Artificial Intelligence (AI) and Information Systems: Perspectives to Responsible AI

Denis Dennehy[1] · Anastasia Griva[2] · Nancy Pouloudi[3] · Yogesh K. Dwivedi[1,4] · Matti Mäntymäki[5] · Ilias O. Pappas[6,7]

1 Introduction

AI provides many transformational benefits to organisations across all industries and sectors (Alshahrani et al., 2021; Dennehy et al., 2022; Dwivedi et al., 2021; Elbanna et al., 2020; Vassilakopoulou et al., 2022). Recent studies have reported that AI can lead to new forms of business value (Enholm et al., 2022; Mikalef and Gupta, 2021), dynamic business-to-business relationships (Dwivedi and Wang, 2022; Keegan et al., 2022), enriched customer experiences (Jain et al., 2022; Griva et al., 2021; Kautish and Khare, 2022), enhanced human capabilities (Dwivedi et al., 2021), resilient supply chains (Zamani et al., 2022) and improved safety in the workplace (Gangadhari et al., 2022).

At the same time, there is a growing awareness of the risks and ethical issues surrounding AI (e.g., Bryson, 2018; Jobin et al., 2019) and the need to move from ethical principles to implementable practices (Schneiderman, 2021; Mäntymäki et al., 2022a; Seppälä et al., 2021) through for example the responsible design (Dennehy et al., 2021) and governance (Mäntymäki et al., 2022b) of AI systems. While organisations are increasingly investing in ethical AI and Responsible AI (RAI) (Zimmer et al., 2022), recent reports suggest that this comes at a cost and may lead to burnout in responsible-AI teams (Heikkilä, 2022). Thus, it is critical to consider how we educate about RAI (Grøder et al., 2022) and rethink our traditional learning designs (Pappas & Giannakos, 2021), as this can influence end-users' perceptions towards AI applications (Schmager et al., 2023) as well as how future employees approach the design and implementation of AI applications (Rakova et al., 2021; Vassilakopoulou et al., 2022).

The use of algorithmic decision-making and decision-support processes, particularly AI is becoming increasingly pervasive in the public sector, also in high-risk application areas such as healthcare, traffic, and finance (European Commission, 2020). Against this backdrop, there is growing concern over the ethical use and safety of AI, fuelled by reports of ungoverned military applications (Butcher and Beridze, 2019; Dignum, 2020), privacy violations attributed to facial recognition technologies used by the police (Rezende, 2022), unwanted biases exhibited by AI applications used by courts (Imai et al., 2020), and racial biases in clinical algorithms (Vyas et al. 2020). The opacity and lack of explainability frequently attributed to AI systems makes evaluating the trustworthiness of algorithmic decisions challenging even for technical experts, let alone the public. Together with the algorithm-propelled proliferation of misinformation, hate speech and polarising content on social media platforms, there is a high risk for erosion of trust in algorithmic systems used by the public sector (Janssen et al., 2020). Ensuring that people can trust in the algorithmic processes is essential not only for reaping the potential benefits from AI (Dignum, 2020) but also for fostering trust and resilience at a societal level.

AI researchers and practitioners have expressed their fears about AI systems being developed that are

✉ Denis Dennehy
denis.dennehy@swansea.ac.uk

1   Digital Futures for Sustainable Business & Society Research Group, School of Management, Swansea University, Wales, UK

2   J.E. Cairnes School of Business and Economics; Lero - the SFI Research Centre for Software, University of Galway, Galway, Ireland

3   ISTLab, Athens University of Economics and Business, Athens, Greece

4   Department of Management, Symbiosis Institute of Business Management, Pune & Symbiosis International (Deemed University), Pune, India

5   University of Turku, Turku, Finland

6   University of Agder, Agder, Norway

7   Norwegian University of Science and Technology, Trondheim, Norway

non-inclusive and enhance inequalities. There are known cases in which AI systems do not always make ethical or accurate choices (Babic et al., 2021) and biased or inaccurate data are used to train the AI algorithms which increase the risk of inequalities and injustice (Agrawal et al., 2020). For example, Amazon[1] trained their AI recruiting tool using masculine language, and thus the tool inherited bias against curriculum vitae submitted by women. This 'bias in – bias out' in AI models embeds the danger of inclusion. Nikon, is another example that illustrates this danger of inclusion. The company trained their AI model which identified people blinking excluding Asian people[2]. There are several examples of discrimination in AI applications which impose the need of critical thinking to question the AI results since it seems inevitable to completely regulate the AIs, which are in essence human opinions embedded in algorithms.

