Socio-Technical Ecosystem Considerations: An Emergent Research Agenda for AI in Cybersecurity

I. Introduction

CYBERSECURITY continues to be a key element of national security. Given that critical infrastructure is powered by digital systems, any disruption to the flow of vital services has the potential effect of causing harm to citizens, organizations, and government entities. The availability of service offerings may be subject to attacks by AI on various aspects of cyber-physical-social systems. And no system is foolproof. Inasmuch as AI can be a threat to cybersecurity, it can also be used to combat attempts at systems penetration through deep learning and other capabilities. However, focusing our attention on AI alone does not allow for a sufficiently accurate understanding of the complex ecosystem that contributes to the conflicting scenarios that are possible in the context of AI in cybersecurity. This Special Issue provides insights into the broader socio-technical ecosystem considerations that are required for the achievement of system robustness, allowing for a more intricate and balanced understanding of the role of AI in cybersecurity, at various levels relevant to individuals, organizations, and communities/ society. Fundamentally, the socio-technical approach calls for an emphasis on people, processes, and technology, and not solely on the shiny gadgetry that we call products. There are four main themes that form the emergent research agenda: 1) the cybersecurity ecosystem that demands an interdisciplinary approach; 2) the state of the art technology development processes and landscape that provide and facilitate cybersecurity mechanisms and competitive advantage through AI; 3) the interconnected and multifaceted nature of the social, technical, and environmental subsystems in the cybersecurity ecosystem; and 4) the emphasis on human requirements and values, inclusive of end-user capacity and awareness, organizational capability and skills development, and societal considerations inclusive of the public interest.

II. Setting the Scene

We live in a saturated, information-rich environment with unprecedented dependence on digital technologies. An element of the expansion of digital technologies is a shift in Artificial Intelligence (AI) technology from research laboratories into the hands of anyone with a digital device (e.g., a smartphone) [1]. AI-powered search, personalization, automation, and augmentation are being deployed across sectors, from education to healthcare, policing, and finance. Wide AI diffusion is then reshaping the way organizations, communities, and individuals function [2]. This trend is being further accelerated by the recent rise of generative AI and large language models.

The potentially radical consequences of AI have impelled nation states across the globe to publish strategies on how they seek to shape, drive, and leverage the disruptive capabilities offered by AI technologies to bolster their prosperity and security [3]. In the context of new partnerships, and within existing alliances, these efforts can be seen as an opportunity for positive alignment so that governance and new capabilities create value for citizens' well-being, privacy [4] and safety [5]. Those same national efforts to lead, nurture and sustain AI to transform citizens' lives can also be viewed as competition, or even as an international AI arms race undermining international stability [6]. Policy initiatives for the governance of AI in the security and defense domains focus on potential security breaches, economic consequences, and political threats [7]. There is an evident disregard for social and environmental factors and dimensions, and the interplay with the technological domain. This lack of attention may shape how AI could be used in cybersecurity for harm, beyond the organizational level, and systems of governance that may or may not respect the rule of law [8]. Furthermore, AI systems themselves introduce new targets for malicious actors.

There has been a resurgence within academia and associated specialist scientific institutes to investigate socio-technical factors, in a broad sense shaping cybersecurity. But there has still been limited focus on the complex external environment and dynamic socio-cyber-physical ecosystem [9]. Most security research relates to the traditional Confidentiality-Integrity-Availability (CIA) triad [10]. While this strategy has continued to strengthen organizational and infrastructural defenses, we must consider the new emergent threats. These include homogeneity in products at their core operating system, large storage area network providers and critical telecommunication exchanges and international banking interchanges, and the supply of electricity and water, in addition to the respective interdependencies [11], [12]. Of particular importance are autonomous systems leveraging advanced machine learning systems that incorporate blackbox models (e.g., billion parameter neural networks), and highly complex technologies that may be microscopic and even embeddable and undetectable [13].

III. An Emergent Research Agenda

The socio-technical approach [14] is promising in the AI in cybersecurity context in that it advocates for a focus on stakeholder centricity in addition to acknowledge technological

Digital Object Identifier 10.1109/TTS.2023.3278908

aspects and dimensions [15]. This special issue invited socio-technical research focused on a deeper examination of value chain stakeholders [16], their roles and responsibilities and corresponding dynamic interactions and interdependencies in the present turbulent environment. Contributions to the special issue were focused on a range of questions. For example, how do different federal and state laws, regulations, policies, guidelines, and economic infrastructure shape the AI and cybersecurity landscape in an international context? How do multi-stakeholder approaches shed light on AI in cybersecurity considerations? How is AI and cybersecurity being applied as a potential global offset? How can cybersecurity specialists respond to these threats once they have been explicitly identified? What are the ethical considerations when incorporating AI into cybersecurity? And what are the governance challenges pertaining to AI?

