# AI in Cybersecurity: The Paradox

## I. AI and the Organizational Cybersecurity Landscape

MODERN artificial intelligence is inherently paradoxical in many ways. While AI aims to increase automation, it also requires more intimate human involvement to reflect on the insights generated (automation paradox). While AI results in job displacement, it also creates new jobs, some simply to provide the necessary support systems for those newly unemployed (transition paradox). And as generative AI takes away control over the creative process, it also offers new creative opportunities (creativity paradox). This article considers another paradox, that relates to the fact that computational systems created using AI can be used both for public good in civilian applications and for harm across a range of application areas and settings [A1]. This contradiction is explored within an organizational and governmental context, where modern AI relies on data which might be externally or internally-sourced [A2]. External data sources [A3] are inclusive of open-source intelligence (OS-INT), such as information available on the Internet and the dark web, and internal data sources may include proprietary data found within an organizational or a wider governmental context [A4]. A further relevant consideration is the expanding role of the Internet of Things to support smart infrastructures, which has created new vulnerabilities [A5].

From a beneficial perspective, an organization can conduct asset monitoring by using artificial intelligence to automatically detect patterns of, for example, anomalous access to network assets. Such AI-enabled security systems can detect the specifics of offensive tactics employed and react with commensurate defensive actions [A6]. From an unfavorable perspective, assets lacking appropriate security are at risk of adverse attempts by hackers, which may go undetected until it is too late, leaving private data compromised or unavailable. Additionally, when the pattern of a novel attack vector is not familiar, then it is desirable that the security system can actively learn and respond with limited, if any, human intervention [A7], [A8].

Indeed, when modern AI is used as a competitive advantage in the context of organizational cybersecurity [A9], there are typically four phases: (1) understanding the cybersecurity lifecycle in place and its major principles, tasks, and processes; (2) collecting and aggregating available internal and external data from a variety of sources; (3) applying AI-enabled analytics to the data collected, allowing inferences and predictions to be made about patterns and trends that may require (urgent) attention; and (4) using and disseminating knowledge on an as required basis to support the cybersecurity goals of an entity. Here we are seeking to comprehend the data that is generated by assets and the metadata that may be available, to conduct near real-time data analysis that may shed light on cybersecurity attacks just-in-time. Although we are still practically some way off in this automated cyberwarfare becoming an every-day reality, we are observing the rapid development of next generation solutions for security using AI. In this way, AI can be used to protect an organization's physical and logical boundaries by automating some aspects of the recognition of cybersecurity threats and vulnerabilities. Previously, these aspects would have typically either gone undetected by even the most well-trained human security agent or required days of algorithmic development to combat a continuously morphing attack.

In known exploits, the AI is a better 'watcher' and 'seeker' than even a human agent, always looking out for exceptions on a network, and even autonomously deciding on the best course of action. But AI is not infallible. It is prone to the same attacks as other computing systems, and unexpected variations can render them ineffective, often granting security analysts a false sense of security. It is evident that security reporting can be enhanced by AI, providing a company's security manager with timely information, for example, that a network shutdown is imminently required before a ransomware attack is unleashed [A10], [A11] or that some relevant intervention is required to ensure protection of valuable data assets, such as customer data or intellectual property.

## II. Cybersecurity Strategies and Methodologies: Toward a Socio-Technical Perspective

When considering how AI can be incorporated into cybersecurity, the first step is to consider what kinds of data is collected by the entity, and what data emanates from assets in a network or in any system these components may interact with to gather personal information [A12]. Patterns of individual behavior can be derived from network access information, and customer data. It is the latter that is particularly valuable to hackers who wish to breach corporate, government or not-for-profit third sector databases [A13]. For example, hackers increasingly provide evidence of their exploits by presenting samples of customer records with sensitive details that have been accessed [A14], and then if a ransom remains unpaid, the whole data set is uploaded to the dark web [A15]. To mitigate such attacks, companies with critical infrastructure need the cybersecurity team actively engaged in the procurement process to proactively consider what kinds of technologies and assets they are investing in within the organization and their provenance. In the same way, organizational

Digital Object Identifier 10.1109/TTS.2023.3280109

data and inter-organizational data sharing needs in the cloud must have commensurate service level agreements (SLAs) ensuring protection of customer data [A16]. One of the fundamental problems with data emanating from system assets is that the data may be open and unencrypted. Once a hacker gains access to a particular record of information that does not need to be decrypted, then that data is viewable in its totality. Additionally, even when data on assets is encrypted, there are means to de-crypt, and even re-identify once de-identified personal information.

