The Impact of Emerging AI Practices on the Cybersecurity Workforce

Miuyin Yong Wong
University of Maryland

Alan Luo
University of Maryland

Yunze Zhao
University of Maryland

Shubham Bhatnagar
University of Maryland

Fabian Monrose
Georgia Institute of Technology

Michelle L. Mazurek
University of Maryland

Abstract

Large language models (LLMs) have frequently been proposed as useful tools for cybersecurity tasks like malware detection, reverse engineering, and incident response. However, little is known about whether and how cybersecurity practitioners have adopted LLMs in enterprise settings, how they perceive these tools and communicate about them with peers, and what benefits or challenges they encounter in practice. To fill this gap, we conducted 28 semi-structured interviews with cybersecurity professionals from 26 organizations, with varying experience levels. We find that although LLM use is common and frequent, organizational policies governing this use are often vague, inconsistent, and unclear to individual practitioners. Practitioners face important challenges related to the fundamental limitations of LLMs—such as unreliable outputs and context window limits—as well as domain-specific challenges including poor performance on specific cybersecurity tasks and restrictive guardrails that inhibit penetration testing. Other key concerns included data leakage and productivity loss. Beyond technical limitations, we find that confusion around informal guidance and limited peer communication is creating a psychological barrier, influencing practitioners' adoption of LLMs and altering the traditional collaborative culture and apprenticeship knowledge-transfer model of the cybersecurity community.

1 Introduction

Large language models (LLMs) are increasingly being applied in many disciplines, and cybersecurity is no exception. Extensive research has evaluated LLM effectiveness on cybersecurity tasks, such as reverse engineering [8], malware detection [17], and capture-the-flag exercises [67]. Though fully automated solutions are still limited, LLMs may still offer benefits to defenders' productivity and efficacy, which could prove especially important as cyberattacks continue to proliferate [22, 38] while cybersecurity professionals remain in short supply [11, 26].

However, there is limited understanding of how cybersecurity practitioners are adopting LLMs in real world settings and how this adoption may be shaping professional practice. Early work examines query data from one Security Operations Center (SOC) to understand how LLMs are used in daily work, finding that SOC analysts primarily used LLMs for sensemaking, context building, and low-level task support, while retaining human decision authority [53]. Other research examines practitioner perspectives on LLM use within software engineering workflows [21, 35]. Klemmer et al. [35] reported mistrust in AI outputs and a lack of standardized verification practices, whereas Giray et al. [21] highlight perceived benefits such as reduced cycle time and improved quality, despite limited objective productivity measures. Although these studies offer valuable insights, it remains unclear whether their findings extend to the broader cybersecurity landscape, such as incident response, threat intelligence, compliance, and digital forensics. Examining LLM adoption across a wider set of roles is essential to identifying potential systemic shifts in professional norms, collaboration patterns, and knowledge transfer that may not be visible from a single vantage point.

Moreover, many groups of cybersecurity practitioners rely on norms of openness, collaboration, and mentorship [19], as demonstrated through competitions such as capture-the-flag tournaments [5], bug report publishing [61], and online communities [60]. This makes it critical to understand how LLM use affects social dynamics in cybersecurity, such as how practitioners communicate about LLM use with their peers, how they perceive others' use of such systems, and the informal guidance and advice emerging around this technology.

To fill this gap, we conducted semi-structured interviews

USENIX Symposium on Usable Privacy and Security (SOUPS) 2026.
August 23–26, 2026, Hannover, Germany.

with 28 cybersecurity practitioners across 26 distinct organizations, with differing levels of cybersecurity experience and different job roles. We aim to answer the following research questions:

RQ1. What policies and guidance for LLM use are organizations providing to their cybersecurity practitioners?

RQ2. What challenges and risks do cybersecurity practitioners encounter when using LLMs, and what strategies do they employ to address or mitigate them?

RQ3. How do peers' shared perceptions, advice, and informal guidance regarding LLMs shape cybersecurity practitioners' adoption and use of LLMs?

Our findings indicate that many organizations lack formal policies on LLM use, and informal guidance by peers is often inconsistent. Yet, participants still reported that their organizations encouraged the use of these models due to expected increases in productivity. Additionally, our participants reveal cybersecurity-specific obstacles associated with LLMs, such as guardrails restricting "offensive" tasks like penetration testing and red-teaming or models being unable to keep up with the quickly-evolving cybersecurity landscape. To our surprise, few participants expressed concern regarding potential adversarial threats, such as data poisoning or prompt injection. This raises the concern that such threats are not yet front of mind for security practitioners, nor are they preparing for them.

Our most significant finding is a psychological barrier that cybersecurity practitioners are experiencing: Participants voiced a fear of being judged for using AI tools, with potential harm to their professional reputations. While this anxiety rarely stops practitioners from using LLMs, we begin to observe a shift in social and interpersonal dynamics within the community. These tensions highlight a pressing need for the security community to engage in a broader dialogue about the role of generative AI, to ensure that we preserve the collaborative practices that have long supported the field.

2 Related Work

LLMs in Cybersecurity. There has been significant effort put into evaluating LLMs for their potential application towards solving or supporting security tasks such as penetration testing [2, 14, 25], phishing simulations [7, 28], threat report management [12], incident response [32], vulnerability detection and repair [15, 18, 52], reverse engineering [8], and anomaly detection [27, 54]. However, the majority of these evaluations have been conducted in controlled environments and have primarily focused on performance metrics. Some prior work has examined the application of LLMs for security tasks from a human-in-the-loop perspective [3, 44, 54, 55], but these studies often only propose frameworks for human-AI collaboration or focus on non-expert/non-professional users.

Although some prior work has examined the impact of AI on the broader cybersecurity landscape [23], there is limited prior research examining how individual practitioners are using and integrating this technology into their day-to-day workflow. Singh et al. [53] conducted the most relevant study to date, analyzing queries from 45 SOC analysts from a single institution over a period of 10 months with the goal of understanding how they adopt LLMs into their daily practices. While this approach yields valuable evidence about SOC analysts' technical use patterns, it is limited in scope and does not capture practitioners' subjective evaluations, sources of frustration, or concerns about trust and accountability. To our knowledge, ours' is the first study to directly examine cybersecurity professionals' perceptions of their own use of LLMs.

Impacts of LLM Usage in Software Engineering. Significant effort has been expended by the community to investigate the performance of LLMs on various software engineering tasks and to systematize this knowledge [16, 29, 66]. More specifically, prior work has focused on evaluating productivity gains for tasks including routine coding [42, 58, 63], template generation, and general library reference inquiries [16, 31, 45, 50, 56]. Additionally, there have also been efforts to examine how developers are adopting generative AI tools in practice, using surveys and interviews [21, 35] to identify common use cases, perceived benefits, challenges, and expectations.

The impacts of LLMs on software engineering and engineers can provide us valuable insights into how security professionals and the security community may similarly be affected. However, security professionals often operate in higher-stakes settings and their decisions and outputs carry significant security implications for their organizations and others. Our work takes an important step forward in capturing perceptions and practices of those security professionals and potentially towards defining norms and standards for AI/LLM usage within the broader security community.

Collaborative Culture in Cybersecurity. The cybersecurity community has a long-standing tradition of resource sharing, collaboration, and collective sensemaking, particularly in support of newcomers and the adoption of new technologies. Community members often share free tools and learning resources [62, 65], engage in online discussions [64], and even provide free infrastructure for the purposes of education [9, 57]. Even in competitive environments, such as vulnerability discovery [60], this prosocial behavior is common, even if it might result in reduced personal gains.

This resource and information sharing behavior extends to large organizations, which often share information relating to cyber risk or threats strategically to collectively gain defensive benefits [36, 37, 46]. Collaboration is also core to different practitioners' motivation to participate, including those in bug bounty hunting [1, 4, 24, 48], security and system administration [30], adversarial cyber testing [41], and vulnerability discovery [43]. Entry into the cybersecurity community is

not always easy or accessible [19], and newcomers often face challenges that mirror those in broader society [20, 34], making participation in the wider community critical to progress.

AI adoption has the potential to reshape these social dynamics, which have long underpinned how the cybersecurity community learns, collaborates, and grows, yet to our knowledge, prior work has not examined this impact. Our work seeks to capture the early impacts and perceptions of community members in this formative stage of LLM adoption.

3 Method

We conducted 28 semi-structured interviews with cybersecurity professionals. This section describes our recruitment methods, interview protocol, and analysis process.

