

Generative AI and Cybersecurity: An Activity Theory Perspective

Early stage paper

Hwee-Joo Kam
University of Tampa
hkam@ut.edu

Chen Zhong
University of Tampa
czhong@ut.edu

Wael Soliman
University of Agder
wael.soliman@uia.no

Allen Johnston
University of Alabama
ajohnston@cba.ua.edu

Abstract

Artificial Intelligence (AI) is widely used in cybersecurity for threat detection. The advent of Generative AI (GAI) extends the capabilities of cybersecurity software, further enhancing cybersecurity measures. This paper studies the intricacies of interactions between humans (i.e., cybersecurity professionals) and AI agents within the cybersecurity domain. Specifically, this paper addresses “complexity meets complexity”, in which the dynamics between complex GAI’s algorithm and complex tasks operated by cybersecurity professionals become the key to this study. Overall, this paper relies on Activity Theory, which offers a comprehensive framework to examine activities between human and AI agents. At this preliminary stage, the findings reveal that cybersecurity professionals leveraged GAI for knowledge exploitation in favor of productivity improvement, and that human-AI interactions engender human-AI augmentation, which is founded on the interplay between AI algorithms and human enquiries.

Keywords: *Generative AI, cybersecurity, human-AI interactions*

Introduction

Generative artificial intelligence (GAI) has gained popularity in the Information Systems (IS) discipline. IS scholars have discussed the ethical perspectives of GAI in societal

(Dwivedi et al., 2023; Stahl & Eke, 2024), organizational (Heyder et al., 2023), and academic research (Schlagwein & Willcocks, 2023) contexts, and examined GAI as a hacking tool (Renaud et al., 2023), an educational tool (Memmert et al., 2023; Van Slyke et al., 2023), and a research tool (Susarla et al., 2023), while uncovering factors behind GAI's adoption across a variety of organizations (Prasad Agrawal, 2023).

Cybersecurity professionals are more likely now than ever to incorporate GAI in their everyday tasks (Sen et al., 2022). Cybersecurity professionals, such as security analysts in a Security Operation Center (SOC), bear the responsibility of identifying cyber threats by scrutinizing intricate and interconnected IT infrastructures. This undertaking requires their interactions with complex tools, notably AI-based intrusion detection systems (IDS) designed for cyber defence. For example, it would not be unusual for SOC analysts to adopt GAI for assistance in testing, threat analysis, and data manipulation (Smith, 2018) associated with log analysis in a Security Information Event Management (SIEM) environment.

Despite the increased presence of GAI in the cybersecurity workspace, there are complexities native to both the human and GAI agents that add uncertainty and concern to the outcomes of their use (Jiang, Kahai, et al., 2022; Liu, 2021). These concerns are particularly troubling as the field of cybersecurity plays a pivotal role in safeguarding national security and critical infrastructure. For instance, humans cannot readily explain the *complex* outcomes generated by AI agents, potentially causing a lack of trust, biased AI outcomes, and a lack of accountability. Similarly, GAI agents do not take into consideration the individual traits of their human counterparts who operates *complex* cybersecurity tasks, potentially leading to information overloading, irrelevant information, and negative impacts on performance. This issue of “*complexity meets complexity*” carries significant weight in shaping any concerns organizational leaders may have toward the use of GAI by cybersecurity professionals in their

organizations. This study argues, however, that by better understanding how beneficial human-AI augmentations occur as a product of human-AI interactions, these concerns may be placated.

In light of this motivation, the research question is, *how are human-AI augmentations formed from the complex interactions of human and AI agents?* To answer this question, this study leverage Third Generation Activity Theory (Engeström, 1999). By collecting Reddit's data about cybersecurity workers' experiences of using ChatGPT, this study reveals how interactions between humans and AI agents engender human-algorithm interactions (Tarafdar et al., 2023) that eventually leads to human-AI augmentations (Rai et al., 2019). Such augmentations enhance both human and GAI knowledge bases and may serve to help pacify concerns associated when *complexity meets complexity*.