Researchers and practitioners state that these fears can be addressed and AI can be more inclusive – by designing 'human-AI hybrids' (Rai et al., 2019). In this context, researchers highlight the need to create Ambient Intelligent (AmI) AI systems to amplify AI-human collaboration (Gams et al., 2019). In such environments the AI system will interact with humans, receive information and learn from them and the environments (Ramos et al., 2008). From a different perspective converting the AI 'black boxes' to 'glass boxes'(Rai, 2020) and creating AI applications that inherit explainable features (XAI) can also facilitate inclusiveness in AI, as this transparency can make it easier to reduce the biases.

AI, like all technology, can be used in diverse ways and users may appropriate the technology in means that designers have not intended (Zamani and Pouloudi, 2020; Zamani et al., 2020). Thus, designers need to consider the intended and unintended consequences (Ransbotham et al., 2016; Majchrzak et al., 2016), by focusing on responsibility and ethical aspects to support this process. The Information Systems (IS) discipline has a sustained record of raising and addressing ethical concerns about IS, and technologies in general (e.g., Mason, 1986; Banerjee et al., 1998; Smith & Hasnas, 1999; Davison, 2000; Mingers & Walsham, 2010; Niederman, 2021). This special issue follows this cumulative tradition of academic discourse and knowledge by seeing vistas beyond technology (Stoodley et al., 2010), specifically AI.

---

[1] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

[2] http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html?_r=0.

2 The Special Issue

In this special issue, we were particularly interested in theory-building studies and empirically grounded theorising related to AI as a technology for an ethical and inclusive society. Following a rigorous review process consisting of a minimum of two and a maximum of four rounds of review, nine articles were selected to be included in this special issue. Each of the selected articles bring a distinct perspective to the emerging IS discourse on AI governance, ethics, and society. Collectively, the articles advance understanding of the socio-technical aspects of AI and its implications for society. The remainder of this editorial briefly describes the contributions that each of the selected articles made to advancing knowledge on AI for an ethical and inclusive society.

Niederman & Baker (2023) provide a reflective perspective on how ethical issues related to AI differ from other technologies. Specifically, they differentiate AI ethics issues from concerns raised by all IS applications by presenting three distinct categories of which AI ethics issues can be viewed. One can view AI as another IS application like any other. They examine this category of AI applications focusing primarily on Mason's (1986) PAPA framework, comprised of *privacy*, *accuracy*, *property*, and *accessibility*, as a way to position AI ethics within the IS domain. One can also view AI as adding a generative capacity to produce outputs that cannot be pre-determined from inputs and code. They examine this by adding 'inference' to the informational pyramid and exploring its implications. AI can also be viewed as a basis for re-examining questions of the nature of mental phenomena such as reasoning and imagination. At this time, AI-based systems seem far from replicating or replacing human capabilities. However, if/when such abilities emerge as computing machinery continues growing in capacity and capability, it will be helpful to have anticipated arising ethical issues and developed plans for avoiding, detecting, and resolving them to the extent possible.

Dattathrani & De (2023) make a strong argument that with the new generation of technologies, such as AI, the notion of agency needs to differentiate between the actions of AI from that of traditional information systems and humans. Indeed human and material agency have been investigated in the IS literature to understand how technology and humans influence each other. Some framings of agency, however, treat humans and technology symmetrically, some privilege the agency of humans over technology, and others do not attribute agency to either humans or non-humans. The authors introduce the dimensions of agency to differentiate agencies while not privileging any actor. They illustrate the application of dimensions by using it as a lens to study the case of a technician using an AI solution for screening patients

for early-stage breast cancer. Through the use of the dimensions of agency, they illustrate how the influence of AI over human practice, such as screening for early-stage breast cancer, is higher than the influence of traditional technology. Their study makes contributions to the theory of agency and concludes with a discussion on potential practical applications of the framework.