3.1 Recruitment

To understand how the emergence of LLMs are impacting the cybersecurity workforce, we recruited participants from a wide range of cybersecurity roles. Recruitment was conducted through cybersecurity mailing lists, LinkedIn posts, personal networks, and snowball sampling from May to September 2025. We did not restrict recruitment based on specific responsibilities, prior experience with LLMs, or years of experience in the field. The only inclusion criterion was holding a cybersecurity position within an organization. This broad recruitment approach allowed us to capture diverse practitioner perspectives and explore potential systemic or community-level shifts in practice, collaboration, and culture.

Although this approach was intended to attract a diverse set of participants, all participants recruited during this first phase reported having more than three years of professional experience and made regular use of LLMs. To obtain a more diverse set of perspectives, particularly from early career practitioners, we launched a second round of recruitment. We used the same channels, as well as alumni mailing lists from universities with cybersecurity programs, to reach professionals with less than 3 years of experience.

Screening Survey. All participants completed a prescreening survey to confirm eligibility. Participants had to be at least 18 years old and currently employed in a cybersecurity position. The survey collected information about the participant's job role, LinkedIn profile, background, and organization details. Eligibility verification required a valid LinkedIn account, which was necessary to filter out bots and prevent participant fraud, and a job description consistent with cybersecurity responsibilities. Responses that did not meet these criteria were not invited to interview. Other information collected in this survey (e.g., technical background or employing organization) was not used to restrict participation but was retained for later use in data analysis and presentation of results.

Ethics Statement. This study was approved by the University of Maryland Institutional Review Board (IRB). Participants' names, LinkedIn accounts, and email addresses were initially collected in order to schedule interviews. Identifying information was stored separately from the interview data and deleted immediately after participants were compensated. During the transcription process, all personal identifiers were also anonymized to ensure confidentiality. Informed consent was provided in the screening survey and again immediately before each interview. Participants were informed that they were permitted to withdraw from the study at any point during the interview and that they could choose not to answer any question if they so chose. In addition, we also never asked participants to disclose sensitive information about their work or organizational security environment. Each participant was compensated with a $50 Amazon gift card.

3.2 Semi-Structured Interviews

We conducted one-hour-long, semi-structured interviews remotely through Zoom. All interviews were conducted by the first author for consistency. The interviewer followed a detailed guide to confirm consent, avoid asking leading questions, and confirm contact information for compensation. The full interview protocol (available in Appendix A) was organized into the following two sections:

Organizational Policy & Informal Guidance. The interview began with introductory questions asking participants when and why they began using LLMs for cybersecurity-related tasks. We then asked whether they had received any formal or informal training on AI use (e.g., prompt engineering). Participants were also asked about their observations of peers' LLM use and how LLM-related knowledge and practices were shared among colleagues. Those with more experience in the field were asked if they would recommend LLM use to less-experienced practitioners, while less-experienced practitioners were asked if they were encouraged to use LLMs. We also asked about informal guidelines and workplace culture surrounding LLM use.

Use Cases, Challenges & Perception of Risk. Participants were asked to provide concrete examples of cybersecurity tasks in which they had integrated LLMs. We then asked participants to describe how they verify and evaluate LLM outputs, followed by questions about their confidence in these processes and challenges that they encounter. Participants were also asked to reflect on perceived risks associated with LLM use. The interview concluded with questions about participants' beliefs about the future of LLMs in cybersecurity.

To refine the protocol, we conducted three pilot studies. The first revealed a few leading and unclear questions, which were revised. The second and third pilots showed no further

issues, so they were retained in our final dataset.

3.3 Thematic Analysis

All interviews were transcribed using Zoom's automatic transcription; errors were corrected manually; and any PII was redacted. For a subset of interviews where Zoom's transcription failed, audio was transcribed using OpenAI's `whisper-large-v2` model, run entirely locally on institutional hardware with no data transmitted to any external party.

The transcripts were analyzed using thematic analysis [10]. To develop the initial codebook, three researchers independently coded five interviews. The researchers then discussed initial themes and codes, arriving at a canonical codebook. [1] The remaining interviews were then doubly-coded using this codebook. The first author coded all interviews for consistency, while the other three researchers split the remaining interviews. The same codebook was applied to both participant groups (more experienced and less experienced) since the interview protocol only differed by two questions.

Whenever new codes or themes emerged, the research team met to discuss and reconcile these codes, after which the first author recoded relevant interviews to maintain consistency. Broad thematic saturation stabilized after coding nine participants, which included a mix of both more and less experienced participants. However, given the exploratory nature of this study, the detailed codebook, and the diversity of participants' responsibilities, more detailed sub-themes continued to emerge throughout the coding process.

We did not calculate inter-rater reliability (IRR) since themes and codes were developed collaboratively and iteratively, and the codes for each interview discussed and reconciled in meetings to ensure agreement [40].

3.4 Threats to Validity

As with any interview study, there are limitations that should be considered when interpreting our findings. Our semi-structured interview protocol asked participants to recall their past experiences with LLMs; it is possible that participants may have forgotten about or misremembered some instances. We opted not to attempt direct observation of workflows, given that many security professionals work with sensitive data or in secure or restrictive settings.

In addition, participants may have been less inclined to describe experiences where they misused LLMs or if they perceived their usage as potentially embarrassing or stigmatizing—for instance, if their usage could be interpreted as over-reliance. We attempted to address this by using non-judgmental language and assuring our participants that their experiences would be anonymized and kept confidential, in order to encourage them to share honestly. We also never asked participants to disclose sensitive information about their work or organizational security environment.

Finally, our participant pool was small but varied. This allowed us to explore perspectives from a range of different cybersecurity job roles, experience levels, and organization sizes, but (as with any qualitative study) our findings may not be generalizable to the broader cybersecurity community, and especially to job roles/responsibilities for which we were unable to recruit. Additionally, most participants were US-based, with smaller numbers from Europe and Asia, so findings may be most applicable to Western contexts.

These limitations are well within the acceptable margin for qualitative work. Our findings still provide a valuable insight into how cybersecurity professionals are using and perceiving LLMs for their security-related tasks in this critical and formative stage of the technology.

4 Study Participants

We received 46 survey responses, from which 41 eligible participants were invited to schedule an interview. In total, 29 interviews were conducted. One interview was excluded due to a participant's misrepresentation of their job responsibilities in the survey. As a result, our findings are based on 28 participants from 26 different organizations.

Table 1 shows a condensed summary of participant characteristics. 16 participants reported having three or less years of experience, while 12 had more than three years. This three-year mark is a natural dividing point to distinguish practitioners whose expertise developed largely before or after widespread LLM adoption, and aligns with a typical point of early-career promotion based on LinkedIn profiles we reviewed. We therefore use this three-year mark in the remainder of the paper to differentiate less-experienced from more-experienced participants. Across our sample, participants collectively represented more than 16 distinct job responsibilities, with the most common being penetration testing, threat detection, and incident response. Participants were employed across diverse organizational types, though technology organizations were the most prevalent. Of these organizations, 2 reported having a single cybersecurity practitioner, 15 had fewer than 50 practitioners, and 11 employed more than 51 practitioners. A complete breakdown of our participants can be found in Appendix 2.

4.1 Reported LLM Use Among Participants

All participants reported using LLMs in their professional work on a regular basis, with 17 describing daily use, 8 reporting use several times per week, and 3 reporting monthly use. Reported uses spanned a wide range of cybersecurity tasks and were often tailored to individual workflows; however, several recurring categories appeared across roles, most of which align with general-purpose LLM use reported in other

[1] See the codebook for details: https://osf.io/97u5r/overview?view_only=9aead8bb2f6447c184f6d5df975079f0.

Category	Response	Count
Years of exp.	>3	12 (43%)
	3+	16 (57%)
Responsibilities	Penetration testing	8 (29%)
	Threat detection	6 (21%)
	Incident response	6 (21%)
	Compliance	4 (14%)
	Risk management	3 (11%)
	Red teaming	3 (11%)
	Threat hunting	3 (11%)
	Triage	3 (11%)
	Consulting	3 (11%)
	Threat intelligence	2 (7%)
Usage freq.	Daily	17 (61%)
	A few times a week	8 (29%)
	A few times a month	3 (11%)
Org. type	Technology	13 (46%)
	Government	4 (14%)
	Finance	4 (14%)
	Academia	3 (11%)
Team size	1	2 (7%)
	2-50	15 (54%)
	51+	11 (39%)

Table 1: Summary of our participants (most common categories). P14, who works at a large organization, declined to specify the size of the cybersecurity team.

domains [21, 35]. For baseline context, we summarize these below as five high-level use-case themes.

Code Generation and Other Technical Applications. The most commonly discussed use case was code generation, used to automate tasks and reduce time spent on repetitive work. More broadly, participants applied LLMs to other technical tasks, including generating phishing content and creating test cases for quality assurance.