Literature Review

Artificial Intelligence (AI) methodologies, such as machine learning and deep learning, have been widely used in cybersecurity for threat detection (Ahmad et al., 2021), phishing detection (Catal et al., 2022), malware detection (Tayyab et al., 2022), and insider detection (Yuan & Wu, 2021). Recently, Generative AI (GAI) has emerged as a branch of AI that focuses on training models to generate new contents, such as, text, images, or other formats, in response to human enquiries (Susarla et al., 2023).

GAI holds the capability to assemble synthetic data, which is instrumental in testing detection models. Additionally, it can build comprehensive and realistic cybersecurity scenarios for training purpose. Large Language Models (LLM) represents a subset of GAI specifically designed for text generation (Devlin et al., 2018). While GAI offers invaluable tools for cybersecurity measures, threat actors use them for malicious purposes, such as automating social engineering (Brundage et al., 2018).

GAI raises some ethical concerns. Trained on vast amounts of data, GAI inherently reflects ethnocentric perspectives that can inadvertently breed societal biases (Susarla et al., 2023). Privacy breaches is another significant concern (Gupta et al., 2023). Also, a known limitation of GAI is “hallucination”, indicating its tendency to create nonsensical or inaccurate content (Nah et al., 2023). As the synergy between humans and AI becomes increasingly crucial, the ability to seamlessly collaborate with AI systems becomes essential. This calls for proficiency in *prompt engineering*, the process of carefully crafting prompts to optimize AI outcomes, which this study intends to address.

Activity Theory

Activity Theory (AT) highlights the dynamic interactions between various actors, such as humans and AI. Its primary objective is to discern essential concepts, like the behavioural model of these actors, while offering insights into the mechanisms underlying specific events. This entails exploring why and how individuals behave in particular ways (Engeström, 1999; Kaptelinin & Nardi, 2006). In AT, the fundamental unit of analysis is termed “activity”. Activity refers to the interplay between a subject and an object, where the subject represents an active entity (i.e., actors) driven by motives to transform the object (Kaptelinin & Nardi, 2006). Specifically, the object is intrinsically linked to the subject's motives, either serving as a source of motivation or fulfilling the subject's needs. The “triangle framework” (as illustrated in Figure 1) represents AT. This framework examines technologically mediated activities, encompassing aspects like technology utilization and interaction (Kaptelinin & Nardi, 2006).

Overall, AT comprises seven key elements: subject, object, instrument (i.e., mediating artifacts), rule, community, division of labour, and outcome. A subject is the social actor actively engaging in activities, while an object is the objectives pursued within an activity system. The community offers a social context in which the subject operates, forming an

integral part of the activity system's structure (Engeström, 1999). As subjects engage with their communities, they gradually realize the group norms and the division of labour, which outlines tasks assigned to other participants within the community. Using mediating artifacts (e.g., tools), subjects attain specific objectives relative to the community's goals. Ultimately, the outcome of this interplay is the transformation of the entire activity system.

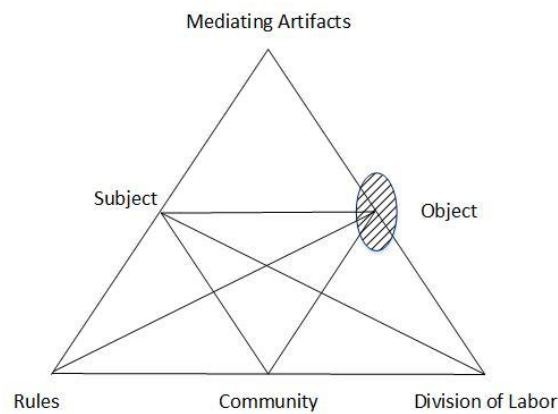


Figure 1. Activity Systems.

