**The Future is i/Thelma : Supercharging Threat Hunting with Large Language Models (LLMs)**

According to CrowdStrike’s and CISO 2025 report, security teams are overwhelmed by a relentless flood of alerts leading to alert fatigue and increasing the risk of missing critical threats. The relentless evolution of cyber threats demands a paradigm shift in defensive strategies. Traditional, reactive security measures are no longer sufficient to counter the sophistication and speed of modern attacks. Modern cyberattacks are engineered to bypass traditional security defenses like standard threat detection. Today’s attackers are highly skilled, leveraging advanced techniques that result in frequent, sophisticated intrusions. These threats often remain undetected for months, quietly operating within networks and increasing the risk of serious breaches [1]. Machine learning (ML) and Artificial Intelligence (AI) are driving the evolution of cybersecurity enhancing capabilities, accelerating threat response, and redefining how organizations defend against modern attacks [2][3][4]. Generative AI (GenAI), a subset of AI capable of creating new content, is emerging as a powerful ally in this fight.

***From Reactive to Proactive: A New Era of Threat Hunting***

Threat hunting can be either reactive triggered by alerts or proactive where analysts actively search for potential threats before any warning signs appear [5]. Threat hunting as a proactive is way to find hidden threats like ransomware, zero-day attacks, and APTs, that traditional security tools often miss [6]. Cyber Threat Intelligence (CTI) helps make threat hunting more proactive by pointing security teams toward likely threats. **In previous research [7], the definition of CTI in the organizational context** refers to the knowledge and insights about real or potential cyber threats that guide security decision-making. This intelligence often includes information about threat actors their goals, tactics, capabilities, and weaknesses. Organizations use CTI to support planning, threat analysis, situational awareness, and to anticipate future cyber events. In general CTI involves gathering and analyzing information to understand and mitigate cyber threats. It often includes insights into attackers’ **Tactics, Techniques, and Procedures (TTPs).** The methods they use to plan, launch, and carry out attacks [8]. Several commercial tools support Cyber Threat Intelligence (CTI). **Recorded Future, Anomali ThreatStream,** and **ThreatConnect** offer real-time insights, threat data aggregation, and automated response. **Mandiant Advantage** and **IBM X-Force Exchange** provide intelligence based on global incident response and curated threat data. **CrowdStrike Falcon X** integrates CTI into endpoint protection, while **Cisco Talos** delivers actionable insights from Cisco’s global threat research. Traditional threat intelligence faces several limitations [9]. It often lacks **context**, making it hard to understand the full threat landscape. **Data overload** from massive information streams can overwhelm analysts. Many systems operate in **silos**, limiting collaboration and sharing across teams. It's also **resource-intensive** to implement and maintain. Other issues include **false positives/negatives**, poor **integration with existing security tools**, and heavy **reliance on signature-based detection**, which may miss advanced or unknown threats. **Traditional threat intelligence has its limits. Generative AI offers a smarter way forward.**  
By analyzing patterns, predicting threats, and enhancing decision-making, Gen AI helps security teams stay ahead of attackers. It’s not just an upgrade it’s a game-changer. LLM applications and threat intelligence can be applied across various areas of cybersecurity : Threat Intelligence for Automated Reasoning [10], This research introduces a novel approach to real-time cyber threat detection by combining Large Language Models (LLMs), specifically GPT-4o, with Retrieval-Augmented Generation (RAG) and live threat intelligence feeds. By integrating data from sources like CVE, CWE, EPSS, and KEV using the Patrowl framework and Milvus for vector search, the system overcomes the static limitations of traditional models. The results show improved detection of emerging threats, making this a strong step toward automated, intelligent cybersecurity systems.; another resarch mention about **CTI systems that use LLMs and NLP models** can be **vulnerable to adversarial attacks [11],** This study examines the **security risks of using LLMs in CTI pipelines,** focusing on how attackers can manipulate these systems by injecting **fake or misleading text**. It identifies three types of attacks **evasion, flooding**, and **poisoning**. shows how these can degrade the accuracy and reliability of CTI systems. Special attention is given to **evasion attacks**, which can bypass detection and open the door for more damaging threats.; and MAD-CTI that focus on automatically scrape, analyze, and classify dark web content related to vulnerabilities, malware, and hacking based on a multi-agent framework powered by Large Language Models (LLMs) [12]. **GenAI is reshaping CTI** by enabling faster, smarter detection of emerging threats. Research has shown how LLMs, real-time data feeds, and multi-agent frameworks can automate threat analysis across sources like the surface web and dark web. While these advances improve scalability and accuracy, they also introduce risks such as adversarial attacks and misinformation.