Natural Language Generation. Participants also often described leveraging the natural language capabilities of LLMs, such as for writing reports, simplifying technical writing, and generating training materials.

Retrieving and/or Summarizing Existing Knowledge. A third common use case was retrieving or summarizing existing knowledge. Participants often used LLMs in place of search engines for software library documentation, market research, threat intelligence, or educational purposes.

Data Processing and Analysis. Participants also noted applying LLMs for data processing and analysis tasks. We distinguish these analysis tasks from information retrieval tasks because participants described using LLMs to generate new information rather than recall existing information. These tasks included using LLMs for network analysis, reverse engineering, reimplementing cryptographic routines, and assisted in exploiting vulnerabilities. P21 shared that they find LLMs particularly helpful in creating intrusion detection signatures, especially for encrypted communications.

Obtaining novel perspectives and guidance. Many participants described using LLMs to obtain novel perspectives or high-level guidance. Notably, LLMs were used as a "coworker" to brainstorm ideas, troubleshoot, provide initial guidance, or check if their intuition was on the right track.

5 Understanding The Landscape

We present results of our thematic analysis exploring how cybersecurity practitioners are adopting LLMs in their daily work. In this section, we present the organizational policies and overall attitudes toward LLMs among peers and management (RQ1). In Section 6, we describe the challenges that participants face, perceived risks of such challenges, and their mitigation strategies (RQ2). Lastly, in Section 7 we shed light on the psychological barriers and shifting social dynamics surrounding LLM adoption (RQ3).

5.1 Organizational Policies on LLM Usage Vary Widely, with Little Oversight

Participants reported a wide spectrum of organizational policies, from those with no formal guidance to those with strict rules and clear recommendations for LLM use. Overall, participants' experiences can be grouped into three broad categories, with roughly a third of participants falling into each.

No Clear or Communicated Policies. Participants in this group were unsure whether any formal policies existed, and reported receiving no explicit guidance. Despite the lack of formal policies, most participants in this group reported that they avoid including personally identifiable or client information in LLM inputs due to privacy concerns.

Relaxed or Generic Policies. Participants in this group reported that their organization had LLM policies in place, but that the policies were relaxed or fairly non-specific, usually relating to protecting proprietary data. P2 described their organization's LLM policy:

> *"So far, that policy is basically, as long as you're not putting any company or customer data, you can use the general LLMs like ChatGPT, Claude."* (P2)

Restrictive Policies. Participants in this group described working in environments with stricter oversight that extended beyond avoiding customer or personally identifiable information. Some were instructed not to submit information that could reveal internal vulnerabilities. Others were limited to specific models, typically customized internally or locally hosted systems or enterprise licensed ones. Participants working in or contracting for government agencies also reported that specific tools, such as DeepSeek, were banned.

A few participants interested in deeper LLM integration reported frustration with API restrictions that limited them to the chat interface. P24 expressed his frustration:

> *"We'd love to train a model on our triage process, so it'd be better at doing what we're actually looking to do with it, but [...] they don't want us doing certain things yet."* (P24)

Limited Enforcement and Employee Responsibility. Regardless of policy strictness, most participants reported that policies and recommendations are generally not tailored specifically to cybersecurity tasks, and there are limited or no controls to ensure compliance. A few participants noted efforts were underway to implement additional controls, while other organizations relied on existing data loss prevention measures. But ultimately, the responsibility for misuse of these models largely falls on the individual employee.

5.2 Organizations Encourage LLM Use, But Informal Guidance Remains Mixed

Despite the frequent absence of clear policies, junior participants commonly reported that they generally felt encouraged to use LLMs to enhance productivity, particularly when the organization had invested heavily in an enterprise model.

While participants generally perceived the organizational sentiment as encouraging, several also reported receiving cautionary warnings. P19 suggested this may reflect a tension between goals: organizations support practitioners in leveraging the potential benefits of LLMs while simultaneously expressing concern about misuse that could violate policies.

Adding to the conflicting guidance, another participant described mixed signals between management and their team. P24 described the situation, stating:

> *"I would say my team was not doing it, but the onboarding process emphasized it [...] They're really pushing automation."* (P24)

They attributed their team's limited use of LLMs to two factors: first, many of their roles do not require coding, and second, a generational difference, with newer team members appearing more open-minded toward LLMs.

6 Risks and Challenges

This section describes the challenges participants encountered when using LLMs, the risks they perceive, and some of the strategies they use to navigate these difficulties (RQ2). While many of these challenges echo findings from adjacent domains [35, 49], we highlight where cybersecurity introduces distinct dynamics.

6.1 Lack of Training and Productivity Loss

A key challenge with any new technology is learning how to use it effectively. Most participants reported receiving no training and having a limited understanding of LLM mechanisms. Three junior participants mentioned receiving prompt engineering training from their organizations; however, they described this training as generic, not specific to cybersecurity, and not particularly useful. P24 even called it a "joke." While we did not directly ask participants about prompt engineering confidence, several more-experienced participants expressed uncertainty in their prompting abilities. To try to keep up, some mentioned learning independently through online courses or blog posts. Although several practitioners appeared eager to learn in their own time, a few noted that integrating these tools into their existing workflows was out of scope due to their heavy workloads.

More experienced participants deal with limited LLM expertise by reverting to manual task completion. Given limited training, the majority of participants shared that they prompt the LLM as if they were talking to a real person. Interestingly, while both experienced and novice participants reported using conversational prompting, they expressed different sentiments about their capabilities.

More experienced participants, such as P1, expressed greater frustration with ineffective prompt engineering: "Getting the prompting correct is a skill that's extremely important [...] it works great or it just doesn't work at all just based upon the prompting." These participants tended to give up on the LLM after several iterations, applying their experience to accomplish tasks manually rather than wasting time continuing to engage. P9 further warned that lack of domain knowledge can aggravate time loss:

> *"Don't use LLMs [...] if you don't know what you're doing, because you'll end up creating a lot of work for later on."* (P9)

Most experienced practitioners therefore recommended restricting LLM use to tasks within one's expertise.

In contrast, more junior participants tend to continue prompting until the task is complete. Junior participants were generally more confident about getting the model to provide the necessary information they needed. At the same time, some also reported not being confident in their ability

to complete tasks without the LLM. As a result, these participants generally continue prompting until they get a result they are happy with. P21 reported:

> *"There's no specific amount of tries or time, it's just until you get the result."* (P21)

He later explained that his continued reliance on an LLM to complete a task reflected a lack of organizational support, which left him to resolve issues on his own.

Tasks that require too much context are not worth it. Even when participants possess the necessary knowledge, they report that the models' lack of contextual understanding can also lead to wasted time. Many reported manually providing required context to remain policy-compliant, but believed the effort outweighed potential benefits in some cases. P10 explained that when extensive context was required, he often found himself questioning the value of using the model at all, since completing the task manually was faster.

> *"The LLM doesn't have an underlying knowledge of why you're doing something. Unless, you explain it, but you would require a lot of context. At some point you get to [thinking], isn't it much easier for me to do this myself?"* (P10)

As a result, many participants prefer to leverage LLMs for smaller, more scoped tasks that require minimal context such as decryption or generating small scripts, where the models tend to perform reliably.

Several participants suggested an ideal way to mitigate this challenge would be to integrate the model into their workplace systems, to allow for a better understanding of their operations. However, only a few reported having this capability in their organization. One participant, P6, explained the benefit of this integration in applying organizational context to produce tailored outputs:

> *"We've got an internal MCP server now as well. [...] it can fetch the HTML of a web page and then it'll think about the HTML and say why is this web page malicious [...] It shortens the time it takes for us to go find something to key in on when we're writing a brand new detection rule."* (P6)

6.2 Overreliance Can Hinder Learning and Critical Thinking

Participants also expressed concerns that overreliance on LLMs could erode critical thinking. Many saw this risk in their own practice, while more experienced participants were particularly concerned about less experienced practitioners becoming overly dependent.

> *"I think it's a crutch for some younger folks. I think it can remove some of the technical burden, right. And it can enable people to not understand the fundamentals as easily. I think that's kind of a more subtle negative."* (P19)

This connects to participants' belief that while LLMs can help users learn information, genuine expertise can only be developed through conventional learning methods.

> *"What if I [had access to this back when I was learning?] This sort of would simplify my work, but I would not have learned to reverse engineer. Staring at my screen for 8 hours a day trying to figure some obfuscation out. [...] If you don't get your reps in to actually understanding that, generative AI will not help you get better at your job. You may masquerade as a useful employee for a while until it plateaus to a certain point where you need some expertise."* (P9)

Perhaps surprisingly, this concern was echoed by several of our less experienced participants who had recently graduated, who reported that they were able to complete their cybersecurity degrees with higher grades but comparatively less effort through the use of LLMs. P22 even reflected that, in retrospect, he would have approached his studies differently.

