyberpunk interface provides a consistent, professionally designed experience that reduces cognitive load and accelerates analyst orientation across all eight security modules.
- Stateless backend architecture enables cost-effective deployment on standard cloud hosting with no persistent database infrastructure requirement, accessible to student project budgets.
- Modular prompt engineering architecture allows new security intelligence modules to be added without changes to the core application architecture, providing long-term extensibility.

X. FUTURE SCOPE

- Real-Time SIEM Integration: Webhook endpoints to ingest live alert streams from Splunk, Elastic SIEM, and Microsoft Sentinel, enabling real-time AI-powered triage and cross-module correlation.
- Threat Hunting Automation: A proactive threat hunting module that accepts natural-language hypothesis statements and generates structured investigation queries for SIEM and EDR platforms.
- Compliance Framework Mapping: Extension of the Zero Trust Policy Builder to generate compliance gap analyses against SOC 2, ISO 27001, GDPR, and PCI-DSS with phased remediation roadmaps.
- Collaborative Analyst Workspace: Multi-user incident investigation workspaces with shared annotation, evidence tagging, and AI-assisted consensus building for distributed SOC teams.
- Mobile Companion Application: A React Native implementation providing push alert notifications, incident status monitoring, and rapid triage capabilities for on-call security personnel.
- Adversarial ML Defence Module: Integration of adversarial robustness testing for ML-based security models, evaluating evasion resistance against common perturbation attack techniques.
- Automated Report Generation: One-click generation of executive-level and technical incident reports in PDF format, incorporating all module findings from a given investigation session.

XI. CONCLUSION

This review paper has presented a comprehensive and critical survey of the state of AI-based cybersecurity monitoring research, with the objective of contextualising SecurAI Sentinel within the existing body of knowledge. Eight significant research contributions were reviewed spanning machine learning intrusion detection, the MITRE ATT&CK framework, dark web intelligence, LLM applications in security, zero-trust architecture, operational ML in cybersecurity, automated threat intelligence, and LLM-based vulnerability analysis.

The analysis reveals consistent technical progress in AI-based security operations yet identifies five persistent research gaps unaddressed by any reviewed system: the absence of unified multi-domain threat correlation, the exclusion of AI-driven red team simulation, dependence on static rule databases, the lack of natural-language explainability, and the absence of integrated zero-trust policy generation.

The proposed SecurAI Sentinel system addresses all five gaps through a coherent technical design combining eight AI-powered security intelligence modules, Google Gemini as the unified reasoning engine, real-time NVD API grounding for CVE data, MITRE ATT&CK alignment as a structural constraint, and a React/Node.js/Express.js full-stack architecture. The system is technically feasible with current commercial APIs, deployable on standard cloud hosting infrastructure, and designed to serve both educational and operational security purposes.

The work contributes to the broader discourse on AI in cybersecurity by demonstrating that the gap between sophisticated security research and practical, accessible deployment is bridgeable through prompt engineering, framework alignment, and user-centred design—without requiring custom model training, large proprietary datasets, or enterprise-scale infrastructure investment.

REFERENCES

[1] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7307098

[2] B. E. Strom et al., "MITRE ATT&CK: Design and Philosophy," MITRE Corporation, Technical Report MTR190051, 2018. [Online]. Available: https://attack.mitre.org

[3] S. Samtani, R. Chinn, H. Chen, and J. F. Nunamaker, "Exploring Emerging Hacker Assets and Key Hackers for Proactive Cyber Threat Intelligence," Journal of Management Information Systems, vol. 34, no. 4, pp. 1023–1053, 2020.

[4] M. A. Ferrag et al., "Revolutionizing Cyber Threat Detection with Large Language Models," IEEE Access, vol. 12, pp. 19895–19936, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10418079

[5] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, National Institute of Standards and Technology, 2020. [Online]. Available: https://doi.org/10.6028/NIST.SP.800-207

[6] G. Apruzzese et al., "The Role of Machine Learning in Cybersecurity," IEEE Transactions on Information Forensics and Security, vol. 18, pp. 2740–2773, 2023.

[7] E. Al-Shaer, A. Mohaisen, and W. Li, "Automated Cyber Threat Intelligence from Heterogeneous Sources," IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 6, pp. 3964–3980, 2022.

[8] R. Bhatt, A. Kumar, and S. Sharma, "LLM-Powered Automated Vulnerability Triage and Security Report Generation," arXiv preprint arXiv:2401.15896, Jan. 2024. [Online]. Available: https://arxiv.org/abs/2401.15896

[9] National Vulnerability Database, "NVD API Documentation and CVE Feed," National Institute of Standards and Technology, 2024. [Online]. Available: https://nvd.nist.gov/developers

[10] T. Hunt, "Have I Been Pwned: Breach Intelligence API," HaveIBeenPwned.com, 2024. [Online]. Available: https://haveibeenpwned.com/API/v3

[11] MITRE Corporation, "MITRE ATT&CK STIX Data and Navigator," Official Portal, 2024. [Online]. Available: https://attack.mitre.org/resources/

[12] Google LLC, "Google Generative AI SDK: Gemini API Reference Documentation," Google AI for Developers, 2024. [Online]. Available: https://ai.google.dev/docs

[13] M. A. Wynn, A. Novak, and P. Haas, "Threat Modeling and Attack Tree Analysis for AI-Integrated Security Systems," ACM SIGSAC, 2023.

[14] CISA, "Known Exploited Vulnerabilities Catalog," Cybersecurity and Infrastructure Security Agency, 2024. [Online]. Available: https://www.cisa.gov/known-exploited-vulnerabilities-catalog

[15] R. A. Olsson, "Network Traffic Analysis and Anomaly Detection: A Survey," Journal of Network and Computer Applications, vol. 119, pp. 1–17, 2023.

[16] Ministry of Electronics and Information Technology, "National Cyber Security Policy and Digital India Cybersecurity Framework," Government of India, 2013. [Online]. Available: https://www.meity.gov.in