Third Generation Activity Theory

Third Generation AT, as shown in Figure 2, revolves around the dynamic interaction of two distinct systems, characterized by dialogue, diverse perspectives, and interrelated activities. This interaction ultimately creates a shared object (Engeström, 2001). To explain this concept, Engeström (2001) presented a healthcare scenario, where patients and the healthcare system acted as two separate but interconnected systems. Patients pursued the goal of healing (object₁), while the healthcare system offered healthcare services (object₁) (Engeström, 2001). As these two systems interacted, patients shared details about their illnesses to heal (object₂), and the healthcare system classified patients' diseases to attain accurate diagnoses (object₂). Eventually, both systems created new information that creates a well-informed diagnosis (object₃), which is a shared object resulting from their interactions.

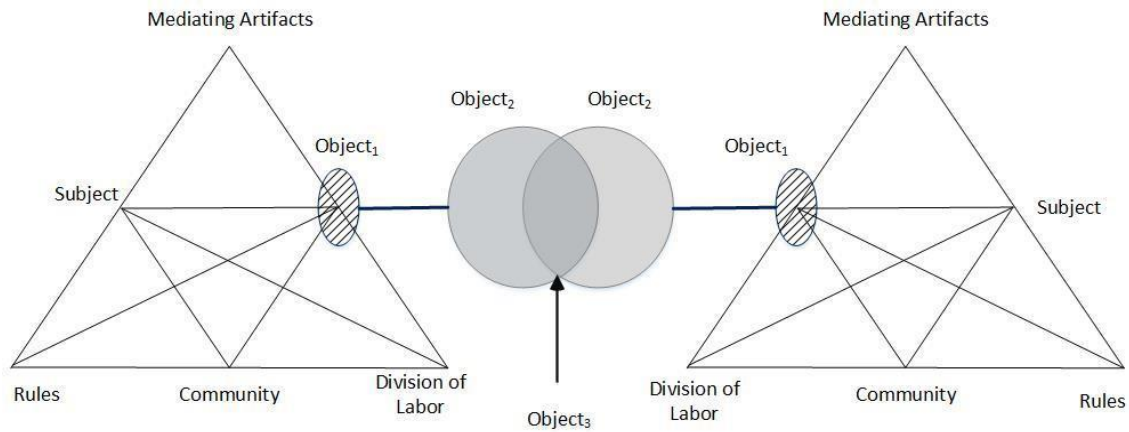


Figure 2. Third Generation Activity Theory.

The Third Generation of AT subsumes five guiding principles (Engeström, 2001). By adhering to these principles, one can systematically explore the theoretical components, such as rules (e.g., group norms), division of labour (e.g., automated tasks), and tools (e.g., AI algorithms) during human-AI interactions.

Principle 1: Unit of Analysis. The focus of analysis is the activity unit that creates actions. This includes actions initiated by both human and AI agents within their respective groups, as well as the interactions that occur between both groups.

Principle 2: Multi-Voicedness. Multi-voicedness encompasses a spectrum of divergent viewpoints, varied work outcomes, and diverse assessments of systems, all expressed by a multitude of individuals. Examining “multi-voicedness” within cybersecurity workers enables us to gain insights into their motivations, performance, and operational dynamics when utilizing GAI technologies.

Principle 3: Historicity. History can be perceived as a sequence of past actions undertaken with the aim of achieving specific objectives. Delving into historical records empowers us to discern recurring patterns, which would eventually uncover the social dynamics within activity systems.

Principle 4: Contradiction. Contradictions represent structural tensions within and across activity systems (Engeström, 2001). These tensions surface when novel circumstances collide with group norms. In a cybersecurity context, contradictions may arise when cybersecurity workers discover that AI (i.e., a new entity) fails to perform certain tasks as required (i.e., norms of cybersecurity operations).

Principle 5: Expansive Cycles. As contradictions escalate, a collaborative process of change emerges, producing object's transformation. For instance, human may find ways to navigate certain AI's constraints (i.e., contradiction); and on the other hand, the AI system could minimize its constraints and improve its responses to human enquiries, fostering enhanced human-AI interactions.

Research Methodology

Researchers in this study obtained data from Reddit using a custom crawler developed with the Pushshift.io Reddit API (Baumgartner, 2017/2023). The search focused on posts about cybersecurity professionals' usage about ChatGPT. The crawler gathered post details, including content, title, score, comment count, hyperlink, subreddit, and creation date. After applying predefined rules, duplicates and unrelated posts were removed, resulting in a dataset of 1,049 unique posts.