> *"I should have worked my [profanity] off to get good grades. Now getting my grades was easier, so I don't know anything, to be honest with you. Maybe if I refresh I might get something, but I'm blank."* (P22)

In fact, several of our less experienced participants shared that while access to LLMs does give them greater courage to tackle new tasks, it also slightly diminishes their confidence in their own technical abilities.

A few participants shared that they have taken measures to mitigate these effects, such as limiting their use of LLMs to automate tasks that do not require much thinking, or to collect information.

6.3 Hallucinations and Lack of Safeguards for Validation Increase Likelihood of Errors

Unsurprisingly, hallucinations and errors were among the most commonly discussed challenges of using LLMs.

Validation strategies are commonly ad-hoc. Similar to findings from prior work [35], participants nearly unanimously reported an absence of formal safeguards or structured procedures to verify model outputs. To mitigate these challenges, participants described using common strategies, such as experimenting with different models to compare responses and evaluating performance against existing trusted solutions.

Multi-level reviews enable verification but dilute responsibility. Some participants noted that their organizations' existing multi-level review processes (predating LLMs) can help catch errors in LLM-generated content before sharing with

critical stakeholders. For some junior participants, these reviews provided the sense of a safety net: a belief that risks from LLM outputs could be tolerated because someone else would catch them later. In contrast, more experienced participants responsible for the final product emphasized that they cannot afford to take chances when their name is on the line.

Incorrect outputs have caused real harms. Two participants reported experiencing verification failures that have already led to direct harms. P27 relied on an LLM-provided script that resulted in data deletion:

> *"I just prompted asking, I need a script where I can delete the browser cache [...] I ran the script, and a couple of people came to our team, saying that they lost their saved passwords and bookmarks."* (P27)

For P27, this negative experience resulted in anxiety about future use (i.e., fear of "breaking things" again). Although increased caution could be viewed positively as a means to limit future errors, P27 reported that in practice, no additional organizational safeguards were implemented and he did not change his behavior, even after multiple incidents had occurred. He explained that given the limited support available from his team, he saw no other alternative but to assume the risk in order to complete his tasks.

P7 reported an incident that created reputational damage:

> *"It had hallucinated the internal IPs. The colleague hadn't paid attention, he just copy, paste it, put in report. And, uh, yeah, the client noticed that, and they went a little crazy."* (P7)

Other participants speculated about similar risks that could arise from undetected output errors. Because of risks like these, many participants expressed a strong view that LLMs should never be used to make permanent decisions or perform independent actions without oversight from humans.

Limited discussion of threats to LLMs. Notably, few participants mentioned possible adversarial threats to LLMs, such as data poisoning and prompt injection, which could further aggravate the risks of incorrect outputs. Those few who did express awareness of these threats did not describe any mitigation measures. This suggests that attacks on LLMs, and the consideration of LLMs as part of an attack surface, did not appear to be front of mind for practitioners in our study.

6.4 Participants Consider Data Leakage Risks But Often Scope Them Narrowly

The most frequently cited risk was potential data leakage. However, unlike other risks, no participant reported experiencing this issue in practice. It remains unclear whether incidents have not occurred, went unnoticed, or were undisclosed due to non-disclosure agreements. Most participants expressed concern about inadvertently violating policies or laws by exposing sensitive data, with some fearing personal repercussions if a leak were traced back to them.

Some participants recognize broader categories of sensitive data, but input limitations are typically narrow in scope. A few participants specifically mentioned the risk of malicious actors accessing data that could enable attacks, such as indicators of compromise (IOCs) collected in their threat intelligence process. P15 reported fearing that if attackers obtain insights about defense systems, they will be given an advantage.

Despite this concern, none of our participants described taking steps to avoid these risks, beyond basic organizational policies to protect, e.g., PII. Some of the data participants reported inputting into LLMs without enterprise licenses or data handling agreements was surprising. One striking example came from P18, who relies on a personal (non-organizational) LLM, and whose job responsibility includes collecting questionnaire data to identify companies' cybersecurity weaknesses and make recommendations.

> *"What I do is once I have those 157 data points [...] I just put that information into ChatGPT and I tell him, what are the missing points? Just tell me. How should I consult my target company, how do I improve their cybersecurity infrastructure. So it does most of the work for me."* (P18)

When asked, he mentioned that no personally identifiable information like the organization's name is included in the data. However, even without PII, this data can disclose the potential vulnerabilities of an organization.

Only P25 described an organizational policy to avoid inputting information that can disclose vulnerabilities, but he was not sure why this policy was needed.

> *"Information that can be used to exploit a system, we're not supposed to put that in there either. Now, why? Why not? I don't know. It's weird.[...] We've been actually surprisingly restrictive. [...] but what are the actual practical consequences of having information in there?"* (P25)

6.5 LLMs Do Not Perform Well for Certain Cybersecurity-Specific Problems

Participants also reported cybersecurity-specific challenges.

Model guardrails inhibit penetration testing. One such challenge, reported by several participants, is that model guardrails can hinder some security tasks. Guardrails that are intended to prevent malicious use cases can cause difficulties for penetration testing and red-teaming, which naturally overlap with malicious activity. Participants like P14 described continuing efforts to work around guardrails that are frequently updated, requiring new strategies:

> *"I have to jump through hoops because of ethical guidelines that these LLMs have [...] I have to regenerate the prompt in a way that it understands that it's not for malicious reasons, so that it gives me the script. And every time it gets smarter [...] so I cannot use the same thing again, so it's a little cumbersome."* (P14)

Documentation is updated faster than models are. As presented in 4.1, practitioners frequently rely on LLMs for guidance and domain-specific knowledge in cybersecurity, such as indicators of compromise (IOCs) and information needed to craft detection rules. While much of this information is available online, it is continuously updated. Participants reported noticing that the models often lag behind these updates, which results in the guidance being less helpful. P4 explains that this occurs to him with a specific detection language query whose documentation is constantly being updated.

> *"It definitely falls short, especially on more niche things. So, like, the detection language that we write our rules in [...] There's public documentation on it but it's updated pretty regularly, [...] so it doesn't do very well at that niche component."* (P4)

Malicious infrastructure raises the stakes of AI actions. Another participant—whose role involves malware analysis—raised a different issue: whether agentic AI systems might inadvertently interact with or access the infrastructure of malicious actors, which could lead to unintended exposure.

> *"I'm a little worried about some of the smarter LLMs trying to reach out to links in malware or whatever. Because when we're dealing with APTs, you never touch attacker infrastructure from an incident response company's network. That's just a complete no-no."* (P8)

At the time of the interview, he did not consider this a current risk, but framed it as a plausible and likely future concern.

6.6 The Benefits Usually Outweigh the Risks, and Even When They Don't, There's No Alternative

While participants acknowledged the aforementioned risks, ultimately, most viewed them as manageable and ultimately outweighed by the benefits. In practice, even those who felt apprehensive often chose to use LLMs to complete their work.

LLMs save time and effort, especially for undesirable tasks. One of the most commonly reported benefits was a reduction in the time required to complete tasks, particularly when practitioners had a clear understanding of what they needed from the LLM. Reduction in time was especially evident in tasks where LLMs demonstrated clear advantages over humans. Participants highlighted examples such as understanding obfuscated code or detecting homoglyphs, which can be particularly challenging for humans. P8 explained that he finds the models very helpful for deciphering obfuscated code and searching through junk code to pick out the malicious pieces of code he needs. Time efficiency was often accompanied by a reduction in cost. In fact, P1 describes an instance where use of LLMs lead to a faster recovery from a ransomware attack:

> *"We were able to programmatically go through and extract the data that we needed to get recovered quicker and keep the client from having to pay millions of dollars in ransom."* (P1)

Perhaps one of the greatest advantages participants report is the ability to offload tasks that they either have little desire to perform or do not feel confident performing. One common example, as mentioned in 4.1 is code generation, which, while often required, is not the primary responsibility of many cybersecurity practitioners. Some participants also reported offloading tasks such as analyzing assembly code.

Overall, risk considered low and manageable. While participants acknowledged the existence of various risks, many expressed confidence in their ability to manage them. This perception can perhaps be explained by the fact that only a few participants could recall any significant negative impacts resulting from the use of LLMs or an automated decision made by an LLM. Some participants argued that although LLMs introduce mistakes, these are not more frequent than those made by humans:

> *"I mean [...] there's always false positives, but not any more or less than, let's say, a human."* (P3)

Others expressed that their trust stemmed from strategic use of LLMs in contexts where risk is limited or inaccuracies pose minimal consequences, particularly because they used models that cannot act autonomously.