Harfouche et al., (2023) highlight that despite the hype surrounding AI, there is a paucity of research that focuses on the potential role of AI in enriching and augmenting organisational knowledge. The authors develop a recursive theory of knowledge augmentation in organisations (the KAM model) based on a synthesis of extant literature and a four-year revised canonical action research project. The project aimed to design and implement a human-centric AI (called Project) to solve the lack of integration of tacit and explicit knowledge in a scientific research centre (SRC). To explore the patterns of knowledge augmentation in organisations, this study extends Nonaka's knowledge management model which includes socialisation, externalisation, combination, and internalisation, by incorporating the human-in-the-loop Informed Artificial Intelligence (IAI) approach. Their proposed design offers the possibility to integrate experts' intuition and domain knowledge in AI in an explainable way. The findings show that organisational knowledge can be augmented through a recursive process enabled by the design and implementation of human-in-the-loop IAI. The study has important implications for both research and practice.

Koniakou (2023) engages in the discourse of AI governance from three angles grounded in international human rights law, namely, Law and Technology, Science and Technology Studies (STS), and theories of technology. The author posits that by focusing on the shift from ethics to governance, it offers a bird-eye view of the developments in AI governance, focusing on the comparison between ethical principles and binding rules for the governance of AI, and critically reviewing the latest regulatory developments. Further, by focusing on the role of human rights, it takes the argument that human rights offer a more robust and effective framework a step further, arguing for the necessity to extend human rights obligations to also directly apply to private actors in the context of AI governance. This study offers insights for AI governance borrowing from the Internet Governance history and the broader technology governance field.

Minkkinen et al., (2023) focus on addressing a gap in knowledge related to the governing AI which requires cooperation, yet the collaboration's form remains unclear. Technological frames provide a theoretical perspective for understanding how actors interpret technology and act upon its development, use, and governance. However, there is limited knowledge about how actors shape technological frames. The authors examine the shaping of the technological frame of the European ecosystem for responsible AI (RAI). Through an analysis of EU documents, we identified four expectations that constitute the EU's technological frame for the RAI ecosystem. Moreover, through interviews with RAI actors, we revealed five types of expectation work responding to this frame: reproducing, translating, and extending (congruent expectation work), and scrutinising and rooting (incongruent expectation work). The authors conceptualise expectation work as actors' purposive actions in creating and negotiating expectations. Their study contributes to the literature on technological frames, technology-centred ecosystems, and RAI while also elucidating the dimensions and co-shaping of technological frames.

Papagiannidis et al., (2023) highlight that despite the use of AI, companies still face challenges and cannot quickly realise performance gains. Adding to the above, firms need to introduce robust AI systems and minimise AI risks, which places a strong emphasis on establishing appropriate AI governance practices. In this paper, we build on a comparative case analysis of three companies from the energy sector and examine how AI governance is implemented to facilitate the development of robust AI applications that do not introduce negative effects. The study illustrates which practices are placed to produce knowledge that assists with decision-making while at the same time overcoming challenges with recommended actions leading to desired outcomes. The study contributes by exploring the main dimensions relevant to AI's governance in organisations and uncovering the practices that underpin them.

Polyviou & Zamani (2023) acknowledge that AI promises to redefine and disrupt several sectors. At the same time, AI poses challenges for policymakers and decision-makers, particularly regarding formulating strategies and regulations to address their stakeholders' needs and perceptions. This paper explores stakeholder perceptions as expressed through their participation in the formulation of Europe's AI strategy and sheds light on the challenges of AI in Europe and the expectations for the future. The findings reveal six dimensions of an AI strategy; ecosystems, education, liability, data availability sufficiency and protection, governance, and autonomy. It draws on these dimensions to construct a desires-realities framework for AI strategy in Europe and provide a research agenda for addressing existing realities. Their study advances the understanding of stakeholder desires on AI and holds important implications for research, practice, and policymaking.

Another interesting, yet theoretically underdeveloped application of AI is the use of AI-powered chatbots in the context of education and the experiences of students who use them. Chen et al., (2023) make the case that chatbots

are increasingly used in various scenarios such as customer service, work productivity, and healthcare, which might be one way of helping instructors better meet student needs. However, few empirical studies in the field of IS have investigated pedagogical chatbot efficacy in higher education, and fewer still discuss their potential challenges and drawbacks. In this research, the authors address this gap in the IS literature by exploring the opportunities, challenges, efficacy, and ethical concerns of using chatbots as pedagogical tools in business education. In this two-study project, they conducted a chatbot-guided interview with 215 undergraduate students to understand student attitudes regarding the potential benefits and challenges of using chatbots as intelligent student assistants. The findings of this study reveal the potential for chatbots to help students learn basic content in a responsive, interactive, and confidential way. The findings also provided insights into student learning needs which we then used to design and develop a new, experimental chatbot assistant to teach basic AI concepts to 195 students. Results of this second study suggest chatbots can be engaging and responsive conversational learning tools for teaching basic concepts and for providing educational resources. The authors discuss possible promising opportunities and ethical implications of using chatbots to support inclusive learning.