Based upon content analysis (Weber, 1990), NVivo software was used to run open coding. Content analysis is a method that draws inferences about psychological behaviors and communication styles in a group (Weber, 1990). It involves categorizing words within the text into fewer categories, facilitating the construction of concepts based on data rather than preconceived notions or biases (Agar, 1980).

When categorizing words into groups, it is important to consider the mutual exclusivity and scope of these categories (Weber, 1990). Using the Third Generation AT's five principles

as the guiding framework, this study argues that a core category derived from data is not strictly mutually exclusive since it can align with multiple principles. This lack of exclusivity is attributed to the interconnected nature of these principles, where the same text may fit into more than one principle. For example,

“I could ask a jailbroken version of the system for advice in committing a ransomware attack...”

On one hand, the text may suggest that individuals try to circumvent ChatGPT’s constraints (i.e., Principle 4: Contradiction). On the other hand, the text may suggest activities of ransomware attacks (i.e., Principle 1: Unit of Analysis). It is also worth noting that similar views were expressed using different words. These similarities could arise from the precise meanings of words or shared connotations (Weber, 1990). For instance, both sentences below link ChatGPT to cyber offense.

“With its advanced NLP and ability to mimic human behavior, ChatGPT can be used for social engineering attacks that are as cunning as they are convincing.”

“XSS attacks in webpages is still a commonly reported attack vector...you have [ChatGPT] with the accumulative cyber security expertise...this could potentially rear its head in an ugly way.”

With shared meanings found in text, this study asserts that content analysis along with a semiotic method (Maasik & Solomon, 2011) is a good data analysis approach. The semiotic method explores the meaning of signs and symbols, enabling a more comprehensive text analysis for drawing meaningful inferences (Myers, 1997). Overall, the data analysis began with open coding, where three researchers met via three Zoom meetings to identify core categories within the text. Disagreements that arose during this coding process were resolved through finding common ground. This study did not compute interrater reliability to assess the consistency of ratings among raters (Cooper & Schindler, 2001). This was due to the face-to-face discussions, which enabled researchers to find common ground and reduce potential biases.

Preliminary Findings

The following preliminary findings lay the foundations for a proposed framework depicted in Figure 3.

Principles	Findings	Suggestions
Principle 1: Activity system as unit of analysis	The key activity was using GAI to run cybersecurity related tasks. Individuals also used GAI to enhance their work productivity.	Individuals use GAI for <i>task automation</i> , with a specific focus on refining work operations to enhance value, as described by the notion of <i>exploitation</i> (Johnson et al., 2022)
Principle 2: Multi-voicedness	Individuals were impressed by GAI's ability to exhibit human-like behaviors and offer accurate responses, making them to believe that GAI as a useful cybersecurity tool. However, they expressed concerns about biases stemming from GAI's reliance on historical data, and they were bothered by occasionally irrelevant and inaccurate responses (Nah et al., 2023).	GAI show <i>data dependence</i> (Weber et al., 2023), <i>anthropomorphism</i> (Sowa et al., 2021), and <i>context-awareness</i> , in which it adopts contextual information to address individuals' enquiries (Ogbuabor et al., 2022). But the complexities of AI (Yang et al., 2020) may cause mishandling of individuals' enquiries, provoking <i>inscrutability</i> where AI becomes unintelligible (Berente et al., 2021).
Principle 3: Historicity	There was a recurring pattern wherein individuals projected GAI's technological frontier following their experiences with its usage.	AI's socio-technical trajectory is based on human-algorithm interaction, which is shaped by AI's <i>algorithm processing</i> and human's <i>enquires generation</i> .
Principle 4: Contradictions	A misfit between AI and individuals' needs (Jiang, Karran, et al., 2022) provoked tensions, propelling individuals to circumvent GAI's security controls.	Tensions reflect a struggle between human control and AI automation (Shneiderman, 2020) in that individuals try to assert <i>autonomy</i> over AI automation.
Principle 5: Expansive cycles	Through human-algorithm interactions, GAI learns to reinforce its security controls for deterring individuals' malicious use, while individuals learn to improve cybersecurity operations based on GAI's responses.	Human-algorithm interactions could foster <i>human-AI augmentation</i> , wherein humans and AI enhance their knowledge bases concurrently (Benbya et al., 2021).