Sometimes practitioners feel they have no other choice. While most participants reported that, overall, the benefits outweigh the risks, several participants also argued that their reliance on LLMs is largely due to necessity. P14 explained that he continuously prompts the model until it produces the desired output because he does not have someone to turn to for assistance in completing his tasks.

> *"Until it works, because there's nobody who's coming to help me."* (P14)

P26, meanwhile, reported asking an LLM when her coworkers do not know the answer.

> *"I asked my coworkers, they also don't know, I'm gonna ask ChatGPT."* (P26)

These observations highlight the complex ways that AI tools have affected workplace relationships, as we will discuss in the next section.

7 The Psychological Barrier

While the challenges above largely mirror those reported in adjacent fields, our most significant and novel finding concerns a psychological barrier. In examining the social dynamics surrounding generative AI adoption, we found that fear of "what will others think of me if I use AI?" was common among participants and carries significant implications. While this concern does not prevent practitioners from adopting AI, it is shaping how they communicate about its use, influencing longstanding social norms of openness, knowledge sharing, and mentorship within the cybersecurity community (RQ3) [19].

Fear of judgment regarding LLM use discourages sharing among practitioners. During our analysis, we noticed participants described two types of behaviors that deviate from what might be expected given the cybersecurity community's well-known culture of openness and sharing [19]. The first was that many participants reported not knowing how their colleagues use LLMs. While they were aware that it was now a commonly-used tool, they did not actively try to engage in conversations about LLM usage. P7 shared that the only reason they learned about their coworkers' LLM usage was because of a team activity:

> *"It was funny, we had an off-site meeting with all my colleagues, [...] and it seemed everyone but one person out of the team, we all did this. We just hadn't talked to each other."* (P7)

The second, and arguably the most surprising, was participants' fear of revealing their use of these tools to their teammates. P25 explained that they hesitated to disclose their use of LLMs out of concern that others might question their behavior or restrict their access:

> *"I would basically hide that I was using the LLM, because [...] I didn't want people to question that, and limit my ability to use it."* (P25)

Other participants fear being negatively perceived as overreliant on these tools.

> *"Everyone kind of keeps it to themselves. [...] I think people who rely on it a lot don't really like to say that they use ChatGPT."* (P26)

Similar fears have been documented among researchers reluctant to disclose LLM use in academic papers for fear of violating cultural norms around individual effort [6].

At first glance, these concerns might seem speculative, because participants who expressed fear of disclosing their LLM usage did not report direct experiences of being judged or stigmatized for doing so. However, conversations with more experienced practitioners revealed that these fears were not entirely misplaced. Several described instances in which the misuse of such tools had negatively influenced how users were perceived:

> *"I've seen AI used in my organization in a way that gives the impression of a lack of care, concern, or even knowledge of their subject matter, which [...] reduces your confidence in them [co-workers]."* (P9)

Moreover, research in organizational behavior has found that in fact, disclosing AI use does erode other people's trust in the AI user [51].

Heightened concerns about job security incentivize nondisclosure. This fear of judgment is further compounded by job uncertainty. Despite the oft-cited demand for skilled cybersecurity professionals, the current job climate does not always reflect this need [39]. While not all participants expressed concerns about their own job security, some observed early signs of displacement in entry-level jobs, particularly for SOC analysts. More generally, participants anticipate near-term job cuts, with P6 describing a widespread fear of AI-driven replacement:

> *"Everybody has it in the back of their mind where eventually, this is going to replace my job."* (P6)

This added uncertainty compounds the psychological barriers to openly using AI tools or seeking guidance. Practitioners are understandably reluctant to give their employer any evidence that an AI tool is particularly useful for their job, or an additional reason to form negative perceptions of their competence or motivations. This transparency dilemma is a topic of significant study in organizational behavior (e.g., [51]); however, it has not, to our knowledge, yet been documented in other technical communities.

Participants' fear of judgment surpasses their desire to learn from others. Interestingly, although many participants mentioned that they do not actively share insights about their LLM use for cybersecurity tasks, one of the most valuable insights they sought from this study was learning how others use these tools. Participants believe that observing others' experiences would be beneficial, as it could help them identify new use cases, adopt effective strategies, or confirm their own assessments concerning LLMs:

> *"If everyone else in the security industry is saying, 'Hey the LLMs are getting better for most security tasks,' [...] I might be able to relax a little bit and trust it a little bit more for simpler things."* (P8)

P7 even expressed a desire for a structured guide on LLM use in cybersecurity, as they currently rely on trial and error for every task.

Participants clearly understand the benefits of knowledge sharing, yet their fear of judgment seems to outweigh their drive to learn and enhance their skills in applying LLMs. Unfortunately, the consequence of this mindset extends far beyond individual learning. By limiting the exchange of insights and strategies, these social pressures threaten to erode the cybersecurity community's collaborative culture. Moreover,

they can also undermine the community's collective ability to respond effectively to ever-evolving threats, including adversarial threats directed at LLMs themselves.

Demand for teamwork and mentorship is declining. We also observe that the integration of AI tools is affecting social dynamics within the community. Importantly, this effect was raised spontaneously by participants, without any prompting. Participants described becoming less inclined to ask colleagues for advice or help on tasks. P12 reported that when seeking guidance, their default approach is now to consult an LLM instead of a human colleague. This pattern of usage can result in a decline in teamwork:

> *"I thought I should be working with the people, but then I'd realized they knew less than the LLM. I was like, 'Well, I probably should just rely on the LLM entirely then, and not work with the people.' [...] It doesn't build up teamwork. It doesn't integrate well with organizational processes."* (P25)

This shift is similarly affecting the mentorship model. We find that the demand for guidance and support from more experienced practitioners appears to be decreasing. This change has some benefits, including faster feedback and avoiding potential negative judgment due to lack of experience or asking basic questions. P28 shared a challenging experience when first starting out in cybersecurity and needing to ask many questions:

> *"When I started with cybersecurity back in 2018, there wasn't much resources [...] It was pretty much heavy Googling and literally asking people, asking senior professionals. Pretty much a dependency, and now it's not the case, LLMs always help to give you that knowledge quicker."* (P28)

Similarly, P17 recalled asking questions on Stack Overflow and receiving a discouraging reception:

> *"If you have any doubts, you have to post in Stack Overflow, and some people [...] will get angry if you ask any basic question, right? So, rather, I would ask this question to ChatGPT."* (P17)

These comments align with prior work that has identified mentoring and apprenticeship, rather than formal education, as the primary mechanisms for developing cybersecurity expertise [1, 61]. At the same time, some of our more experienced practitioners reported a conflict between giving tasks to junior co-workers—who might take longer to accomplish the task but would learn from the experience—and using an LLM.

While LLMs offer considerable assistance, our findings (e.g., Section 6.1) reveal that they cannot fully replace human support, especially when the models fall short and less-experienced participants get stuck. In these situations, participants still require guidance from human mentors; without it, they risk wasting significant time or making errors that could have serious consequences. P9 described just this type of mentorship situation:

> *"[He] threw the code at the LLM, and it gave him this blob of code that was, like, 40 lines long, super complex. We spent 13 minutes reverse engineering, line by line, in Ghidra. [...] And he realized how simple it was [... it] just looks confusing in your first go. So he just threw it to AI to get the answer, instead of just working through it."* (P9)

In this case, the availability of a mentor helped overcome LLM challenges.

8 Discussion

Through in-depth interviews with a broad range of cybersecurity practitioners, we were able to uncover a less-discussed but potentially highly impactful challenge: the psychological and social dynamics shaping generative AI adoption. Fear of judgment, uncertainty about policies, and questions of competence appear to influence how professionals engage with AI. Rather than collaborating and sharing knowledge as the community has traditionally done, many practitioners are navigating AI adoption largely on their own. These findings make it clear that integrating AI into cybersecurity is not just a technical challenge—it is fundamentally a human one.

8.1 Signals of Cultural and Community Shifts

Our findings suggest a departure from long-standing traditions of collaboration, discourse, and resource sharing within cybersecurity communities, with respect to generative AI and LLMs. This departure appears to be partly driven by a perceived stigmatization of LLM usage within the security community. As a result, our participants are often reluctant to share their full use of LLMs for work tasks with their coworkers, and even less willing to share best practices or prompting strategies. In these instances, fear and stigma actively prevented knowledge transference and collaboration, creating a tense environment in which most professionals acknowledge (and likely expect) widespread LLM usage by their peers, yet choose not to disseminate their own practices.