The OODA loop (Observe, Orient, Decide, Act) is one methodology that machine learning specialists have implemented in the context of commercial security lifecycles, although it was primarily developed for operational level military campaigns, and now adopted in a variety of contexts [A17]. In the first step of "Observe", a variety of different types of network traffic are monitored. Given the amount of data traffic being sent and received today, the human security analyst does not have the same propensity as a machine to detect non-traditional traffic patterns. These anomalous patterns are called exceptions in a network and may involve measuring abnormal levels of traffic, wireless signal strength, biometric recognition authentication attempts, and even the number of API calls [A18]. In the second step, "Orient", patterns may emerge that are identifiable by machines and humans, requiring some form of tactical cybersecurity adaptation, to curb the impending threat. In the third step, AI in cybersecurity specialists may "Decide" to write algorithms in just a few hours, that can observe incoming and outgoing data traffic in a given context, helping to detect anomalous sensor or network-based traffic patterns in milliseconds, and ongoingly report on these through situational awareness databases. In the fourth and final step "Act", these AI-based algorithms can be run at short intervals without human intervention but may demand human action in response to a machine-determined alert linked to an exception. Some propose the level of accuracy of such algorithms are at 99% detection [A17], but the human-in-the-loop may well detect and act on that crucial 1% of activity that goes undetected by a machine and could ultimately be catastrophic to an organization, if not discovered in a timely manner.

While the OODA loop might well be instituted at the tactical level for the incorporation of AI for cybersecurity, the strategic level perspective of the overall socio-technical system, within which the AI and cybersecurity strategies and methodologies are embedded, remains an important consideration in view of security that goes beyond the AI or technological element. This is so, as AI is *not* the silver bullet solution that will end all cybersecurity woes, as we will continue to experience sophisticated social engineering attacks on people, and emergent machine attacks on neural networks. AI is merely one aspect of a larger socio-technical system that incorporates humans (inclusive of non-cybersecurity specialists), human relationships, tools and techniques, the environment, and their corresponding interrelationships [A19]. As a result, we are likely to experience a multitude of socio-technical challenges, some of which are due to AI being analogous to a double-edged sword; a technology that has or can have both favorable and unfavorable consequences [A20]. These consequences are paradoxical in nature, requiring a fundamental understanding of the complex socio-technical setting, inclusive of who is steering the AI, and *for* and *against* whom, allowing for an enhanced understanding of the rationale behind its application.

## III. The Paradoxical Nature of AI

During the *9th International Conference on Ethics in Biology, Engineering & Medicine* in Florida in 2018, Jonathan D. Moreno referred to AI as an "offset", comparing it to a handful of technologies, the first of which was the atomic bomb itself [A21], [A22]. Some might believe that Moreno was overstating the potential use and impact of AI, but increasingly leading voices in the space have called for bans for generative AI [A23], have signed petitions for a pause on AI development [A24], and have suggested slower innovation cycles to ensure we have safeguards for how we are deploying AI [A25]. But if we are to consider Moreno's analogy of AI akin to an atomic bomb, what might well be the fallout if AI is left unchecked [A26], especially in cybersecurity applications? This question could and should be analyzed from a variety of perspectives and at multiple levels to better understand the risks associated with personal and local community-level security, organizational security, and even national, supranational, and international security [A27].

Whereas previously our technological systems were not reliant on cyber-physical-social infrastructures, today there are few human-made standalone systems that do not rely on internetworking, are not interdependent, and are not highly complex [A2]. Even a human as a complex living system can now be imbued with external extensions dependent, for example, on personal digital communications and data storage capabilities in the form of a smartphone or smartwatch, that have made the simple act of exchange- if not survival itself- subject to vulnerabilities and cybersecurity breaches. All these electronic interconnections between entities and their assets are potential points of failure, posing risks that are specific to a technological artifact such as an unexpected malfunction, or external in the form of interceptions from an AI bot. Such considerations and the challenges they present will only become more intricate as the operational configurations increase in complexity. An illustrative example of such complexity is in the deployment of brain implants for non-medical applications, and the potential use of other implantables in soldiers [A28].

Reflecting on Alan M. Turing's seminal contribution to analyzing the 'Enigma' Code, a basic cipher by today's standards, it enabled the development of Gordon Welchman's Bombe at Bletchley Park [A29] that could significantly reduce the work done by human code-breakers during the war [A30]. The situation today is fundamentally different, and it is not realistic to expect that a single system will be capable of responding to different forms of AI attacks on cybersecurity. Indeed, one must accept that the world is a far more complex place today, no longer founded on slow machines operating in isolation such as the Enigma but rather highly interconnected systems

of far greater power and capability, potentially each operating a unique flavor of AI.

This situation is only to become more challenging in the future. Consider for example the concept of self-replication first hypothesized by John von Neumann in 1948 [A31] and recently demonstrated through the development of Xenobots, i.e., programmed engineered organisms [A32], [A33]. This approach can be applied to self-replicating code in the form of AI that can learn and change its behavior through every interaction. It is worth reflecting on the limits to this kind of attack strategy, and how and whether it is possible to defend against such unknowns. While large language models (LLMs) have demonstrated numerous applications in the context of generative AI, it remains to be seen whether AI in cybersecurity will take on a similar path, autonomously. Where once we had self-replicating malware in the form of worms, self-contained programs that would create copies of themselves without user intervention utilizing networks to propagate onto other systems [A34], now we are referring to something with the ability to change its own state on execution. For every affirmative action AI can make there are a host of negative potentialities. These include data quality and bias concerns, privacy and content ownership concerns, environmental concerns, explainability challenges and unintended consequences [A34], [A35].

It seems no matter how we position AI, the very techniques and methods that can be applied to create an application with positive impact for humanity, can also be inverted, and appropriated to produce negative outcomes. Supplementary to scenarios of measured misapplication, are the unanticipated use cases, secondary uses and implications that emerge when a particular solution or service has been deployed. So, what then? Do we go at it alone, nation state against nation state to determine who can build the most sophisticated AI bots to defend against and attack their neighbor's critical infrastructure [A37]? Is it paradoxical to engage in the development of offensive AI applications for cybersecurity, while simultaneously investing heavily in next generation AI for education, safety, and care, with the purpose of building capacity among our public servants, professionals, blue collar workers, underrepresented persons in the labor force, our students, and our children [A38]? Will building improved technical approaches to cyber-defense and offence suffice?

It is quite clear that the best formal methods alone will not overcome the challenges relevant to cybersecurity. Scenarios are extremely relevant here as we ponder on the possibilities and the consequences of a given sequence of events occurring [A39]. Having built infrastructure contingent on the Internet, there is a responsibility to understand the complex landscape of AI in cybersecurity and be skilled and trained in planning for and developing a range of scenarios, inclusive of potential adverse uses of AI in the cybersecurity context. For example, it has been demonstrated that attackers can perturb machine learning models, rendering them to malfunction [A7]. As such, cybersecurity professionals and other relevant community stakeholders must develop a nuanced understanding of the paradoxical potential of AI in the cybersecurity context, a prerequisite of which is developing an appreciation of the nature of present-day socio-technical systems and their complex characteristics.

## IV. On Increased Complexity, Interconnectedness and Sophistication

The more complex our systems become [A40], the greater the attack plane that can be targeted. A tit-for-tat, 'catch me if you can' attitude, will only lead to greater exposures, and misdirected cybersecurity attacks with mass-scale, even global implications [A47]. Gamifying cybersecurity is not a solution either, because what is at stake are people's lives, their livelihoods, their associations, and dependencies [A41]. Organizations that make substantial investments to maintain a positive reputation, protecting customer data, raising capital, hiring labor, also do not wish to be subjected to penetration whether external or internal in nature. We need to discover and address the root causes of cybersecurity issues, which can only be achieved by exploring and analyzing the complex socio-technical system within which AI in cybersecurity, and the related challenges, exist. This does not require merely taking into consideration national and organizational-level risk assessment, but rather considering risk at the individual and or household level. Geopolitical pressures at the national level will have flow on effects, and governments must remain cognizant that interferences by state and non-state actors on critical infrastructure providers and major organizations, will have a direct impact on individual citizens and their respective households and communities at large. This fragmentation will require new architectures for international AI governance [A42]. Cybersecurity outages and exposures of any kind can have a long-lasting effect on trust, both in state and private-interest entity relationships, acting to destabilize people's mental models with respect to the competency and assurances of providers [A43].

## V. Humans in the Loop

Within an environment open to destabilization, and factoring in the multiplicity of scenarios, it is easy to assume that the future of AI in cybersecurity is one void of human intervention: an entirely autonomous vision. This is a consequential misconception when discussing the potential of AI in cybersecurity. That is, a human can, and in most instances should, be kept in the loop. Specialists must work with AI and keep striving for its appropriate and optimal use, and not become complacent or over-reliant on third party ML-based solutions. External data sources can provide new sources of intelligence with respect to the latest cybersecurity attacks, the development of new information on the latest forms of attack, and the construction of a customized cybersecurity knowledge repository that can act as an aid to decision-making for risk managers and security specialists [A44].

However, knowledge of the limitations and risks associated with these sources should be incorporated into cybersecurity strategies, and corresponding education and training programs. Additionally, a reframing is required in terms of how we envisage and define the future of AI in cybersecurity, one in which socio-technical considerations and an understanding of both complexity and the paradoxical nature of AI in cybersecurity are in the foreground. Simply put, AI can be utilized for both defense and offence, depending on execution and

operation; applying to use case scenarios targeting individuals, organizations, and nation states. The takeaway here is that we will always require skilled professionals in cybersecurity and risk management, and increasingly a multidisciplinary, even interdisciplinary, team where possible to account for the paradoxical nature of AI. It should never be about scenarios solely featuring AI vs AI alone, but rather an emphasis on the human who is an informed security analyst to know how best to utilize AI within a broader, complex socio-technical system and context [A45]. It is therefore a necessity to be aware of the required level of training and experience of a multiskilled cybersecurity team [A46], made up of more than cybersecurity computing professionals equipped with technical AI skills, as this will dictate the success of an organization's security function, and the relative impact on individuals, the organizations themselves and society.

## VI. Conclusion

Artificial intelligence is considered antithetical in that it can impact the world simultaneously in positive and negative ways. In the context of cybersecurity, this implies that AI can provide valuable tools to ensure public safety and at the same time be used for harm. This contradiction creates complex dynamics that cannot be resolved with simple solutions and moreover if they remain unchecked, can potentially cause significant harm. The increased complexity, interconnectedness and sophistication that characterize these AI-enabled systems may be interpreted as suggesting that humans may no longer have a role in their operation, management, and control. Yet, meaning-making still remains a uniquely human function and for this reason humans must continue to be firmly embedded in the decision-making process. In order to better understand "AI in Cybersecurity" as elaborated upon in the subject matter of this Special Issue, this editorial presented the paradox: too much AI in cybersecurity can increase vulnerability of a complex socio-technical system, and not enough AI in cybersecurity can lead to dire circumstances in an organizational or governmental setting. If we claim AI is the answer to our cybersecurity woes then there is a lack of acknowledgment of the detrimental risks that AI places on cybersecurity infrastructure, and if we say that AI is not the answer to at least some of our problems, then there is limited to no recognition that AI has altered the rules of engagement.

## Appendix: Related Articles


[A1] H. Ueno, "Artificial intelligence as dual-use technology," in *Fusion of Machine Learning Paradigms*. (Intelligent Systems Reference Library), vol. 236, I. K. Hatzilygeroudis, G. A. Tsihrintzis, and L. C. Jain, Eds. Cham, Switzerland: Springer, 2023. [Online]. Available: https://doi.org/10.1007/978-3-031-22371-6_2

[A2] S. Samtani, M. Kantarcioglu, and H. Chen, "Trailblazing the artificial intelligence for cybersecurity discipline: A multi-disciplinary research roadmap," *ACM Trans. Manage. Inf. Syst.*, vol. 11, no. 4, pp. 1–19, 2020.

[A3] J. Samuel, J. Jaskolka, and G. O. M. Yee, "Leveraging external data sources to enhance secure system design," in *Proc. Reconciling Data Anal. Autom. Privacy, Security Big Data Challenge (RDAAPS)*, Hamilton, ON, Canada, 2021, pp. 1–8, doi: 10.1109/RDAAPS48126.2021.9452029.

[A4] R. A. Guzzo, "Organizational data and its implications for research and theory," in *Data, Methods and Theory in the Organizational Sciences*. New York, NY, USA: Routledge, 2022, pp. 2–27.

[A5] C. A. White, "Root causes of insecure Internet of Things and holistically addressing them," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Las Vegas, NV, USA, 2020, pp. 1066–1074, doi: 10.1109/CSCI51800.2020.00198.

[A6] A. Jamalipour and S. Murali, "A taxonomy of machine-learning-based intrusion detection systems for the Internet of Things: A survey," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9444–9466, Jun. 2022, doi: 10.1109/JIOT.2021.3126811.

[A7] M. Usama, M. Asim, S. Latif, J. Qadir, and A. Al-Fuqaha, "Generative adversarial networks for launching and thwarting adversarial attacks on network intrusion detection systems," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, 2019, pp. 78–83. [Online]. Available: https://ieeexplore.ieee.org/document/8766353

[A8] P. F. de Araujo-Filho, G. Kaddoum, D. R. Campelo, A. G. Santos, D. Macêdo, and C. Zanchettin, "Intrusion detection for cyber–physical systems using generative adversarial networks in fog environment," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6247–6256, Apr. 2021.

[A9] J. M. Torres, C. I. Comesaña, and P. J. García-Nieto, "Review: Machine learning techniques applied to cybersecurity," *Int. J. Mach. Learn. Cybern.*, vol. 10, pp. 2823–2836, Jan. 2019. [Online]. Available: https://doi.org/10.1007/s13042-018-00906-1

[A10] M. Taddeo and L. Floridi, "Regulate artificial intelligence to avert cyber arms race," *Nature*, vol. 556, no. 7701, pp. 296–298, 2018.

[A11] I. Yaqoob et al., "The rise of ransomware and emerging security challenges in the Internet of Things," *Comput. Netw.*, vol. 129, pp. 444–458, Dec. 2017.

[A12] F. A. Qazi, "Study of zero trust architecture for applications and network security," in *Proc. IEEE 19th Int. Conf. Smart Communities Improving Qual. Life Using ICT, IoT AI (HONET)*, Marietta, GA, USA, 2022, pp. 111–116, doi: 10.1109/HONET56683.2022.10019186.

[A13] D. Kolevski, K. Michael, R. Abbas, and M. Freeman, "Cloud data breach disclosures: The consumer and their personally identifiable information (PII)?" in *Proc. IEEE Conf. Norbert Wiener 21st Century (CW)*, Chennai, India, 2021, pp. 1–9, doi: 10.1109/21CW48944.2021.9532579.

[A14] A. Foster. "Alleged hacker apologises to Optus after data of 10,000 customers reportedly released." Sep. 27, 2022. [Online]. Available: https://www.news.com.au/technology/online/hacking/alleged-optus-hacker-claims-10000-customer-records-leaked/news-story/618fa8fa7de7fea00e281958c36a67f4

[A15] J. Taylor. "Optus reveals at least 2.1 million ID numbers exposed in massive data breach." The Guardian. Oct. 2022. [Online]. Available: https://www.theguardian.com/business/2022/oct/03/optus-commissions-independent-review-of-data-breach

[A16] D. Kolevski and K. Michael, "Cloud computing data breaches a socio-technical review of literature," in *Proc. Int. Conf. Green Comput. Internet Things (ICGCIoT)*, 2015, pp. 1486–1495, doi: 10.1109/ICGCIoT.2015.7380702.

[A17] F. P. Osinga, *Science, Strategy and War: The Strategic Theory of John Boyd*. New York, NY, USA: Routledge, 2007.

[A18] V. Lenders. "The role of artificial intelligence in cyber-defence." Applied Machine Learning Days. EPFL. May 3, 2022. [Online]. Available: https://www.youtube.com/watch?v=GpvpSBaif6M

[A19] R. Abbas and K. Michael, "Socio-technical theory: A review," in *TheoryHub Book*, S. Papagiannidis Ed. Newcastle upon Tyne, U.K.: TheoryHub, 2023. [Online]. Available: http://open.ncl.ac.uk/

[A20] C. Benzaïd and T. Taleb, "AI for beyond 5G networks: A cyber-security defense or offense enabler?" *IEEE Netw.*, vol. 34, no. 6, pp. 140–147, Nov./Dec. 2020, doi: 10.1109/MNET.011.2000088.

[A21] J. D. Moreno, "Mind wars," presented at Int. Conf. Ethics Biol., Eng. Med., Miami, FL, USA, Apr. 2018.

[A22] N. E. Sondak and V. K. Sondak, "Neural networks and artificial intelligence," in *Proc. 20th SIGCSE Tech. Symp. Comput. Sci. Educ.*, 1989, pp. 241–245.

[A23] *Generative AI Carriers Serious Online Risks: Emerald Expert Briefings*, Oxford Analytica, Oxford, U.K., 2023.

[A24] L. Clarke, "Call for AI pause highlights potential dangers," *Science*, vol. 380, no. 6641, pp. 120–121, 2023.

[A25] J. L. Schenker. "Europe, China and the U.S. consider ways to restrain new AI models." The Innovator. Apr. 17, 2023. [Online]. Available: https://theinnovator.news/europe-china-and-the-u-s-consider-ways-to-restrain-new-ai-models/

[A26] S. Akter et al., "Algorithmic bias in data-driven innovation in the age of AI," *Int. J. Inf. Manage.*, vol. 60, Oct. 2021, Art. no. 102387.

[A27] S. Kowalski and J. Mwakalinga, "Modelling the enemies of an IT security system—A socio-technical system security model," in *Proc. 2nd Int. Multi-Conf. Complexity, Inform. Cybern.*, 2011, pp. 251–256.
[A28] K. Michael, "DARPA's ADAPTER Program: Applying the ELSI approach to a semi-autonomous complex socio-technical system," in *Proc. IEEE Conf. Norbert Wiener 21st Century (21CW)*, 2021, pp. 1–10.
[A29] C. Severance, "Alan Turing and Bletchley park," *Computer*, vol. 45, no. 6, pp. 6–8, Jun. 2012, doi: 10.1109/MC.2012.197.
[A30] D. Lenton, "Rebuilding the Bombe [Enigma code breaking machine]," *IEE Rev.*, vol. 47, no. 6, pp. 7–10, Nov. 2001, doi: 10.1049/ir:20010601.
[A31] M. Sipper, "Fifty years of research on self-replication: An overview," *Artif. Life*, vol. 4, no. 3, pp. 237–257, Jul. 1998, doi: 10.1162/106454698568576.
[A32] E. Ramanujam, L. Rasikannan, P. A. Anandhalakshmi, and N. A. Kamal, "Xenobots: A remarkable combination of an artificial intelligence-based biological living robot," *Int. J. Sociotechnol. Knowl. Devel.*, vol. 14, no. 1, pp. 1–11, 2022.
[A33] P. Ball, "Living robots," *Nat. Mater.*, vol. 19, no. 3 p. 265, 2020.
[A34] "Computer security: Virus highlights need for improved Internet management," United States Gen. Account. Office, Washington, DC, USA, Rep. GAO/IMTEC-89-57, Jun. 1989.
[A35] I. Ozkaya, "Application of large language models to software engineering tasks: Opportunities, risks, and implications," *IEEE Softw.*, vol. 40, no. 3, pp. 4–8, May/Jun. 2023, doi: 10.1109/MS.2023.3248401.
[A36] R. Pringle, K. Michael, and M. G. Michael, "Unintended consequences of living with AI: The paradox of technological potential?" *IEEE Technol. Soc. Mag.*, vol. 35, no. 4, pp. 17–21, Dec. 2016, doi: 10.1109/MTS.2016.2632978.
[A37] J. Dennis, C. Grady, and S. Rajtmajer, "Comparative assessment of cyber-physical threats to megacities," in *Proc. IEEE Int. Symp. Technol. Soc. (ISTAS)*, Waterloo, ON, Canada, 2021, p. 1, doi: 10.1109/ISTAS52410.2021.9629170.
[A38] B. Dupont, "Cybersecurity futures: How can we regulate emergent risks?" *Technol. Innov. Manage. Rev.*, vol. 3, no. 7, pp. 6–11, 2013.
[A39] B. C. Jantzen, "Cyberwarfare," in *Spaces for the Future*. New York, NY, USA: Routledge, 2017, pp. 141–153.
[A40] L. Robertson, A. M. Aneiros, and K. Michael, "A theory of exposure: Measuring technology system end user vulnerabilities," in *Proc. IEEE Int. Symp. Technol. Soc. (ISTAS)*, 2017, pp. 1–10.
[A41] G. Génova, V. M. Pelayo, and M. R. G. Martín, "A lesson from AI: Ethics is not an imitation game," *IEEE Technol. Soc. Mag.*, vol. 41, no. 1, pp. 75–81, Mar. 2022, doi: 10.1109/MTS.2022.3147531.
[A42] P. Cihon, M. M. Maas, and L. Kemp, "Fragmentation and the future: Investigating architectures for international AI governance," *Global Policy*, vol. 11, no. 5, pp. 545–556, 2020.
[A43] J. M. Bauer and W. H. Dutton, "The new cybersecurity agenda: Economic and social challenges to a secure Internet," in *World Development Report 2016 Digital Dividends*. Washington, DC, USA: World Bank, 2016, pp. 1–35. [Online]. Available: https://openknowledge.worldbank.org/handle/10986/7735
[A44] S. Zeadally, E. Adi, Z. Baig, and I. A. Khan, "Harnessing artificial intelligence capabilities to improve cybersecurity," *IEEE Access*, vol. 8, pp. 23817–23837, 2020.
[A45] T. Stevens, "Knowledge in the grey zone: AI and cybersecurity," in *Digital War*, vol. 1. Cham, Switzerland: Springer, 2020, pp. 164–170.
[A46] J. R. S. Blair, A. O. Hall, and E. Sobiesk, "Educating future multidisciplinary cybersecurity teams," *Computer*, vol. 52, no. 3, pp. 58–66, Mar. 2019, doi: 10.1109/MC.2018.2884190.
[A47] K. Cukier. "Babbage: What if generative AI destroys biometric security?" Babbage From the Economist [podcast]. May 2023. [Online]. Available: https://shows.acast.com/theeconomistbabbage/episodes/babbage-what-if-generative-ai-destroys-biometric-security