***Beyond the Alarm: How AI Became Threat Hunting's Smartest Partner***

In the world of cybersecurity, for too long we've been playing defense. We wait for an alarm to blare. a digital tripwire before springing into action. But what about the intruders who never touch the tripwire? The silent threats that move through a network like a ghost? To catch them, you need to go hunting. And in 2025, the best hunters have an artificially intelligent partner.

Think of a traditional security analyst as a detective at a massive crime scene, buried under mountains of evidence and false leads. Threat hunting is the proactive process of sifting through that evidence to find clues of a crime *in progress*, not just one that's already happened. It’s a brilliant but exhausting task.

This is where AI changes the game for Threat Hunting:

1. OTuHunt [13]

**OTuHunt** is an automated threat hunting and cyber threat intelligence platform specifically designed for **Operational Technology (OT)** and **Industrial Control Systems (ICS)** environments. It leverages advanced **Natural Language Processing (NLP)** techniques to extract **Indicators of Compromise (IoCs)** and **Tactics, Techniques, and Procedures (TTPs)** from unstructured Cyber Threat Intelligence (CTI) reports. These extracted elements are then mapped to the **MITRE ATT&CK for ICS** framework and transformed into **SIEM-compatible queries**, enabling faster and more accurate threat detection. The platform consists of two integrated modules: the **MSSP Platform**, which automates the CTI extraction and report generation process for managed security providers, and the **Threat Hunter Platform**, which assists security analysts in creating detection queries and finalizing threat hunting reports. OTuHunt addresses the critical need for OT-specific, proactive threat detection by bridging the gap between traditional IT-based threat hunting tools and the unique requirements of industrial environments.

1. ThreatScout [14]

**ThreatScout** as a comprehensive solution to the challenges of modern threat hunting. ThreatScout is an automated, closed-loop threat hunting system that leverages machine reasoning to proactively detect advanced and stealthy cyber threats. The system is designed to automate the core steps of threat hunting: it begins by automatically generating threat hypotheses based on incoming system telemetry such as logs, alerts, and observed behaviors using rule-based reasoning powered by the RETE algorithm. These hypotheses are then tested through case-based reasoning, which selects or constructs appropriate test flows to guide targeted data collection across relevant network nodes. Once sufficient data is gathered, the hypotheses are validated using a separate set of rules and the same reasoning engine, confirming or refuting potential threats. ThreatScout also incorporates an iterative feedback mechanism that refines hunting rules over time, allowing the system to adapt to evolving attack patterns. Architecturally, it consists of two main modules: the Data Knowledge Module, which manages threat intelligence, system abstractions, and rule sets; and the Analysis Module, which includes intelligent agents for hypothesis generation, testing, and validation. By integrating seamlessly with existing SOC tools such as SIEM and SOAR, ThreatScout aims to reduce manual workload, improve detection accuracy, and shift cybersecurity operations toward a more proactive and intelligent defense model.

1. TRUSTY [15]

TRUSTY (Threat hUnting using data analySiS in criTical infrasTructures sYstems) is a web-based cybersecurity platform designed to enhance threat detection within Industrial Internet of Things (IIoT) environments, particularly in critical infrastructure systems. It leverages industrial honeypots that emulate common industrial communication protocols such as Modbus/TCP, IEC 60870-5-104, BACnet, MQTT, and EtherNet/IP to attract and analyze malicious cyber activity. Unlike traditional Intrusion Detection and Prevention Systems (IDPS), which often struggle with zero-day attacks and false positives, TRUSTY enhances detection by combining honeypot data with advanced analytics. The platform uses a strategic deployment mechanism modeled as a Multi-Armed Bandit (MAB) problem and solved using the Thompson Sampling (TS) algorithm. This approach helps determine the optimal number of honeypots to deploy, balancing the trade-offs between attacker and defender strategies and minimizing resource use while maximizing protection. TRUSTY’s architecture includes a Honeypot Sensor layer for data collection and a Honeypot Analyzer Server that processes and visualizes data using tools like Logstash, Elasticsearch, and Kibana. The platform provides rich contextual insights such as geolocation, network and application-layer information, and calculated risk levels. Additionally, it offers a publicly available dataset containing logs, network traffic, and statistics from honeypot activity, which can be used for further research or intrusion detection training. Evaluation results show that TRUSTY’s deployment strategy achieves significantly higher accuracy up to 89% compared to random selection methods. Overall, TRUSTY represents a comprehensive and intelligent solution for proactive threat hunting in critical IIoT systems.