Several interacting factors may be contributing to this perceived stigma. First, many participants described organizational policies as unclear, inconsistent, or entirely absent. This lack of clarity leaves practitioners uncertain about acceptable use and concerned that they might unintentionally violate policies and be held personally accountable for potential damages. Limited training and few opportunities to experiment further compound this uncertainty, reducing confidence and making practitioners hesitant to discuss their LLM use with peers.

Beyond organizational policies and training, social and cultural norms within cybersecurity may also reinforce the perceived stigma around LLM use. As reflected by participants such as P9 (Section 6.2), professional norms have long em-

phasized technical expertise and the value of "earning your stripes" through experience. In many disciplines, this meritocracy is a source of professional pride and identity, and effortful learning is often treated as a signal of competence and legitimacy [59]. In cybersecurity, mentorship has historically been a key avenue for transmitting scarce, specialized knowledge [20]. Consequently, the emergence of LLMs can be perceived as disrupting traditional learning dynamics by enabling newcomers to bypass established pathways of expertise, thereby potentially threatening what psychologists describe as occupational or professional identity [33]. Such perceived identity threat may bias how practitioners interpret and evaluate LLM use by others, likely contributing to the observed stigmatization of those who rely on these tools.

It's also important to note that workload pressures and limited access to supportive mentorship further shape these dynamics. Several participants noted consulting an LLM instead of a senior co-worker because guidance from the AI was faster or more convenient. One of our participants also shared using LLMs as a strategy to avoid toxic elements within the community, which highlights that traditional mentorship is not always available or beneficial [20]. In this sense, LLMs are both a reflection of and a contributor to shifting community behaviors, providing alternatives when human guidance is constrained by time, availability, or interpersonal challenges.

Together these factors leave professionals in a difficult position: learning to leverage powerful tools with minimal guidance while navigating potential stigma, risk of errors, and the fear of being left behind.

8.2 Need for Future Work

Our interview study offers an initial glimpse into how perceived stigma around LLM use may influence collaboration and everyday practice in the cybersecurity workforce. Given the historically central role of collaboration and mentorship in this domain, future research is nonetheless needed to understand the long-term community impacts of LLM adoption, including when such shifts may be beneficial, detrimental, or both. This need is further underscored by the lack of consensus on how generative AI affects workplace social dynamics in adjacent fields. Giray et al. [21], for instance, reported no agreement among software engineers on whether gen-AI reduced human interactions or increases isolation. Meanwhile, Dell'Acqua et al. [13] found that individuals working with AI can perform at levels comparable to an entire team, suggesting a diminished need for human collaboration. Our findings diverge from both: participants in our study described emerging feelings of isolation alongside a need for human collaboration.

These discrepancies may reflect differences in study populations—particularly cross-national difference in work culture—as well as methodological approach. Both prior studies relied on surveys, whereas our interview-based methodology proved instrumental in surfacing the nuanced, experience-based concerns reported in Section 6—such as stigma and reluctance to disclose AI use—dimensions that surveys alone can struggle to capture [47]. We therefore strongly encourage the adoption of qualitative interview methods in future work seeking to examine these dynamics in greater depth. With these considerations in mind, we outline several promising directions for future research below.

Further research is needed to inform the development of clear guidelines and best practices for effective and appropriate LLM use in cybersecurity contexts. Importantly, well defined norms could facilitate more specific, evidence-based critique of potentially inappropriate LLM use cases, and help displace the kind of generalized stigma our participants described. Complementing such efforts, larger-scale and longer-term studies would allow for a more granular examination of practitioners' perspectives. Large-scale studies could illuminate relationships between organizational culture, prevailing norms, and experiences of stigma or collaboration, while longitudinal research could track how teams and security subcommunities adapt to LLMs over time.

Future work should also explore when and how to foster collaborative environments in the era of AI, also recommended by [21]. For example, do practitioners want to reintroduce more human collaboration, and if so, on which tasks, and how? Key questions include whether existing knowledge sharing mechanisms such as regular meetings, workshops, challenges, and discussion boards remain effective; what practitioners expect from collaborative tools; and which approaches best support learning and transparency. Research could also examine whether specific interface designs or tools can facilitate collaboration or reduce stigma, enabling more open and productive LLM use within teams.

Addressing these questions can guide the development of best practices, organizational strategies, and technological solutions that promote intentional, responsible, transparent, and socially supported adoption of LLMs in cybersecurity.

9 Conclusion

We interviewed 28 cybersecurity practitioners to investigate the impact of LLMs on the cybersecurity workforce. We examined organizational policies and guidance, as well as perceived risks and challenges, and the strategies practitioners use to address them. We find that policies governing LLM usage are often unclear, and guidance generally absent. Practitioners apply LLMs to a wide range of security tasks, but face significant challenges related to trust. More importantly, we find that cybersecurity practitioners experience a fear of judgement regarding AI use, which is shifting the culture from collaboration toward individualism.

Acknowledgments

We thank all of our participants for their time and our anonymous reviewers for their thoughtful feedback, which strengthened this work. This project was supported in part by NSF grant CNS-1943215.

References

[1] Omer Akgul, Taha Eghtesad, Amit Elazari, Omprakash Gnawali, Jens Grossklags, Michelle L. Mazurek, Daniel Votipka, and Aron Laszka. Bug Hunters' Perspectives on the Challenges and Benefits of the Bug Bounty Ecosystem. In *Proc. USENIX Security*, 2023.

[2] Haitham S. Al-Sinani and Chris J. Mitchell. AI-Enhanced Ethical Hacking: A Linux-Focused Experiment, October 2024. arXiv:2410.05105 [cs].

[3] Massimiliano Albanese, Xinming Ou, Kevin Lybarger, Daniel Lende, and Dmitry Goldgof. Towards AI-Driven Human-Machine Co-Teaming for Adaptive and Agile Cyber Security Operation Centers, May 2025. arXiv:2505.06394 [cs].

[4] Erica Azad. Inside the Mind of a Hacker: 2024 Edition | @Bugcrowd, October 2024.

[5] Tyler Balon and Ibrahim (Abe) Baggili. Cybercompetitions: A survey of competitions, tools, and systems to support cybersecurity education. *Education and Information Technologies*, 28(9):11759–11791, 2023.

[6] Aorigele Bao and Yi Zeng. AI disclosure, moral shame, and the punishment of honesty. *Accountability in Research*, 33(3):1–14, 2025.

[7] Mazal Bethany, Athanasios Galiopoulos, Emet Bethany, Mohammad Bahrami Karkevandi, Nicole Beebe, Nishant Vishwamitra, and Peyman Najafirad. Lateral Phishing With Large Language Models: A Large Organization Comparative Study. *IEEE Access*, 13:60684–60701, 2025.

[8] Artur Boronat and Jawad Mustafa. MDRE-LLM: A Tool for Analyzing and Applying LLMs in Software Reverse Engineering. In *Proc. SANER*, 2025.

[9] Hack The Box. https://www.hackthebox.com/.

[10] Virginia Braun, Victoria Clarke, and Nikki Hayfield. 'A starting point for your journey, not a map': Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. *Qualitative Research in Psychology*, 19(2):424–445, 2022.

[11] William Crumpler and James A. Lewis. The Cybersecurity Workforce Gap. Technical report, Center for Strategic and International Studies (CSIS), 2019.

[12] Hoang Cuong Nguyen, Shahroz Tariq, Mohan Baruwal Chhetri, and Bao Quoc Vo. Towards Effective Identification of Attack Techniques in Cyber Threat Intelligence Reports using Large Language Models. In *Proc. WWW*, 2025.

[13] Fabrizio Dell'Acqua, Charles Ayoubi, Hila Lifshitz, Raffaella Sadun, Ethan Mollick, Lilach Mollick, Yi Han, Jeff Goldman, Hari Nair, Stewart Taub, et al. The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise. Technical report, National Bureau of Economic Research, 2025.

[14] Gelei Deng, Yi Liu, Víctor Mayoral-Vilches, Peng Liu, Yuekang Li, Yuan Xu, Tianwei Zhang, Yang Liu, Martin Pinzger, and Stefan Rass. PentestGPT: Evaluating and Harnessing Large Language Models for Automated Penetration Testing. In *Proc. USENIX Security*, 2024.

[15] Yangruibo Ding, Yanjun Fu, Omniyyah Ibrahim, Chawin Sitawarin, Xinyun Chen, Basel Alomair, David Wagner, Baishakhi Ray, and Yizheng Chen. Vulnerability Detection with Code Language Models: How Far Are We?, July 2024. arXiv:2403.18624 [cs].

[16] Angela Fan, Beliz Gokkaya, Mark Harman, Mitya Lyubarskiy, Shubho Sengupta, Shin Yoo, and Jie M. Zhang. Large Language Models for Software Engineering: Survey and Open Problems. In *Proc. ICSE-FoSE*, 2023.