A. Topics

Submissions were especially invited on intersecting topics of AI and/in Cybersecurity, topics that specifically tie to various facets of the socio-technical ecosystem; for instance, relating to social, technical, and environmental dimensions and considerations in an interlinked manner. This was a result of a review of scholarship that identified emergent research topics, which collectively form a socio-technical research agenda in the context of AI in cybersecurity. Topics include:

1) Responsible innovation, science, and technology ethics [17], [18]
2) Science and technology policy, regulation, and governance [19]
3) AI as double-edged sword [20], [21], [22], [23]
4) The application of socio-technical theory to complex problems related to emerging technology [24], [25], [26]
5) AI and cybersecurity ecosystems [27]
6) Design and innovation processes [28], [29], [30]
7) The security of AI algorithms in a socio-technical context [31]
8) Public understanding of and engagement with AI and cybersecurity, mindsets, mental models, capacity [32], [33]
9) Socio-technical cybersecurity frameworks [34]
10) Socio-technical security modelling and optimization [35], [36]
11) Developing organizational AI and cybersecurity capability models in vertical sectors, and cybersecurity skillsets [37]
12) Socio-technical futures, techno-scientific imaginaries, power, discrimination, contradiction [38], [39] [40]
13) The role of scenarios, vignettes, stories and qualitative approaches to AI and cybersecurity understanding [41], [42], [43]
14) Anticipatory/futures-literate approaches to AI and cybersecurity [44], [45], [50]
15) Algorithmic and technological biases and inequalities [46]
16) Impacts of AI and cybersecurity unleashed by nation states [47]
17) Impacts of AI and cybersecurity on nascent wearable and implantable technologies [48], [49], [51], [52]
18) Data/AI-driven cybersecurity for attack and defense [53]
19) Intelligence challenges related to AI and cybersecurity [54]
20) Holistic and exploratory approaches to AI- big picture national perspectives [55]
21) Public interest technologies in AI and cybersecurity [26]
22) The role of regulation and or soft/laws on the future practices of AI, considering both national (e.g., governance of AI) and international (AI for defense) perspectives [56], [57], [58], [59]
23) The role of education and training in raising societal awareness of cybersecurity threats [60], [61]
24) Opportunities and challenges for socio-technical systems enhancement, multidisciplinary approaches [62], [63].

B. Socio-Technical Themes

The four main themes that are identified from the literature forming an emergent research agenda of AI in cybersecurity need considerable attention in terms of further exploratory research, deeper examination of the opportunities and threats, theoretical extension to deal with the complexity of the unfolding challenges, and methodological rigor to better propose ways to overcome the challenges and reduce the risks. First, the cybersecurity ecosystem demands an interdisciplinary approach, and making this happen in practice will first require new approaches to training in the tertiary education sector, new approaches to hiring personnel that may not have traditional computer science skills (e.g., psychologists), and a broadening of the scope of the cybersecurity function in organizations and government. Second, state of the art technology development processes must be diffused especially at the meso level, where the vast majority of significant security breaches occur. Understanding the cybersecurity technology landscape better will provide and facilitate cybersecurity mechanisms and competitive advantage through AI. Third, the interconnected and multifaceted nature of the social, technical, and environmental subsystems in the cybersecurity ecosystem needs greater attention and to be understood as interconnected phenomena and not standalone subsystems. Finally, there needs to be an emphasis on human requirements and values, inclusive of end-user capacity and awareness, organizational capability and skills development, and national public stakeholder interests. Security is everybody's responsibility, but not everyone will have the same ability to enact change. For this reason, solutions targeted at different sectors of society will be increasingly important: raising individual human capacity to detect cybersecurity scams, for organizations to be better equipped with capabilities to deal with cybersecurity attacks, and for governments to develop their regulatory and governance functions nationwide, even internationally within an inter-governmental context.

IV. In This Special Issue

This special issue aimed to bring together researchers from different disciplines to explore the intersections of socio-technical futures, responsible innovation, and the role of AI

in cybersecurity as an exemplary area for inquiry and debate. Six papers were accepted into the Special Issue representing wide-ranging topics: from what can be considered micro-level applications in the form of brain-machine interfaces and facial recognition systems requiring direct human interaction, in the cyberspace context, followed by an interstate emphasis, after which issues of principles and broader governance measures are covered, in support of both meso- and macro- level emphases.