1. **Threat Trekker [16]**

**Threat Trekker** is an intelligent cyber threat hunting system designed to proactively detect and classify security threats in complex network environments. Unlike traditional security tools that react to known attack signatures, Threat Trekker assumes attackers may already be inside the network and continuously scans for unusual patterns or anomalies in log and traffic data. It uses **Random Forest**, a powerful machine learning algorithm, to analyze features such as packet sizes, durations, IP addresses, and connection states. This algorithm was chosen for its high accuracy, robustness to overfitting, and ability to handle both structured and imbalanced data—common in cybersecurity scenarios. By integrating with data connectors and streaming platforms like Apache Kafka, Threat Trekker enables real-time analysis and feeds insights back into systems like SIEMs (Security Information and Event Management), enhancing detection capabilities and enabling quicker, more informed responses. This makes it a scalable and adaptive solution for modern cybersecurity defense.

The evolution of AI-driven threat hunting tools such as OTuHunt, ThreatScout, TRUSTY, and Threat Trekker illustrates the profound shift in cybersecurity from reactive defense to intelligent, proactive threat mitigation. These systems, each with their unique strengths ranging from OT-specific NLP extraction and hypothesis-driven machine reasoning to strategic honeypot deployment and anomaly-based machine learning demonstrate how AI has become an indispensable partner in uncovering sophisticated, stealthy threats that traditional systems often miss. Together, they embody the core message of Beyond the Alarm: that AI doesn’t just sound the alert it collaborates with human analysts, anticipates attacker behavior, and enables a smarter, faster, and more adaptive approach to threat hunting across diverse digital terrains.

***i/Thelma : An Agent Framework for Threat Hunting Automation***

Despite advancements in cybersecurity tools and frameworks, threat hunting remains a largely manual and expert-driven process. Traditional systems rely heavily on indicators of compromise (IOCs), which are rigid and easily evaded by sophisticated attackers. While artificial intelligence and machine learning have been introduced to aid detection, these models often require large datasets, lack interpretability, and are not flexible enough to adapt to evolving threats. Moreover, existing automated threat hunting tools focus primarily on detection and analysis, but fail to address the "last-mile" tasks such as formulating threat hypotheses and crafting customized hunting queries tasks that still demand significant human expertise. The heterogeneous nature of computing environments, along with diverse data formats and query languages, further complicates the automation process. As a result, there is a critical need for a solution that can automate these final, labor-intensive steps of the threat hunting process while remaining adaptable, interpretable, and effective in complex environments.

Thelma (Threat-Hunting Enhanced Language Models for Hunt Automation) is an AI-powered agent framework designed to automate key steps in the threat hunting process (see Figure 1). It uses two large language models (LLMs): one to **generate custom threat hunting queries** from natural language descriptions (LLM A), and another to **prioritize which systems to investigate** based on current threat status (LLM B). By reading threat playbooks and adapting to different computing environments, Thelma can craft relevant queries and guide security analysts efficiently. This automation helps reduce manual effort, speeds up response time, and supports less experienced analysts in detecting complex cyber threats.

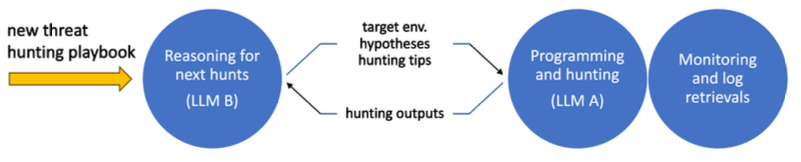


Figure 1. Thelma Agent Framework for Threat Hunting Automation

Thelma showed promising results in automating threat hunting tasks. The first model, LLM A, which generates threat hunting queries, achieved high accuracy, producing correct results in 19 out of 26 cases (about 73% accuracy). It performed best when clear and detailed threat descriptions were provided. The second model, LLM B, which prioritizes endpoints for investigation, had low prediction accuracy (around 10%), but maintained a high recall rate (over 60%), meaning it was effective at minimizing missed threats. Overall, Thelma effectively automates query generation but needs further refinement in reasoning and environment-based decision-making.

Thelma demonstrates that large language models can automate key parts of the threat hunting process, especially generating accurate, context-aware queries from natural language. This reduces the manual workload on security analysts and speeds up response times. While query generation is effective, the system's reasoning model for prioritizing threats needs improvement. With more training data and richer contextual inputs, Thelma has strong potential to become a valuable tool for scalable, AI-driven cybersecurity operations.

**Reference**

[1] B. Nour, M. Pourzandi, and M. Debbabi, “A Survey on Threat Hunting in Enterprise Networks,” *IEEE Commun. Surv. Tutorials*, vol. 25, no. 4, pp. 2299–2324, 2023, doi: [10.1109/comst.2023.3299519](https://doi.org/10.1109/comst.2023.3299519).

[2] “Artificial Intelligence and Cybersecurity: Past, Presence, and Future,” in *Advances in Intelligent Systems and Computing*, Singapore: Springer Singapore, 2020, pp. 351–363. doi: [10.1007/978-981-15-0199-9\_30](https://doi.org/10.1007/978-981-15-0199-9_30).

