Symposium On Usable Privacy and Security (SOUPS), 2026

Like a Hammer, It Can Build, It Can Break: Large Language Model Uses, Perceptions, and Adoption in Cybersecurity Operations on Reddit

Souradip Nath, Chih-Yi Huang, Aditi Ganapathi, Kashyap Thimmaraju
Jaron Mink, and Gail-Joon Ahn

ASU Arizona State University

Technische Universität Berlin

August 25, 2026

The 2025 Pivot: The Rise of Agentic Security

Cybersecurity in 2025: Agentic AI to change enterprise security and business operations in year ahead

Cybersecurity at AI speed: How agentic AI is supercharging SOC teams in 2025





Microsoft Google Cloud IBM CISCO paloalto NETWORKS

CROWDSTRIKE DARKTRACE

Dropzone AI torq arcanna CML ZERO twine prophet

Research Gap: The Bigger Picture

- **Prior works on ML, AI and LLMs in SOCs incorporate human perspective**
  - Small-scale investigation through in-depth interview studies [1,4,5]
  - Scoped around a specific tool [2] , task [3] or aspect [4, 5]

- **Broader perspective across tools, use cases, and adoption is missing**
  - How LLM tools are adopted and perceived at scale and longitudinally?

- **Our Work: A large-scale, multi-dimensional Reddit-based investigation of practitioner practices, opinions and experiences.**
  - RQ1: LLM tools and use cases
  - RQ2: Perceptions of LLM tools
  - RQ3: Adoption patterns

[1] "Everybody's Got ML, Tell Me What Else You Have: Practitioners' Perception of ML-Based Security Tools and Explanations." Mink et al., IEEE S&P 2023
[2] "LLMs in the SOC: An Empirical Study of Human-AI Collaboration in Security Operations Centers." Singh et al., IEEE S&P 2026
[3] "Integrating Large Language Models into Security Incident Response." Kramer et al., SOUPS 2025
[4] "Navigating Autonomy: Unveiling Security Experts' Perspectives on Augmented Intelligence in Cybersecurity." Roch et al., SOUPS 2024
[5] "Too Much to Trust? Measuring the Security and Cognitive Impacts of Explainability in AI-Driven SOCs." Rastogi et al., arXiv 2025

Our Research Methodology

Reddit Discourse Analysis

892 Posts | **76** Threads | **3** Subreddits

r/cybersecurity
r/Information_Security
r/ciso

33
Months

Dec 2022 – Aug 2025

- **Subreddit Discovery** through Topical Relevance
- Keyword-based **Thread Extraction (n=1,310)**
- LLM-assisted **Thread Relevance Classification (n=76)**
- **Post Collection (n=1,703)**
- Manual **Post Relevance Classification (n=892)**
- **Qualitative Analysis of Posts** through Deductive and Inductive Double Coding, followed by a Reflexive Thematic Analysis
- **Quantitative Analysis of Codes** to assess statistical patterns within prevalences and associations between codes

RQ1 Findings: Tools

Which LLM tools are gaining more attention?



Claude

Copilot

- **General-purpose LLMs > Security-focused LLMs**
  - Tools, such as, ChatGPT, Copilot, and Claude are significantly more referenced over enterprise-grade security solutions, such as, Security Copilot, Dropzone, Intezer
- **A growing but fragmented ecosystem of AI-for-Cybersecurity Vendors**
  - 30 distinct security-focused tools were mentioned, 15 of them only once
- **In-house custom workflows are rare, but exist**
  - Only 3.4% posts mentions of custom LLM-driven pipelines for automating SIEM workflows, with custom python scripts and open-source LLM frameworks

> *I have developed **custom Python scripts** using **open-source agent frameworks** to feed SIEM query results into an agent to iteratively process the data and generate a consolidated review of the search output."*

RQ1 Findings: Use Cases

How LLM tools are being used within SOC workflows?

- Security-focused LLMs integrated in **Triage and Response** workflows
- General-purpose LLMs **for Scripting, Reporting,** and **Knowledge Support**
  - Code generation and query debugging
  - Technical documentation and reporting
  - Learning & faster information retrieval

Table 3: **Reported LLM Tools Use Cases Across Tool Categories (General-Purpose, Security-Specific).**

Types of LLM Use Cases	# General	# Security	*p*-value
Triage & Incident Response (*n*=139)	**30 (40.54%)**	**44 (59.46%)**	**< 0.001***
Scripting & Query Support (*n*=88)	**51 (89.47%)**	**6 (10.53%)**	**< 0.01***
Reporting & Documentation (*n*=84)	**43 (89.58%)**	**5 (10.42%)**	**< 0.05***
Threat Analysis (*n*=54)	16 (57.14%)	12 (42.86%)	0.948
Knowledge Support (*n*=53)	31 (88.57%)	4 (11.43%)	0.119
Training, Compliance, Others (*n*=18)	9 (81.82%)	2 (18.18%)	–

Counts and percentages are reported row-wise and represent the relative distribution of tool types within each use case category. Statistical significance was assessed using a chi-square test of independence followed by Bonferroni-corrected post-hoc analysis of adjusted residuals. Rows denoting significant associations are bolded.

RQ1 Findings: Use Cases

How LLM tools are being used within SOC workflows?

- **From Investigation Buddies to Fully Autonomous Triage and Response**
  - Uses of LLM tools demonstrate a varying level of autonomy granted to the LLM tools.