The first paper [A1] is by Juris Doctors Lucille Nalbach Tournas and Walter G. Johnson of the Sandra Day O'Connor College of Law at Arizona State University. Tournas and Johnson investigate the realm of novel neurotechnologies in the form of brain-computer interfaces (BCIs), that aim to offer new treatment options for those living with mental or neurological diseases. Legal tools to regulate and manage these advancements are underdeveloped and the authors propose soft law mechanisms for management of the development of BCIs. They note that such soft tools need to be guided by standards and codes of conduct that respect the privacy, agency, identity, and dignity of individuals and communities. Three examples are reviewed, across intergovernmental, civil society, and standard-setting bodies. Furthermore, neurotechnologies along with any embeddable "inside-the-body" technologies pose new challenges to governance when human ICT appliances are invisible [64].

The second paper [A2] is by Sankini Rancha Godage, Sushma Venkatesh, Kiran Raja, Raghavendra Ramachandra, and Christoph Busch all of whom are with the Norwegian University of Science and Technology (NTNU), and Frøy Løvåsdal who is with the Norwegian Police Directorate. The paper focuses on emergent morphing attack detection mechanisms and examines how these attacks fuse two or more facial images into one, with the final image resembling the contributed faces. This research challenges a prevalent misconception that human examiners' or observers' capacity for facial morph detection depends on their expertise and experience with facial recognition systems. In fact, the analysis offers intriguing insights as to the failure to recognize a sizable number of morphing attacks by experienced observers. Human observers also tend to detect morphed images to a lower accuracy as compared to the four automated morphing attack detection algorithms evaluated in this work. The study intends to aid in the development of training programs to prevent security failures. Generative AI applications that are image-based are increasingly producing high quality images that have been morphed between one or more real people, inclusive of people that do not even exist. It is possible that systems and human observers alike may be duped by such AI capabilities, providing more than one bona-fide identity access to a physical or virtual space.

The third paper [A3] is written by practicing engineer Catherine B. Smith. Influence operations in cyberspace raise questions about how narratives are strategically disseminated and circulated online, and about whether state-of-the-art machine learning (ML) techniques for narrative understanding and automated narrative generation are or will soon become part of adversarial nation-states' arsenal. Already we are witnessing the importance of narrative in impacting the mindsets and mental models of everyday citizens; some of these campaigns are so successful that they impact peoples' actions and behaviors [65]. Smith emphasizes that to accurately assess the threat of ML we need to clarify some ambiguities surrounding narratives in cyberspace. As one example, she asks how to define "narrative" in a way that makes sense across the organizations and academic disciplines that are involved in defending against disinformation? Smith also questions what blind spots exist in our shared lexicon that have stymied the United States' response to disinformation attacks so far? The author posits a new systems-dynamic model of narrative in cyberspace is required and demonstrates the use of the model in the context of an existing case study in disinformation.

The fourth paper [A4] is written by Thomas Reinhold, Philipp Kuehn, and Christian Reuter from the Science and Technology for Peace and Security (PEASEC) group and Daniel Günther and Thomas Schneider from the Cryptography and Privacy Engineering Group (ENCRYPTO) at the Technical University of Darmstadt. The authors of this paper examine malicious cyber-operations that may threaten global IT security. Advancements in the field of artificial intelligence have accelerated such malicious cyber operations, with the use of artificial intelligence-enabled cyber weapons, automated cyber defense measures, and artificial intelligence-based threat and vulnerability detection. State actors, with their long-term strategic security interests, often stockpile knowledge of vulnerabilities and exploits to enable their military or intelligence service cyberspace operations. The authors propose a privacy-preserving approach to enhancing global IT security that allows multiple state parties to privately compare their stock of vulnerabilities and exploits to check for items that occur in multiple stockpiles without revealing them. The ExTRUST approach is particularly useful when vulnerabilities are found in some organizations throughout a supply chain and the exposure may be carried through to all members of that supply chain [66].