Table 1. Summary of Data Analysis.

Figure 3 combines the activity systems of human and AI agents. Enquires processing by AI agents generates an object that interacts with human's enquiries, and likewise, enquires generation by human agents creates an object that interacts with AI's algorithms. This study argues that these interactivities gradually engender *human-AI augmentations*. Specifically, this study asserts that humans gain knowledge from AI-generated responses, while the high quality

of human enquiries enables AI systems to learn from human data. Through ongoing interaction between human and AI agents, humans gain a deeper understanding of the AI agents' strengths and learn how to effectively collaborate with them, a process often referred to as prompt engineering. Meanwhile, AI agents leverage insights gained from human behavior sequences to refine their decision-making abilities (Zhong et al., 2020). This dynamic suggests that interactions between humans and algorithms foster human-AI augmentations, enabling both humans and AI to expand their knowledge bases. This synergy between humans and AI signifies a symbiotic relationship between both entities, creating human-AI augmentations (Rai et al., 2019).

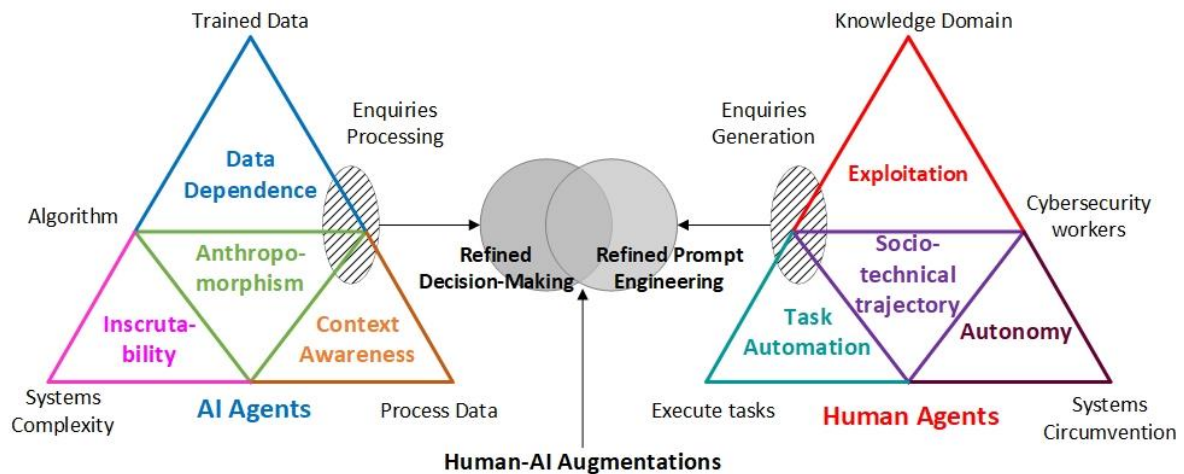


Figure 3. GAI's Activity Systems

Conclusion and Future Research

Based on the preliminary findings, this study argues that beneficial human-AI augmentations are the outcomes of complex human-algorithm interactions and knowledge repositories of both human and AI agents. That is, a synergistic effect between human and AI is shaped by human's and AI's knowledge base (Sundar, 2020), in which collective knowledge from both human and AI agents are disseminated, shared, and synthesized through human-AI interactions. Moreover, cybersecurity professionals mainly use GAI not so much for knowledge

exploration or creation but rather for knowledge exploitation favoring productivities improvement. By highlighting the beneficial human-AI augmentations shaped by human-AI interactions in the pursuit of cybersecurity task completion, this study hopes to help placate the concerns related to the *complexity meets complexity* phenomenon that is human-AI engagement. In the future, researchers will run Zoom meetings with cybersecurity professionals to observe how they use GAI to complete a given task and how their use affects the synergies gained for both the human and AI agents. This would further enrich the findings and offer a significant contribution to IS scholarship.