[3] H. Binyamini, R. Bitton, M. Inokuchi, T. Yagyu, Y. Elovici, and A. Shabtai, “An Automated, End-to-End Framework for Modeling Attacks From Vulnerability Descriptions,” 2020, *arXiv*. doi: [10.48550/ARXIV.2008.04377](https://doi.org/10.48550/ARXIV.2008.04377).

[4] A. Boukhtouta, T. Madi, M. Pourzandi, and H. A. A., “Cloud Native Applications Profiling using a Graph Neural Networks Approach,” in *2022 IEEE Future Networks World Forum (FNWF)*, Montreal, QC, Canada: IEEE, Oct. 2022. doi: [10.1109/fnwf55208.2022.00046](https://doi.org/10.1109/fnwf55208.2022.00046).

[5] R. Puzis, P. Zilberman, and Y. Elovici, “ATHAFI: Agile Threat Hunting And Forensic Investigation,” 2020, *arXiv*. doi: [10.48550/ARXIV.2003.03663](https://doi.org/10.48550/ARXIV.2003.03663).

[6] S. Singh, P. K. Sharma, S. Y. Moon, D. Moon, and J. H. Park, “A comprehensive study on APT attacks and countermeasures for future networks and communications: challenges and solutions,” *J Supercomput*, vol. 75, no. 8, pp. 4543–4574, Aug. 2019, doi: [10.1007/s11227-016-1850-4](https://doi.org/10.1007/s11227-016-1850-4).

[7] S. Ainslie, D. Thompson, S. Maynard, and A. Ahmad, “Cyber-threat intelligence for security decision-making: A review and research agenda for practice,” *Computers & Security*, vol. 132, p. 103352, Sep. 2023, doi: [10.1016/j.cose.2023.103352](https://doi.org/10.1016/j.cose.2023.103352).

[8] A. Mahboubi *et al.*, “Evolving techniques in cyber threat hunting: A systematic review,” *Journal of Network and Computer Applications*, vol. 232, p. 104004, Dec. 2024, doi: [10.1016/j.jnca.2024.104004](https://doi.org/10.1016/j.jnca.2024.104004).

[9] “Generative AI for Threat Intelligence and Information Sharing,” in *Advances in Digital Crime, Forensics, and Cyber Terrorism*, IGI Global, 2024, pp. 191–234. doi: [10.4018/979-8-3693-8944-7.ch006](https://doi.org/10.4018/979-8-3693-8944-7.ch006).

[10] S. Paul, F. Alemi, and R. Macwan, “LLM-Assisted Proactive Threat Intelligence for Automated Reasoning,” 2025, *arXiv*. doi: [10.48550/ARXIV.2504.00428](https://doi.org/10.48550/ARXIV.2504.00428).

[11] S. Shafee, A. Bessani, and P. M. Ferreira, “False Alarms, Real Damage: Adversarial Attacks Using LLM-based Models on Text-based Cyber Threat Intelligence Systems,” 2025, *arXiv*. doi: [10.48550/ARXIV.2507.06252](https://doi.org/10.48550/ARXIV.2507.06252).

[12] S. Shah and V. K. Madisetti, “MAD-CTI: Cyber Threat Intelligence Analysis of the Dark Web Using a Multi-Agent Framework,” *IEEE Access*, vol. 13, pp. 40158–40168, 2025, doi: [10.1109/access.2025.3547172](https://doi.org/10.1109/access.2025.3547172).

[13] F. Alaliwat et al., "OTuHunt: An Aggregated Threat Hunting & Intelligence Platform for OT/ICS Environment and MSSP Services," 2025 12th International Conference on Information Technology (ICIT), Amman, Jordan, 2025, pp. 39-46, doi: 10.1109/ICIT64950.2025.11049292.

[14] Z. M. Ferdjouni, B. Nour, M. Pourzandi, and M. Debbabi, “ThreatScout: Automated Threat Hunting Solution Using Machine Reasoning,” *IEEE Secur. Privacy*, pp. 2–13, 2024, doi: [10.1109/msec.2024.3492132](https://doi.org/10.1109/msec.2024.3492132).

[15] P. Radoglou-Grammatikis *et al.*, “TRUSTY: A Solution for Threat Hunting Using Data Analysis in Critical Infrastructures,” in *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, Rhodes, Greece: IEEE, Jul. 2021, pp. 485–490. doi: [10.1109/csr51186.2021.9527936](https://doi.org/10.1109/csr51186.2021.9527936).

[16] Á. C. Bienzobas and A. Sánchez-Macián, “Threat Trekker: An Approach to Cyber Threat Hunting,” 2023, *arXiv*. doi: [10.48550/ARXIV.2310.04197](https://doi.org/10.48550/ARXIV.2310.04197).