The fifth paper [A5] is by ten members of Australia's national science agency and innovation catalyst, the Commonwealth Scientific and Industrial Research Organization (CSIRO). The paper is led by Conrad Sanderson. The authors of this paper aim to address a gap which remains between high-level principles and practical techniques that can be readily adopted to design and develop responsible AI systems. This paper presents outcomes from semi-structured interviews concerning practices and experiences of researchers and engineers from CSIRO, involved in the design and development of AI systems across application areas, such as environmental health and monitoring, health and wellbeing infrastructure management, industry innovation, and ensuring AI is ethical [67]. These align with a set of high-level AI ethics principles published by the Australian Government [68]. The principles noted in the CSIRO outcomes are: (1) privacy protection and security, (2) reliability and safety, (3) transparency and explainability, (4) fairness, (5) contestability, (6) accountability, (7) human-centered values, and (8) human, social and environmental well-being.

In the sixth and final paper [A6], Matti Minkkinen and Matti Mäntymäki tackle AI governance by discerning between 'easy' and 'hard' problems. It is a fitting conclusion to the Special Issue as the authors note that while there is widespread consensus that artificial intelligence (AI) needs to be governed, scholarly discussion on AI governance is dispersed among numerous disciplines and problem domains. The authors aim to clarify this situation by distinguishing between two problem areas, metaphorically titled the "easy" and "hard" problems of AI governance. The "easy problem" of AI governance concerns how organizations' design, development, and use of AI systems align with laws, values, and norms stemming from legislation, ethics guidelines, and the surrounding society. The "hard problem" of AI governance concerns AI as a general-purpose technology that transforms organizations and societies. The authors note that while societies should not lose track of the "hard problem" of AI governance, there is significant value in solving the "easy problem," a) because it can be provisionally solved by tackling bias, harm, and transparency issues, and b) because solving the "easy problem" contributes to solving the "hard problem", as responsible organizational AI practices will ideally create virtuous rather than vicious cycles.

ACKNOWLEDGMENT

The authors would like to thank the Alan Turing Institute and the National Cybersecurity Centre (NCSC) in the United Kingdom for their ongoing support, especially the NCSC's Sociotechnical and Risk Group (StRG), who helped to organize six workshops that ran from January to June 2021 from which this special issue idea was born.

We would also like to thank Terri Bookman for her editorial assistance.

APPENDIX: RELATED ARTICLES

[A1] L. N. Tournas and W. G. Johnson, "Regulating brain–computer interfaces: Ensuring soft law doesn't go flat," *IEEE Trans. Technol. Soc.*, early access, Sep. 26, 2022, doi: 10.1109/TTS.2022.3208821.

[A2] S. R. Godage, F. Løvåsdal, S. Venkatesh, K. Raja, R. Ramachandra, and C. Busch, "Analyzing human observer ability in morphing attack detection—Where do we stand?" *IEEE Trans. Technol. Soc.*, early access, Dec. 22, 2022, doi: 10.1109/TTS.2022.3231450.

[A3] C. B. Smith, "The semantic attack surface: A systems-dynamic model of narrative in cyberspace," *IEEE Trans. Technol. Soc.*, early access, Sep. 29, 2022, doi: 10.1109/TTS.2022.3210782.

[A4] T. Reinhold, P. Kuehn, and C. Reuter, "ExTRUST: Reducing exploit stockpiles with a privacy-preserving depletion system for inter-state relationships," *IEEE Trans. Technol. Soc.*, submitted for publication.

[A5] C. Sanderson et al., "AI ethics principles in practice: Perspectives of designers and developers," *IEEE Trans. Technol. Soc.*, early access, Mar. 15, 2023, doi: 10.1109/TTS.2023.3257303.

[A6] M. Minkkinen and M. Mäntymäki, "Discerning between the 'easy' and 'hard' problems of AI governance," *IEEE Trans. Technol. Soc.*, early access, Apr. 17, 2023, doi: 10.1109/TTS.2023.3267382.

REFERENCES

[1] Y. Gil and B. Selman. "A 20-year community roadmap for artificial intelligence research in the U.S. computing community consortium (CCC) and association for the advancement of artificial intelligence (AAAI)." Accessed: Aug. 6, 2019. [Online]. Available: https://arxiv.org/ftp/arxiv/papers/1908/1908.02624.pdf

[2] U.K. Government. "AI sectoral deal." Accessed: May 21, 2019. [Online]. Available: https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal

[3] Select Committee on Artificial Intelligence of the National Science and Technology Council. "The national artificial intelligence research and development strategic plan: 2019 update (nitrd.gov)." Jun. 2019. [Online]. Available: https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf

[4] K. Michael, S. Kobran, R. Abbas, and S. Hamdoun, "Privacy, data rights and cybersecurity: Technology for good in the achievement of sustainable development goals," in *Proc. IEEE Int. Symp. Technol. Soc. (ISTAS)*, 2019, pp. 1–13.