References

- Agar, M. H. (1980). *The professional stranger: An informal introduction to ethnography*. Academic Press.
- Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1), e4150. <https://doi.org/10.1002/ett.4150>
- Baumgartner, J. M. (2023). *Pushshift Reddit API Documentation* [Python]. <https://github.com/pushshift/api> (Original work published 2017)
- Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems*, 22(2), 281–303. <https://doi.org/10.17705/1jais.00662>
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly*, 45, 1433–1450. <https://doi.org/10.25300/MISQ/2021/16274>
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitsoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., ... Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (arXiv:1802.07228). arXiv. <https://doi.org/10.48550/arXiv.1802.07228>
- Catal, C., Giray, G., Tekinerdogan, B., Kumar, S., & Shukla, S. (2022). Applications of deep learning for phishing detection: A systematic literature review. *Knowledge and Information Systems*, 64(6), 1457–1500. <https://doi.org/10.1007/s10115-022-01672-x>
- Cooper, D. R., & Schindler, P. S. (2001). *Business Research Methods* (7th ed). Irwin/McGraw-Hill.

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. <https://doi.org/10.48550/ARXIV.1810.04805>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on Activity Theory* (1st ed., pp. 19–38). Cambridge University Press. <https://doi.org/10.1017/CBO9780511812774.003>
- Engeström, Y. (2001). Expansive Learning at Work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133–156. <https://doi.org/10.1080/13639080020028747>
- Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. *IEEE Access*, 11, 80218–80245. <https://doi.org/10.1109/ACCESS.2023.3300381>
- Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, 32(3), 101772. <https://doi.org/10.1016/j.jsis.2023.101772>
- Jiang, J., Kahai, S., & Yang, M. (2022). Who needs explanation and when? Juggling explainable AI and user epistemic uncertainty. *International Journal of Human-Computer Studies*, 165, 102839. <https://doi.org/10.1016/j.ijhcs.2022.102839>
- Jiang, J., Karran, A. J., Coursaris, C. K., Léger, P.-M., & Beringer, J. (2022). A Situation Awareness Perspective on Human-AI Interaction: Tensions and Opportunities. *International Journal of Human-Computer Interaction*, 0(0), 1–18. <https://doi.org/10.1080/10447318.2022.2093863>
- Johnson, P. C., Laurell, C., Ots, M., & Sandström, C. (2022). Digital innovation and the effects of artificial intelligence on firms’ research and development – Automation or augmentation, exploration or exploitation? *Technological Forecasting and Social Change*, 179, 121636. <https://doi.org/10.1016/j.techfore.2022.121636>
- Kaptelinin, V., & Nardi, B. A. (2006). *Acting with Technology: Activity Theory and Interaction Design*. The MIT Press.
- Liu, B. (2021). In AI We Trust? Effects of Agency Locus and Transparency on Uncertainty Reduction in Human–AI Interaction. *Journal of Computer-Mediated Communication*, 26(6), 384–402. <https://doi.org/10.1093/jcmc/zmab013>

- Maasik, S., & Solomon, J. (2011). *Signs of Life in the USA: Readings on Popular Culture for Writers*. Bedford/St. Martin's.
- Memmert, L., Tavanapour, N., & Bittner, E. (2023). Learning by Doing: Educators' Perspective on an Illustrative Tool for AI-Generated Scaffolding for Students in Conceptualizing Design Science Research Studies. *Journal of Information Systems Education*, 34(3), 279–292.
- Myers, M. D. (1997). Qualitative Research in Information Systems. *Management Information Systems Quarterly*, 21(2), 241–242.
- Nah, F. F.-H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304. <https://doi.org/10.1080/15228053.2023.2233814>
- Ogbuabor, G. O., Augusto, J. C., Moseley, R., & van Wyk, A. (2022). Context-aware system for cardiac condition monitoring and management: A survey. *Behaviour & Information Technology*, 41(4), 759–776. <https://doi.org/10.1080/0144929X.2020.1836255>
- Prasad Agrawal, K. (2023). Towards Adoption of Generative AI in Organizational Settings. *Journal of Computer Information Systems*, 0(0), 1–16. <https://doi.org/10.1080/08874417.2023.2240744>
- Rai, A., Constantinides, P., & Sarker, S. (2019). Next-Generation Digital Platforms: Toward Human–AI Hybrids. *MIS Quarterly*, 43(1), iii–ix.
- Renaud, K., Warkentin, M., & Westerman, G. (2023). From ChatGPT to HackGPT: Meeting the Cybersecurity Threat of Generative AI. *MIT Sloan Management Review*.
- Schlagwein, D., & Willcocks, L. (2023). 'ChatGPT et al.': The ethics of using (generative) artificial intelligence in research and science. *Journal of Information Technology*, 38(3), 232–238. <https://doi.org/10.1177/02683962231200411>
- Sen, R., Heim, G., & Zhu, Q. (2022). Artificial Intelligence and Machine Learning in Cybersecurity: Applications, Challenges, and Opportunities for MIS Academics. *Communications of the Association for Information Systems*, 51(1). <https://doi.org/10.17705/1CAIS.05109>
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Smith, G. (2018). The intelligent solution: Automation, the skills shortage and cyber-security. *Computer Fraud & Security*, 2018(8), 6–9. [https://doi.org/10.1016/S1361-3723\(18\)30073-3](https://doi.org/10.1016/S1361-3723(18)30073-3)
- Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work: Human – AI collaboration in managerial professions. *Journal of Business Research*, 125, 135–142. <https://doi.org/10.1016/j.jbusres.2020.11.038>

- Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, 74, 102700. <https://doi.org/10.1016/j.ijinfomgt.2023.102700>
- Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1), 74–88. <https://doi.org/10.1093/jcmc/zmz026>
- Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus Effect of Generative AI: Charting the Path for Responsible Conduct of Scholarly Activities in Information Systems. *Information Systems Research*, 34(2), 399–408. <https://doi.org/10.1287/isre.2023.ed.v34.n2>
- Tarafdar, M., Page, X., & Marabelli, M. (2023). Algorithms as co-workers: Human algorithm role interactions in algorithmic work. *Information Systems Journal*, 33(2), 232–267. <https://doi.org/10.1111/isj.12389>
- Tayyab, U.-H., Khan, F. B., Durad, M. H., Khan, A., & Lee, Y. S. (2022). A Survey of the Recent Trends in Deep Learning Based Malware Detection. *Journal of Cybersecurity and Privacy*, 2(4), Article 4. <https://doi.org/10.3390/jcp2040041>
- Van Slyke, C., Johnson, R., & Sarabadani, J. (2023). Generative Artificial Intelligence in Information Systems Education: Challenges, Consequences, and Responses. *Communications of the Association for Information Systems*, 53(1), 1–21. <https://doi.org/10.17705/1CAIS.05301>
- Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. (2023). Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI. *Information Systems Frontiers*, 25(4), 1549–1569. <https://doi.org/10.1007/s10796-022-10297-y>
- Weber, R. P. (1990). *Basic Content Analysis*. SAGE. <https://dx.doi.org/10.4135/9781412983488>
- Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3313831.3376301>
- Yuan, S., & Wu, X. (2021). Deep learning for insider threat detection: Review, challenges and opportunities. *Computers & Security*, 104, 102221. <https://doi.org/10.1016/j.cose.2021.102221>
- Zhong, C., Yen, J., & Liu, P. (2020). Can Cyber Operations Be Made Autonomous? An Answer from the Situational Awareness Viewpoint. In S. Jajodia, G. Cybenko, V. S. Subrahmanian, V. Swarup, C. Wang, & M. Wellman (Eds.), *Adaptive Autonomous Secure Cyber Systems* (pp. 63–88). Springer International Publishing. https://doi.org/10.1007/978-3-030-33432-1_4