[17] Ruirui Feng, Hui Chen, Shuo Wang, Md Monjurul Karim, and Qingshan Jiang. LLM-MalDetect: A Large Language Model-Based Method for Android Malware Detection. *IEEE Access*, 13:81347–81364, 2025.

[18] Mohamed Amine Ferrag, Ammar Battah, Norbert Tihanyi, Ridhi Jain, Diana Maimuţ, Fatima Alwahedi, Thierry Lestable, Narinderjit Singh Thandi, Abdechakour Mechri, Merouane Debbah, et al. SecureFalcon: Are We There Yet in Automated Software Vulnerability Detection with LLMs? *IEEE Transactions on Software Engineering*, 2025.

[19] Nathan Fisk, Nicholas M. Kelly, and Lori Liebrock. Cybersecurity communities of practice: Strategies for creating gateways to participation. *Computers & Security*, 132:103188, 2023.

[20] Kelsey R. Fulton, Samantha Katcher, Kevin Song, Marshini Chetty, Michelle L. Mazurek, Chloé Messdaghi, and Daniel Votipka. Vulnerability Discovery for All: Experiences of Marginalization in Vulnerability Discovery. In *Proc. IEEE S&P*, 2023.

[21] Görkem Giray, Onur Demirörs, Marcos Kalinowski, and Daniel Mendez. An Empirical Study of Generative AI Adoption in Software Engineering, 2025. arXiv:2512.23327 [cs].

[22] Olivér Gulyás and Gábor Kiss. Impact of cyber-attacks on the financial institutions. *Procedia Computer Science*, 219:84–90, 2023.

[23] Wenbo Guo, Yujin Potter, Tianneng Shi, Zhun Wang, Andy Zhang, and Dawn Song. Frontier AI's Impact on the Cybersecurity Landscape, April 2025. arXiv:2504.05408 [cs].

[24] HackerOne. 8th Annual Hacker-Powered Security Report 2024-25. https://www.hackerone.com/resources/pf/col/home/8th-hacker-powered-security-report. Accessed 2025-11-13.

[25] Andreas Happe and Jürgen Cito. Getting pwn'd by AI: Penetration Testing with Large Language Models. In *Proc. ESEC/FSE*, 2023.

[26] Gary A. Harris. How State Universities Are Addressing the Shortage of Cybersecurity Professionals in the United States. *Journal of Cybersecurity Education, Research and Practice*, 2024(1), 2024.

[27] Mohammed Hassanin, Marwa Keshk, Sara Salim, Majid Alsubaie, and Dharmendra Sharma. PLLM-CS: Pre-trained Large Language Model (LLM) for Cyber Threat Detection in Satellite Networks. *Ad Hoc Networks*, 166:103645, 2025.

[28] Julian Hazell. Spear Phishing With Large Language Models, December 2023. arXiv:2305.06972 [cs].

[29] Xinyi Hou, Yanjie Zhao, Yue Liu, Zhou Yang, Kailong Wang, Li Li, Xiapu Luo, David Lo, John Grundy, and Haoyu Wang. Large Language Models for Software Engineering: A Systematic Literature Review. *ACM Transactions on Software Engineering and Methodology*, 33(8):1–79, 2024.

[30] Adam Jenkins, Pieris Kalligeros, Kami Vaniea, and Maria K. Wolters. "Anyone Else Seeing this Error?": Community, System Administrators, and Patch Information. In *Proc. EuroS&P*, 2020.

[31] Haolin Jin, Linghan Huang, Haipeng Cai, Jun Yan, Bo Li, and Huaming Chen. From LLMs to LLM-based Agents for Software Engineering: A Survey of Current, Challenges and Future, April 2025. arXiv:2408.02479 [cs].

[32] Gavin Jones, Dimitrios Kasimatis, Nikolaos Pitropakis, Richard Macfarlane, and William J. Buchanan. Analysing the role of LLMs in cybersecurity incident management. *International Journal of Information Security*, 24(6), 2025.

[33] Ekaterina Jussupow, Kai Spohrer, and Armin Heinzl. Identity Threats as a Reason for Resistance to Artificial Intelligence: Survey Study with Medical Students and Professionals. *JMIR Formative Research*, 6(3):e28750, 2022.

[34] Samantha Katcher, Liana Wang, Caroline Yang, Chloé Messdaghi, Michelle L. Mazurek, Marshini Chetty, Kelsey R. Fulton, and Daniel Votipka. A Survey of Cybersecurity Professionals' Perceptions and Experiences of Safety and Belonging in the Community. In *Proc. SOUPS*, 2024.

[35] Jan H. Klemmer, Stefan Albert Horstmann, Nikhil Patnaik, Cordelia Ludden, Cordell Burton, Carson Powers, Fabio Massacci, Akond Rahman, Daniel Votipka, Heather Richter Lipford, Awais Rashid, Alena Naiakshina, and Sascha Fahl. Using AI Assistants in Software Development: A Qualitative Study on Security Practices and Concerns. In *Proc. CCS*, 2024.

[36] Jeffrey Korte. Mitigating cyber risks through information sharing. *Journal of Payments Strategy & Systems*, 11(3):203–214, November 2017.

[37] Stefan Laube and Rainer Böhme. Strategic Aspects of Cyber Risk Information Sharing. *ACM Computing Surveys*, 50(5):77:1–77:36, 2017.

[38] Martti Lehto. Cyber-Attacks against Critical Infrastructure. In Martti Lehto and Pekka Neittaanmäki, editors, *Cyber Security: Critical Infrastructure Protection*, pages 3–42. Springer, 2022.

[39] Andrey Leskin. Why cybersecurity jobs are hard to find amid a worker shortage. https://www.darkreading.com/cybersecurity-operations/cybersecurity-jobs-hard-find-amid-worker-shortage, March 2025. Accessed: 2025-11-13.

[40] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 2019.

[41] Martha McNeil and Thomas Llansó. An Analysis of Adversarial Cyber Testing Practice. In *Proc. SSS*, 2020.

[42] Amr Mohamed, Maram Assi, and Mariam Guizani. The Impact of LLM-Assistants on Software Developer Productivity: A Systematic Review and Mapping Study, March 2026. arxiv:2507.03156 [cs].

[43] Timothy Nosco, Jared Ziegler, Zechariah Clark, Davy Marrero, Todd Finkler, Andrew Barbarello, and W. Michael Petullo. The Industrial Age of Hacking. In *Proc. USENIX Security*, 2020.

[44] Rita Olla, Emily Hand, Sushil J. Louis, Ramona Houmanfar, and Shamik Sengupta. A Cybersecurity Game to Probe Human-AI Teaming. In *Proc. CoG*, 2024.

[45] Ipek Ozkaya. Application of Large Language Models to Software Engineering Tasks: Opportunities, Risks, and Implications. *IEEE Software*, 40(3):4–8, 2023.

[46] Ali Pala and Jun Zhuang. Information Sharing in Cybersecurity: A Review. *Decision Analysis*, 16(3):172–196, 2019.

[47] Michael Quinn Patton. *Qualitative research & evaluation methods: Integrating theory and practice*. Sage Publications, 2014.

[48] Yangheran Piao, Temima Hrle, Daniel W. Woods, and Ross Anderson. Study Club, Labor Union or Start-Up? Characterizing Teams and Collaboration in the Bug Bounty Ecosystem. In *Proc. IEEE S&P*, 2025.

[49] Han Qiao, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. To use or not to use: Impatience and overreliance when using generative ai productivity support tools. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2025.

[50] June Sallou, Thomas Durieux, and Annibale Panichella. Breaking the silence: the threats of using llms in software engineering. In *Proc. ICSE-NIER*, 2024.

[51] Oliver Schilke and Martin Reimann. The transparency dilemma: How AI disclosure erodes trust. *Organizational Behavior and Human Decision Processes*, 2025.

[52] Ze Sheng, Fenghua Wu, Xiangwu Zuo, Chao Li, Yuxin Qiao, and Lei Hang. LProtector: An LLM-driven Vulnerability Detection System, November 2024. arXiv:2411.06493 [cs].

[53] Ronal Singh, Shahroz Tariq, Fatemeh Jalalvand, Mohan Baruwal Chhetri, Surya Nepal, Cecile Paris, and Martin Lochner. Llms in the SOC: An Empirical Study of Human-AI Collaboration in Security Operations Centres. In *Proc. IEEE S&P*, 2026.

[54] Shahroz Tariq, Mohan Baruwal Chhetri, Surya Nepal, and Cecile Paris. A^2C: A modular multi-stage collaborative decision framework for human–AI teams. *Expert Systems with Applications*, 282:127318, July 2025.