[5] K. Michael and R. Abbas, "Responsible AI: Ensuring reliable, safe & trustworthy systems," in *Proc. 13th Workshop Soc. Implications Nat. Security (SINS)*, Mar. 2020. [Online]. Available: https://www.katinamichael.com/sins20

[6] K. Michael, R. Abbas, and J. Pitt, "Maintaining control over AI," *Issues Sci. Technol. Forum*, vol. 37, no. 3, p. 19, 2021. [Online]. Available: https://issues.org/debating-human-control-over-artificial-intelligence-forum-shneiderman/

[7] L. Helen, *A Sociotechnical Approach to Cyber Security: How a Multi-Disciplinary Approach can Help Us Deliver Security That Works in the Real World*, Nat. Cyber Security Centre, London, U.K., Jul. 2020. [Online]. Available: https://www.ncsc.gov.uk/blog-post/a-sociotechnical-approach-to-cyber-security

[8] M. C. Davis, R. Challenger, D. N. W. Jayewardene, and C. W. Clegg, "Advancing socio-technical systems thinking: A call for bravery," *Appl. Ergon.*, vol. 45, no. 2, pp. 171–180, 2014. [Online]. Available: https://doi.org/10.1016/j.apergo.2013.02.009

[9] R. Abbas and K. Michael, "The design and implementation of the COVID safe app in Australia: A socio-technical overview," in *Proc. Int. COVID Congr.*, Aug. 2020. [Online]. Available: https://www.katinamichael.com/seminars/2020/8/10/the-design-and-implementation-of-the-covidsafe-app-in-australia

[10] G. Dhillon and J. Backhouse, "Current directions in IS security research: Towards socio-organizational perspectives," *Inf. Syst. J.*, vol. 11, no. 2, pp. 127–153, 2001. [Online]. Available: https://doi.org/10.1046/j.1365-2575.2001.00099.x

[11] T. Bonaci, K. Michael, P. Rivas, L. J. Robertson, and M. Zimmer, "Emerging technologies, evolving threats: Next-generation security challenges," *IEEE Trans. Technol. Soc.*, vol. 3, no. 3, pp. 155–162, Sep. 2022, doi: 10.1109/TTS.2022.3202323.

[12] K. D. Stephan, K. Michael, M. G. Michael, L. Jacob, and E. P. Anesta, "Social implications of technology: The past, the present, and the future," *Proc. IEEE*, vol. 100, pp. 1752–1781, May 2012, doi: 10.1109/JPROC.2012.2189919.

[13] (Community, Computing, Consortium, Catalyst, Phoenix, Arizona). *Assured Autonomy: Path Toward Living With Autonomous Systems We Can Trust*. (Oct. 2020). [Online]. Available: https://cra.org/ccc/wp-content/uploads/sites/2/2020/10/Assured-Autonomy-Workshop-Report-Final.pdf

[14] R. P. Bostrom and J. S. Heinen, "MIS problems and failures: A socio-technical perspective, Part II: The application of socio-technical theory," *MIS Quart.*, vol. 1, no. 4, pp. 11–28, 1977.

[15] J. Pitt and J. Ober, "Democracy by design: Basic democracy and the self-organization of collective governance," in *Proc. IEEE 12th Int. Conf. Self-Adapt. Self-Org. Syst. (SASO)*, 2018, pp. 1–9, doi: 10.1109/SASO.2018.00013.

[16] P. Carayon, "Human factors of complex sociotechnical systems," *Appl. Ergon.*, vol. 7, no. 4, pp. 525–535, Jul. 2006. doi: 10.1016/j.apergo.2006.04.011.

[17] European Parliament. "The ethics of artificial intelligence: Issues and initiatives." Mar. 2020. [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

[18] K. Vogel, B. Balmer, S. Weiss, I. Kroener, M. Matsumoto, and R. Brian, *The Handbook of Science and Technology Studies*. Newcastle upon Tyne, U.K.: SAGE, 2016, pp. 973–1002.

[19] J. Pitt, J. Dryzek, and J. Ober, "Algorithmic reflexive governance for socio-techno-ecological systems," *IEEE Technol. Soc. Mag.*, vol. 39, no. 2, pp. 52–59, Jun. 2020, doi: 10.1109/MTS.2020.2991500.

[20] T. H. Davenport and R. Ronanki, "Artificial intelligence for the real world," *Harvard Bus. Rev.*, vol. 96, no. 1, pp. 108–116, 2018.