> *During a complex investigation, I utilize **LLMs as a companion**, throwing ideas at them. I **let them examine the data** to help with correlation, and I handle the critical thinking.... I often **provide logs, screenshots, or event timelines** to LLMs to help me piece together what might be happening, either to **validate my findings**, or help me **zoom in on a problem area.**"*

- **Decision Support > Human-in-the-loop > Fully Autonomous Integration**

RQ2 Findings: Perceptions of LLM Tools







Capabilities



Reliability



Autonomy



Efficiency



Security & Privacy



Cost

Table 4: **Reported LLM Factors and Opinions.**

Factors	# Positive	# Negative	*p*-value
Capabilities	**174 (64.21%)**	97 (35.79%)	**< 0.001***
Efficiency	**54 (84.38%)**	10 (15.62%)	**< 0.001***
Reliability	3 (5.17%)	**55 (94.83%)**	**< 0.001***
Security & Privacy	6 (11.54%)	**46 (88.46%)**	**< 0.001***
Autonomy	6 (12.77%)	**41 (87.23%)**	**< 0.001***
Cost	5 (14.29%)	**30 (85.71%)**	**< 0.001***

Percentages are reported row-wise and represent the relative proportion of opinions within each LLM factor. Statistical significance was assessed using a chi-square test followed by Bonferroni-corrected post-hoc analysis of adjusted residuals. Rows denoting significant associations are bolded.

RQ2 Findings: Perceptions of LLM Tools

Why Practitioners are Excited about LLM tools for SecOps?

- Better contextualization and visibility of alerts
- Improved interpretability of signals and reduced FPs
- Quantifiable efficiency gains, better MTTT, MTTD, MTTR

> *I use Agentic LLMs, and **they're incredible**. Not only does it turn **half a million alerts/events a month into 1-3 relevant daily alerts**, but also **gives important context.** During an investigation, I can ask, 'Does this event fire every time a user logs in or is this a new alert?' or, 'Did they get challenged for MFA?'... **It's not just reducing my workload, it's finding things I couldn't see.**"*

RQ2 Findings: Perceptions of LLM Tools

Why Practitioners are still Cautious about LLM tools for SecOps?

- LLMs lacks precision and truth telling
- Hallucinations and non-determinism cause concern
- While improving, human oversight is still needed

> *In my professional circle, **we all agree that we will need to leverage AI**, but **only for advice, not for taking autonomous action… Use it like an intern**. Allow it to gather information and raise its hand when it notices something, but **don't let it touch anything critical.**"*

RQ2 Findings: Perceptions of LLM Tools

Why Practitioners are still Cautious about LLM tools for SecOps?

- Difficulties and uncertainties around data governance
- Security landscape of LLMs is still immature

> *It sounds like what all AI companies are trying to do:* ***get as much data to train models that they can sell to other customers.****"*

- LLM inference costs can become prohibitively expensive
- Unsustainable financial overhead with limited returns

> *Although Security Copilot has several interesting features,* ***it is still not worth the price.*** *Even minimal usage can be* ***equivalent to one full-time employee's salary.****"*

RQ3 Findings: Adoption

How does the adoption of LLM Tools look like?

- LLM adoption shifts from curiosity to operational use
- Non-Adoption is constant, driven by,
  - Inflated Vendor Promises
  - Sufficiency of Traditional Solutions
  - Organizational Restrictions
  - Anti-LLM Sentiment

Table 5: **Reported Adoption of LLM Tools Over Time.**

	Counts			p1 vs p2		p2 vs p3	
Adoption Stage	**p1**	**p2**	**p3**	**Delta**	***p*-val**	**Delta**	***p*-val**
Evaluating	20	16	87	+4.40%	1.000	**-33.65%**	**0.002***
Using	12	7	181	-7.36%	1.000	**+31.09%**	**0.013***
Not Using	3	3	44	+2.97%	1.000	+2.56%	1.000

p1: Dec 2022–Oct 2023, **p2**: Nov 2023–Sep 2024, **p3**: Oct 2024–Aug 2025
Pairwise comparisons were conducted using a chi-square test followed by Bonferroni-corrected post-hoc analysis of adjusted residuals. Bold values indicate statistically significant differences. Delta percentages are based on the proportions of adoption stages.

> *Organizations are eager to replace analysts with AI,* ***making it difficult to be enthusiastic*** *about technologies that may* ***ultimately reduce the need for human roles.***”

Implications and Future Directions

- **LLM Adoption is shaped by control and commitment**
  - Top-down vs. Bottom-up Adoption for general-purpose vs. security-focused LLMs
  - Warrants further investigation into LLM adoption through **Multi-Stakeholder Studies**
- **Reliability is the hard ceiling for autonomy**
  - Distrust is grounded in experience with hallucinations and inconsistent outputs
  - Unreliable Outputs → Verification Overhead → Lack of Autonomous Workflows
  - How **LLMs can better communicate their uncertainties** without subsidizing those with hallucinated outputs?

Key Takeaways

- General-purpose LLMs are readily adopted across a range of operational activities
- Awareness of the rapidly expanding security-focused LLM ecosystem remains fragmented
- LLM uses span a spectrum of autonomy, from decision support to rare autonomous workflows
- LLMs are capable and efficient, but not yet fully trusted without human oversight
- Adoption of LLM tools is accelerating, yet inflated vendor promises keep non-adoption steady

Thank you

Read the paper

Souradip Nath

snath8@asu.edu | souradipnath.com