Despite the concerns raised by scholars and practitioners about AI, the pervasiveness of social recommender systems (SRSs) in e-commerce platforms highlights a trend that consumers who are willing to delegate their decisions to algorithms (Schneider & Leyer, 2019). SRS are increasingly becoming embedded in e-commerce ecosystems due to their ability to reduce consumers' decision time and effort by filtering out excess information and providing personalised recommendations (Tsai & Brusilovsky, 2021). As previous studies have largely focused on the technical aspects of the recommendation systems, there is limited understanding about the nature of the social information that improves the recommendation performance (Shokeen & Rana, 2020).

Bawack & Bonhoure (2023) investigate this phenomenon to identify the behavioural factors that influence consumers' intention to purchase products or brands recommended by SRSs. The authors adopt a meta-analytic research approach to conduct an aggregative literature review that uses quantitative methods to test specific research hypotheses based on prior empirical findings. Through the analysis of 72 articles, the authors identify 52 independent variables which are organised into 12 categories. Emerging from the analysis of the articles the authors propose a theoretical model on the behavioural factors that affect consumers' intentions to purchase products recommended by SRSs. As the study has important implications for research, the authors provide an agenda for future research that could advance

theory-building efforts and theory-driven designs in SRS research and practice.

Each of the articles of this special issue, as well as other recent studies (e.g., Akter et al., 2021; Bankins et al., 2022; Gupta et al., 2022; Shneiderman, 2021) have advanced knowledge on the ethical issues and governance of AI. Despite these important contributions, significant learning remains about how to use AI for social good (Ashok et al., 2022; Coombs et al., 2021; Dwivedi et al., 2021; Kumar et al., 2021; Fossa Wamba et al., 2021). To this end, we make a call for future research. First, there is a need for a concerted effort within and between academic disciplines (e.g., IS, arts, engineering), policymakers, governments, and the wider society to discover innovative ways to use AI to achieve the sustainable development goals (SDGs). Second, while significant attention has been given to understanding the application of AI in a variety of contexts, there is a limited discourse about how to use AI for future-oriented inquiry, whereby IS researchers can explore future scenarios through immersive virtual experiences to better understand how to design resilient IS and incorporate these insights in future-oriented inquiry (Brooks & Saveri, 2017; Chiasson et al. 2018). Third, future scholarship on AI governance could investigate auditing of AI systems (Minkkinen et al., 2022b) as a mechanism to foster transparency, accountability, and trust.

We hope that this special issue provides scholars with a foundation in which integrity and rigor for scientific research will promote high-quality IS, and ethical principles will translate into professional and organisational practice (Calzarossa et al., 2010; Mäntymäki et al., 2022a).

References

Agrawal, A., Gans, J., & Goldfarb, A. (2020). How to win with machine learning. *Harvard Business Review*, *98*(5), 126–133.

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, *60*, 102387.

Alshahrani, A., Dennehy, D., & Mäntymäki, M. (2021). *An attention-based view of AI assimilation in public sector organizations: the case of Saudi Arabia*. Government Information Quarterly.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, *62*, 102433.

Babic, B., Cohen, I. G., Evgeniou, T., & Gerke, S. (2021). When Machine Learning Goes Off the Rails.Harvard Business Review, (January-February).

Brooks, L. A., & Saveri, A. (2017). Expanding imagined affordance with futuretypes: Challenging algorithmic power with collective 2040 imagination. In Proceedings of the 50th Hawaii International Conference on System Sciences (HICCS).

Bryson, J. J. (2018). Patiency is not a virtue: the design of intelligent systems and systems of ethics. *Ethics and Information Technology*, *20*(1), 15–26.

Calzarossa, M. C., De Lotto, I., & Rogerson, S. (2010). Ethics and information systems - guest editors introduction. *Information Systems Frontiers*, *12*(4), 357–359. https://doi.org/10.1007/s10796-009-9198-4.

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial Intelligence (AI) student assistants in the Classroom: Designing Chatbots to Support Student Success. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10291-4.

Chiasson, M., Davidson, E., & Winter, J. (2018). Philosophical foundations for informing the future (S) through IS research. *European Journal of Information Systems*, *27*(3), 367–379.