[21] K. Crawford, *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT, USA: Yale Univ. Press, 2021.

[22] M. Taddeo, T. McCutcheon, and L. Floridi, "Trusting artificial intelligence in cybersecurity is a double-edged sword," *Nat. Mach. Intell.*, vol. 1, no. 12, pp. 557–560, 2019.

[23] S. Murugesan, "The AI-cybersecurity nexus: The good and the evil," *IT Prof.*, vol. 24, no. 5, pp. 4–8, 2022.

[24] R. Abbas and K. Michael, *Socio-Technical Theory: A Review*, Univ. Newcastle, Newcastle, NSW, Australia, 2022.

[25] K. Michael and R. Abbas, "What happens to COVID-19 data after the pandemic? Socio-technical lessons," *IEEE Trans. Technol. Soc.*, vol. 3, no. 4, pp. 242–247, Dec. 2022, doi: 10.1109/TTS.2022.3226119.

[26] R. Abbas, J. Pitt, and K. Michael, "Socio-technical design for public interest technology," *IEEE Trans. Technol. Soc.*, vol. 2, no. 2, pp. 55–61, Jun. 2021, doi: 10.1109/TTS.2021.3086260.

[27] M. G. Jacobides, S. Brusoni, and F. Candelon, "The evolutionary dynamics of the artificial intelligence ecosystem," *Strategy Sci.*, vol. 6, no. 4, pp. 412–435, 2021.

[28] J. Ellul, "The technological order," *Technol. Culture*, vol. 3, no. 4, pp. 394–421, 1962.

[29] F. W. Geels, "Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study," *Res. Policy*, vol. 31, nos. 8–9, pp. 1257–1274, 2002.

[30] E. Mumford, "Socio-technical design: An unfulfilled promise or a future opportunity?" in *Proc. Int. Conf. Soc. Org. Perspect. Res. Practice Inf. Technol.*, Jun. 2000, pp. 33–46.

[31] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial intelligence and the 'good society': The U.S., EU, and U.K. approach," *Sci. Eng. Ethics*, vol. 24, no. 2, pp. 505–528, 2018.

[32] K. Michael and R. Abbas, "Lessons from COVIDSafe: Toward public interest technologies of the future," in *Proc. Int. COVID-19 Congr.*, Aug. 2020, p. 9. [Online]. Available: https://www.katinamichael.com/seminars/2020/8/10/lessons-from-covidsafe-toward-public-interest-technologies-of-the-future

[33] J. R. Schoenherr, R. Abbas, K. Michael, P. Rivas, and T. D. Anderson, "Designing AI using a human-centered approach: Explainability and accuracy toward trustworthiness," *IEEE Trans. Technol. Soc.*, vol. 4, no. 1, pp. 9–23, Mar. 2023, doi: 10.1109/TTS.2023.3257627.

[34] M. Malatji, S. Von Solms, and A. Marnewick, "Socio-technical systems cybersecurity framework," *Inf. Comput. Security*, vol. 27, no. 2, pp. 233–272, 2019. [Online]. Available: https://doi.org/10.1108/ICS-03-2018-0031

[35] U. D. Ani, J. Watson, N. Tuptuk, S. Hailes, and A. Jawar. "Socio-technical security modelling: Analysis of state-of-the-art, application, and maturity in critical industrial infrastructure environments/domains." 2022. [Online]. Available: https://arxiv.org/pdf/2305.05108

[36] M. Malatji, *A Socio-Technical Systems Cybersecurity Optimisation Process: The Systems Engineering Management Approach*. Johannesburg, South Africa: Univ. Johannesburg, 2012. [Online]. Available: http://hdl.handle.net/102000/0002

[37] 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.

[38] S. Jasanoff and S. H. Kim, "Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea," *Minerva*, vol. 47, no. 2, p. 119, 2009.

[39] G. E. Marcus, *Technoscientific Imaginaries: Conversations, Profiles, and Memoirs*, vol. 2. Chicago, IL, USA: Univ. Chicago Press, 1995.

[40] J. Sadowski and R. Bendor, "Selling smartness: Corporate narratives and the smart city as a sociotechnical imaginary," *Sci. Technol. Human Values*, vol. 44, no. 3, pp. 540–563, 2019.

[41] R. Abbas, "Socio-technical theory: The role of scenarios in informing design choices. Threats and opportunities for AI in cybersecurity," in *Proc. Workshop Exam. Socio-Techn Ecosyst. (STeS) Considerations*, Feb. 2021, pp. 1–8.