KATINA MICHAEL
School for the Future of Innovation in Society
Arizona State University
Tempe, AZ 85281 USA

School of Computing and Augmented Intelligence
Arizona State University
Tempe, AZ 85281 USA
E-mail: katina.michael@asu.edu

ROBA ABBAS
School of Business
University of Wollongong
Wollongong, NSW 2522, Australia
E-mail: roba@uow.edu.au

GEORGE ROUSSOS
Department of Computer Science and Information Systems
Birbeck College
University of London
WC1E 7HX London, U.K.
E-mail: g.roussos@bbk.ac.uk

**Katina Michael** (Senior Member, IEEE) is a Professor with Arizona State University and a Senior Global Futures Scientist with the Global Futures Laboratory and has a joint appointment with the School for the Future of Innovation in Society and the School of Computing and Augmented Intelligence, Arizona State University. Prior to academia, she was employed with Nortel Networks, Anderson Consulting, and OTIS Elevator Company. She has been funded by the National Science Foundation, the Canadian Social Sciences and Humanities Research Council, and the Australian Research Council. She is the Director of the Society Policy Engineering Collective and the Founding Editor-in-Chief of the IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY. She is also the Founding Chair of the inaugural Master of Science in Public Interest Technology.

**Roba Abbas** (Member, IEEE) is a Senior Lecturer of Operations and Systems with the Faculty of Business and Law, University of Wollongong, Australia, and was a Visiting Professor with the School for the Future of Innovation in Society, Arizona State University, USA. Her research is focused on methodological approaches to complex socio-technical systems design, emphasizing transdisciplinarity, co-design and the intersection of society, technology, ethics, and regulation. She is also the Co-Editor-in-Chief of the IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY, a Former Associate Editor of the *IEEE Technology and Society Magazine*, and the Technical Committee Chair of the Socio-Technical Systems Committee of the IEEE.

**George Roussos** (Senior Member, IEEE) received the B.S. degree in mathematics from the University of Athens, Athens, Greece, the M.S. degree in numerical analysis and computing from the University of Manchester Institute of Science and Technology, Manchester, U.K., and the Doctor of Philosophy degree from the Imperial College of Science Technology and Medicine, University of London, London, U.K. Before joining as a Lecturer with the Birkbeck College, University of London, he worked as a Research and Development Manager for a multinational information technology corporation in Athens, Greece, where he was responsible for the strategic development of new IT products in the areas of knowledge management and mobile Internet. As an Internet Security Officer with the Ministry of Defense, Athens, he designed the Hellenic armed forces Internet exchange and domain name systems; and as a Research Fellow with Imperial College, London, U.K., he conducted research in distributed systems. He is currently investigating the effects of social activity on system architectures and exploring mechanisms to support navigation and findability.