Coombs, C., Stacey, P., Kawalek, P., Simeonova, B., Becker, J., Bergener, K., & Trautmann, H. (2021). What is it about humanity that we can't give away to intelligent machines? A european perspective. *International Journal of Information Management*, *58*, 102311.

Banerjee, D., Cronan, T. P., & Jones, T. W. (1998). Modeling IT ethics: a study in situational ethics. *MIS Quarterly*, *21*(1), 31–60.

Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context. *Information Systems Frontiers*, *24*, 1–19. https://doi.org/10.1007/s10796-021-10223-8.

Bawack, R. E., & Bonhoure, E. (2023). Influencer is the new recommender: insights for theorising social recommender systems. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10262-9.

Butcher, J., & Beridze, I. (2019). What is the state of artificial intelligence governance globally? *The RUSI Journal*, *164*(5–6), 88–96.

Dattathrani, S., & De, R. (2023). The Concept of Agency in the era of Artificial Intelligence: dimensions and degrees. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10336-8.

Davison, R. M. (2000). Professional ethics in information systems: a personal perspective. *Communications of the Association for Information Systems*, *3*(1), 8. https://doi.org/10.17705/1CAIS.00308.

Dennehy, D., Schmarzo, B., & Sidaoui, M. (2022). Organising for AI-powered innovation through design: the case of Hitachi Vantara. *International Journal of Technology Management*, *88*(2–4), 312–334. https://doi.org/10.1504/IJTM.2022.121507.

Dennehy, D., Pappas, I., Fosso Wamba, S., & Michael, K. (2021). Socially responsible Information Systems Development: the role of AI and business analytics, Editorial. *Information Technology & People*, *34*(6), 1541–1550. https://doi.org/10.1108/ITP-10-2021-871.

Dignum, V. (2020). Responsibility and artificial intelligence. The oxford handbook of ethics of AI, 4698, 215.

Dwivedi, Y. K., & Wang, Y. (2022). Guest editorial: Artificial intelligence for B2B marketing: Challenges and opportunities. *Industrial Marketing Management*, *105*, 109–113.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Williams, M. D. (2021). Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994.

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: a literature review. *Information Systems Frontiers*, *24*, 1709–1734. https://doi.org/10.1007/s10796-021-10186-w.

Elbanna, A., Dwivedi, Y., Bunker, D., & Wastell, D. (2020). The search for smartness in working, living and organising: beyond the 'Technomagic.'. *Information Systems Frontiers*, *22*(2), 275–280. https://doi.org/10.1007/s10796-020-10013-8.

Gams, M., Gu, I. Y. H., Härmä, A., Muñoz, A., & Tam, V. (2019). Artificial intelligence and ambient intelligence. *Journal of Ambient Intelligence and Smart Environments*, *11*(1), 71–86.

Gangadhari, R. K., Khanzode, V., Murthy, S., & Dennehy, D. (2022). Modelling the relationships between the barriers to implementing machine learning for accident analysis: the Indian petroleum industry.Benchmarking: An International Journal.

Griva, A., Bardaki, C., Pramatari, K., & Doukidis, G. (2021). Factors affecting customer analytics: evidence from three retail cases. *Information Systems Frontiers*, *24*, 493–516. https://doi.org/10.1007/s10796-020-10098-1.

Grøder, C. H., Schmager, S., Parmiggiani, E., Vasilakopoulou, P., Pappas, I., & Papavlasopoulou, S. (2022). Educating about Responsible AI in IS: Designing a course based on Experiential Learning. *International Conference on Information Systems (ICIS)*, 10, Copenhagen, Denmark.

Gupta, M., Parra, C., & Dennehy, D. (2022). 'Questioning racial and gender Bias in AI recommendations: do Individual-Level Cultural values Matter?'. *Information Systems Frontiers*, *24*, 1465–1481. https://doi.org/10.1007/s10796-021-10156-2.

Harfouche, A., Quinio, B., Saba, M., & Saba, P. B. (2023). The recursive theory of knowledge augmentation: integrating human intuition and knowledge in Artificial Intelligence to augment organizational knowledge. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10352-8.

Heikkilä, M. (2022). Responsible AI has a burnout problem. MIT Technology Review. October 28, 2022. https://www.technologyreview.com/2022/10/28/1062332/responsible-ai-has-a-burnout-problem/

Imai, K., & Jiang, Z. (2020). