Investigating the Intersection of AI and Cybercrime Article Review

Charlotte Lacy

Old Dominion University

CYSE 201S: Cybersecurity and the Social Sciences

Professor Diwakar Yalpi

November 10, 2025

Principles of Social Sciences

In the article, “Investigating the Intersection of AI and Cybercrime: Risks, Trends, and Countermeasures,” it frames AI-enabled cybercrime as a social phenomenon shaped by human routines, institutions, and power relations rather than only technical factors. This study uses the Routine Activities Theory to situate cybercrime in social-science concepts such as motivated offenders, suitable targets, and absent guardianship. This shows how every day online behaviors, media narratives, and institutional responses shape opportunities for crime and victimization.

Analysis of Experiment

There are 3 research questions that were posed in this article: How is information involving malicious use of AI distributed and used on both the dark web and the clear web, and what are the mechanisms for its transfer between these domains? What role does media dissemination play in spread of AI-facilitated cybercrime? How can individual cyber hygiene practices be improved to reduce risks associated with AI-based threats? (Shetty et al., 2024) There was not a clear hypothesis said but the study pursues exploratory objectives, such as cataloging malicious prompts, identifying dissemination patterns, and eliciting expert perspectives. The independent variables in the article are the type of AI tools and prompts, such as ChatGPT and DAN prompts, forum locations like the dark web vs surface web, and media/discourse conditions like coverage and narratives. The dependent variables were measures of malicious activity prevalence and dissemination like percentages of malicious prompts, forum reach and expert-perceived risk/impact on victimization and cyber hygiene.

The types of research methods used in this study is a mixed method design. This means that quantitative and qualitative methods were used. Quantitative data includes content collection

and analysis with collected prompts from forums and descriptive statistics of forum locations, tool types, and content categories. Qualitative data includes semi-structured expert interviews transcribed and analyzed using thematic analysis. The type of data collected were 102 AI-generated malicious prompts from eight forums, some being Reddit, Youtube, Hidden Answers, and FlowGPT. Another form of data being interview transcripts and written statements from six experts. The analysis was also a mix of quantitative and qualitative methods. For quantitative methods, the study used descriptive statistics and correlation exploration between forum user counts and prompt frequency, tabulation of content types, some being malware, phishing, and jailbreak prompts. For qualitative methods, thematic analysis was used to extract themes about online lifestyles, guardianship, media narratives, regulation, and recommendations.

Relating to the Real World

From the powerpoint, the article uses skepticism by critically questioning media narratives that may exaggerate or distort AI risk. The article's research methods use archival because it analyzes existing dark-web and clear-web posts. In one powerpoint presentation, it shows that cybersecurity interacts with families, peer networks, schools, political institutions. The article expands on this by showing how media systems and online communities circulate harmful AI content. This topic relates to the challenges and contributions of marginalized groups by noting socioeconomic and vulnerability dimensions such as concerns about deepfakes and non-consensual uses that may disproportionately harm women and children. The global, multilingual nature of forums which points to cross-national and language barriers that can exacerbate inequalities in protection and reporting.

Conclusion

Overall, these studies make significant contributions to society by deepening our understanding of how emerging AI technologies, particularly large language models, are transforming the cyber threat landscape. By mapping and cataloging concrete evidence of malicious uses of LLMs and jailbreak prompts across both the dark and surface web, the research provides valuable insights into evolving cybercriminal behaviors. It also refines existing theory by demonstrating how AI tools reshape opportunity structures for cybercrime. Together, these contributions support a safer and more adaptive digital environment.

References

Shetty, S., Choi, K.-S., & Park, I. (2024). Investigating the Intersection of AI and Cybercrime: Risks, Trends, and Countermeasures. *International Journal of Cybersecurity Intelligence & Cybercrime*, *7*(2). https://doi.org/10.52306/2578-3289.1187

Link: https://vc.bridgew.edu/ijcic/vol7/iss2/3/