[55] Shahroz Tariq, Ronal Singh, Mohan Baruwal Chhetri, Surya Nepal, and Cecile Paris. Bridging Expertise Gaps: The Role of LLMs in Human-AI Collaboration for Cybersecurity, May 2025. arxiv:2505.03179 [cs].

[56] Jiessie Tie, Bingsheng Yao, Tianshi Li, Hongbo Fang, Syed Ishtiaque Ahmed, Dakuo Wang, and Shurui Zhou. "Should I Give Up Now?" Investigating LLM Pitfalls in Software Engineering. *ACM Transactions on Software Engineering and Methodology*, March 2026.

[57] TryHackMe. https://tryhackme.com/.

[58] Jonathan Ullrich, Matthias Koch, and Andreas Vogelsang. From Requirements to Code: Understanding Developer Practices in LLM-Assisted Software Engineering. In *Proc. IEEE RE*, 2025.

[59] John Van Maanen and Edgar H. Schein. Toward a theory of organizational socialization. 1977.

[60] Daniel Votipka, Mary Nicole Punzalan, Seth M. Rabin, Yla Tausczik, and Michelle L. Mazurek. An Investigation of Online Reverse Engineering Community Discussions in the Context of Ghidra. In *Proc. EuroS&P*, 2021.

[61] Daniel Votipka, Rock Stevens, Elissa Redmiles, Jeremy Hu, and Michelle Mazurek. Hackers vs. Testers: A Comparison of Software Vulnerability Discovery Processes. In *Proc. IEEE S&P*, 2018.

[62] Daniel Votipka, Eric Zhang, and Michelle L. Mazurek. HackEd: A Pedagogical Analysis of Online Vulnerability Discovery Exercises. In *Proc. IEEE S&P*, 2021.

[63] Wei Wang, Huilong Ning, Gaowei Zhang, Libo Liu, and Yi Wang. Rocks Coding, Not Development: A Human-Centric, Experimental Evaluation of LLM-Supported SE Tasks. In *Proc. FSE*, 2024.

[64] Muting Wu, Raul Aranovich, and Vladimir Filkov. Evolution and Differentiation of the Cybersecurity Communities in Three Social Question and Answer Sites: A Mixed-Methods Analysis. *PLoS ONE*, 16(12):e0261954, 2021.

[65] Miuyin Yong Wong, Matthew Landen, Manos Antonakakis, Douglas M Blough, Elissa M Redmiles, and Mustaque Ahamad. An Inside Look Into the Practice of Malware Analysis. In *Proc. CCS*, 2021.

[66] Zibin Zheng, Kaiwen Ning, Qingyuan Zhong, Jiachi Chen, Wenqing Chen, Lianghong Guo, Weicheng Wang, and Yanlin Wang. Towards an understanding of large language models in software engineering tasks. *Empirical Software Engineering*, 30(2):50, 2025.

[67] Yuwen Zou, Yang Hong, Jingyi Xu, Lekun Liu, and Wenjun Fan. Leveraging Large Language Models for Challenge Solving in Capture-the-Flag. In *Proc. TrustCom*, 2024.

A Interview Protocol

Thank you again for participating in this study. As a reminder, this interview will be audio recorded. If at any point you feel uncomfortable, you can let me know and we can pause, skip a question, or stop altogether. Do you give me consent to begin?

I'd like to begin by reviewing some of the information you provided in the questionnaire. This is a standard step we take with all participants. It's important for the integrity of the study that we have accurate background information, and I want to make sure we categorize your responses correctly. Regardless of the details, we'll still be able to conduct the interview, this is just about making sure our records are precise.

To confirm, how many years of professional experience do you have in a cybersecurity role? What formal education, training, or certifications have you completed to prepare for this role, and in which year(s) did that take place? To get a clearer picture of your current work, could you please describe your current position and specifically what cybersecurity-related tasks or responsibilities you handle in that role?

A.1 Organizational Policy & Informal Guidance

- When did you first start using LLMs for cyber-response work and what prompted that adoption (personal curiosity, company directive, colleague suggestion, etc.)
- Did you have reservations at the start? If so, what were they?
- Since you began using them, has your perspective changed on LLMs changed
- Which LLM services or platforms do you currently use, and how are you accessing it? (e.g., via API, web interface, or another method).
- How common is LLM usage among your peers and coworkers?
- Are there any formal or informal policies at your company about using AI or LLMs?
- To experienced participant: Would you recommend that someone new to cybersecurity use LLMs? Why or why not?
- To novice participant: Were you encouraged or discouraged by senior staff or management to use LLMs for your day to day tasks that involve cybersecurity knowledge (such as analyzing malware, writing signatures, drafting threat intelligence reports, or similar tasks)? Did they provide a reasoning? Did you ever receive conflicting advice?
- Have you taken any courses on prompting? How confident are you about your ability to prompt?
- Do members of your team currently use LLMs to help generate threat-intel or executive reports? If so, is there any formal policy or guidance governing that use?

A.2 Use cases, Challenges & Perception of Risk

- What specific tasks do you apply them to, and how has that changed your day-to-day workflow? Can you provide an example of a prompt that you use?
- Task by task, how well do LLMs perform for you? Where do they excel and where do they fall short?
- Can you recall an instance where an LLM helped detect or mitigate a threat? If so, what happened?
- Can you recall an instance where the use of an LLM had a negative impact your organization?
- What's the first thing you do when you receive an LLM output?
- Do you have a process for verifying or double-checking those results? If so, could you walk me through the steps?
- Are you confident in your ability to detect if it's correct? How do you measure if its good or bad? is it easy or hard?
- Are there specific aspects of an LLM's output that raise or lower your confidence in its accuracy?
- Have you encountered problematic outputs? How did you detect and mitigate them?
- How do you decide when to use LLMs and when to manually do the task?
- Roughly how much time or resources did you invest, and was the benefit worth that effort? Why or why not?
- Do you see any possible risks or concerns when using LLMs in your day-to-day job? If so, what are they?
- Are you ever worried about data leakage when entering proprietary or confidential information into an LLM? If yes, how do you mitigate that risk?
- Have you ever had an instance where an LLM LLM-assisted report contained an error? If so could you walk me through that instance? How was the problem detected, corrected, and communicated upstream?

A.3 Future Outlook

- Do you foresee LLMs automating other parts of your role soon? Which ones?
- Which tasks do you believe will always require human judgment rather than automation, and why?
- Overall what improvements would you most like to see in LLMs?
- Is there anything in particular you're curious to learn from the other people taking part in this study?

B Participants

ID	Years of exp.	Responsibilities	Usage frequency	Org. Type	Team Size
P1	30	Incident response	Daily	Technology	11-50
P2	9	Threat detection; Incident response	A few times a week	Cybersecurity	51+
P3	20	Risk management	Daily	Technology	1
P4	15	Threat detection	Daily	Technology	11-50
P5	12	Penetration testing; Red teaming	Daily	Finance	11-50
P6	3	Threat detection	Daily	Technology	11-50
P7	15	Penetration testing; Red teaming	Daily	Cybersecurity	2-10
P8	10	Malware analysis	A few times a month	Technology	51+
P9	6	Compliance; Risk management	A few times a week	Government	2-10
P10	35	Threat modeling; Consulting	A few times a week	Technology	2-10
P11	9	Secure development	Daily	Technology	51+
P12	6	Penetration testing; Threat hunting	A few times a month	Government	11-50
P13	10	Security training	Daily	Aerospace	51+
P14	1	Compliance; Penetration testing	A few times a week	Technology	Large
P15	3	Threat intelligence	A few times a week	Technology	11-50
P16	<1	Penetration testing; Threat intelligence	A few times a month	Healthcare	11-50
P17	3.5	Threat detection	Daily	Technology	51+
P18	2	Compliance; Consulting	Daily	Technology	1
P19	2	Red teaming; Threat hunting	A few times a week	Government	51+
P20	1	Penetration testing	Daily	Technology	51+
P21	1.5	Threat detection; Triage	Daily	Academia	2-10
P22	3	Penetration testing	A few times a week	Technology	11-50
P23	1	Incident response; Threat hunting; Triage	Daily	Finance	51+
P24	<1	Incident response; Triage	Daily	Finance	51+
P25	1	Compliance; Risk management	Daily	Academia	11-50
P26	1	IT; Security audit	Daily	Academia	11-50
P27	3.5	Threat detection; Incident response	Daily	Finance	2-10
P28	1.5	Incident response	A few times a week	Consulting	51+

Table 2: Participants. P14, who works at a large organization, declined to specify the size of the cybersecurity team.