[42] G. Lively, *Narratology*. Oxford, U.K.: Oxford Univ. Press, 2019.

[43] K. M. Vogel, "The need for greater multidisciplinary, sociotechnical analysis: The bioweapons case," *Stud. Intell.*, vol. 57, no. 3, pp. 1–10, 2013.

[44] V. Adams, M. Murphy, and A. E. Clarke, "Anticipation: Technoscience, life, affect, temporality," *Subjectivity*, vol. 28, no. 2, pp. 246–265, 2009.

[45] B. D. Johnson, "Science fiction prototyping: Designing the future with science fiction," *Synth. Lectures Comput. Sci.*, vol. 3, no. 1, pp. 1–190, 2011.

[46] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, p. 19, 2016.

[47] D. G. Dresner, T. Chappelow, M. Oldham, R. Mackenzie, and W. Roebuck, *U.K. National Security in a Digital World Inquiry*, U.K. Parliament, London, U.K., 2016. [Online]. Available: https://committees.parliament.uk/writtenevidence/78022/html/

[48] L. Perusco and K. Michael, "Control, trust, privacy, and security: Evaluating location-based services," *IEEE Technol. Soc. Mag.*, vol. 26, no. 1, pp. 4–16, Mar. 2007, doi: 10.1109/MTAS.2007.335564.

[49] D. Gokye and K. Michael, "Digital wearability scenarios: Trialability on the run," *IEEE Consumer Electron. Mag.*, vol. 4, no. 2, pp. 82–91, Apr. 2015, doi: 10.1109/MCE.2015.2393005.

[50] B. D. Johnson, "Science fiction for scientists! An introduction to SF prototypes and brain machines," in *Proc. Intell. Environ. (Workshops)*, Nov. 2010, pp. 195–203.

[51] K. Michael, "DARPA's ADAPTER program: Applying the ELSI approach to a semi-autonomous complex socio-technical system," in *Proc. 3rd 21st Century Wiener Conf.*, Jul. 2021, pp. 1–10.

[52] K. Cukier. "Babbage: Podcast: What if generative AI destroys biometric security?" The Economist. May 2023. [Online]. Available: https://www.economist.com/biometrics-pod

[53] N. Dhir, H. Hoeltgebaum, N. Adams, M. Briers, A. Burke, and P. Jones. "Prospective artificial intelligence approaches for active cyber defense." Apr. 20, 2021. [Online]. Available: https://arxiv.org/abs/2104.09981

[54] M. Taddeo, "Information warfare: A philosophical perspective," *Philosophy Technol.*, vol. 25, no. 1, pp. 105–120, 2012.

[55] M. C. Horowitz, G. C. Allen, E. Saravalle, A. Cho, K. Frederick, and P. Scharre, *Artificial Intelligence and International Security*, Center New Amer. Security, Washington, DC, USA, 2018.

[56] G. E. Marchant, "The growing gap between emerging technologies and the law," in *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight* (International Library of Ethics, Law and Technology), vol. 7. Dordrecht, The Netherlands: Springer, 2011, pp. 19–33. [Online]. Available: https://doi.org/10.1007/978-94-007-1356-7_2

[57] G. E. Marchant, K. W. Abbot, and B. Allenby, *Innovative Governance Models for Emerging Technologies*. Northampton, MA, USA: Edward Elgar, 2013.

[58] C. I. Gutierrez, G. E. Marchant, and K. Michael, "Effective and trustworthy implementation of AI soft law governance," *IEEE Trans. Technol. Soc.*, vol. 2, no. 4, pp. 168–170, Dec. 2021.

[59] G. I. Gutierrez, G. E. Marchant, and K. Michael, "Ideas on optimizing the future soft law governance of AI," *IEEE Technol. Soc. Mag.*, vol. 40, no. 4, pp. 10–13, Dec. 2021.

[60] T. Kohno and B. D. Johnson, "Science fiction prototyping and security education: Cultivating contextual and societal thinking in computer security education and beyond," in *Proc. 42nd ACM Techn. Symp. Comput. Sci. Educ.*, Mar. 2011, pp. 9–14.

[61] S. Willis, G. Byrd, and B. D. Johnson, "Challenge-based learning," *Computer*, vol. 50, no. 7, pp. 13–16, 2017.

[62] R. Abbas and A. Munoz, "Designing antifragile social-technical information systems in an era of big data," *Inf. Technol. People*, vol. 34, no. 6, pp. 1639–1663, 2021.

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

[64] R. E. Burnett, "The human information appliance in combat, intelligence, and diplomacy space," *IEEE Technol. Soc. Mag.*, hboxvol. 32, no. 2, pp. 39–51, Jun. 2013, doi: 10.1109/MTS.2013.2259651.

[65] G. H. Bower and D. G. Morrow, "Mental models in narrative comprehension," *Science*, vol. 247, no. 4938, pp. 44–48, 1990, doi: 10.1126/science.2403694.

[66] G. Dhillon and J. Backhouse, "Technical opinion: Information system security management in the new millennium," *Commun. ACM*, vol. 43, no. 7, pp. 125–128, 2000. [Online]. Available: https://doi.org/10.1145/341852.341877

[67] CSIRO. "AI at CSIRO, commonwealth scientific and industrial research organization." Accessed: May 18, 2023. [Online]. Available: https://www.csiro.au/en/research/technology-space/ai

[68] AG. "Australia's AI ethics principles." Accessed: May 17, 2023. [Online]. Available: https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles

MARIAROSARIA TADDEO
Oxford Internet Institute
University of Oxford
OX1 3JS Oxford, U.K.
Turing Fellow
The Alan Turing Institute
NW1 2DB London, U.K.
E-mail: mariarosaria.taddeo@oii.ox.ac.uk

PAUL JONES
The Alan Turing Institute
NW1 2DB London, U.K.
Email: pjones@turing.ac.uk

ROBA ABBAS
School of Business
University of Wollongong
Wollongong, NSW 2500, Australia
E-mail: roba@uow.edu.au

KATHLEEN VOGEL
School for the Future of Innovation in Society
Arizona State University
Tempe, AZ 85281 USA
Turing Fellow
The Alan Turing Institute
NW1 2DB London, U.K.
E-mail: kathleen.vogel@asu.edu

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

Mariarosaria Taddeo is an Associate Professor and a Senior Research Fellow with the Oxford Internet Institute, University of Oxford, and a Defence Science and Technology Ethics Fellow with Alan Turing Institute. Her research has been published in major journals like *Nature*, *Nature Machine Intelligence*, *Science*, and *Science Robotics*. Her work focuses mainly on digital ethics, the ethical analysis of artificial intelligence (AI), ethics of AI for national defence, cybersecurity, cyber conflicts, and ethics of digital innovation. She has received multiple awards, the 2010 Simon Award for Outstanding Research in Computing and Philosophy and the 2016 World Technology Award for Ethics. In 2018, InspiringFifty named her among the most inspiring 50 Italian women working in technology. ORBIT listed her among the top 100 women working on Ethics of AI in the world. She is one of the 12 2020 "Outstanding Rising Talents" named by the Women's Forum for Economy and Society. Since 2020, she has been serving in the Ethics Advisory Panel to the U.K. Ministry of Defence. Since 2016, She has been serving as an Editor-in-Chief for *Minds & Machines* (Springer Nature) and *Philosophical Studies Series* (Springer Nature).

Paul Jones (Member, IEEE) received the Ph.D. degree from North Carolina State University. He is a Visiting Research Scholar with the Alan Turing Institute, London, U.K., and also a Technical Director of Cybersecurity Research with the U.K. Government. He is an Associate with the Laboratory for Analytic Sciences, Raleigh, NC, USA. His main research interests are in artificial intelligence for cybersecurity, the security of AI itself, data-driven methods for understanding human sensemaking processes, and advancing human–machine collaborative intelligence in defence and security contexts.

Roba Abbas (Member, IEEE) is a Senior Lecturer of Operations and Systems with the Faculty of Business and Law, University of Wollongong, Australia. She 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 a 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.

Kathleen Vogel is a Professor with the School for the Future of Innovation in Society, Arizona State University. She is a 2023 Irregular Warfare Initiative Non-Resident Fellow, a joint collaboration of Princeton University's Empirical Studies of Conflict Project and the Modern War Institute at West Point. And previously, she was a Rutherford Fellow with The Alan Turing Institute from 2018 to 2019. She is the author of *Phantom Menace or Looming Danger?: A New Framework for Assessing Bioweapons Threats* (Johns Hopkins University, 2013). Her research focuses on the production of knowledge and big data in intelligence assessments. Her work has kept a close engagement between academia, intelligence, and policy. She is also a Co-Editor of the IEEE TRANSACTIONS ON TECHNOLOGY AND SOCIETY.

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.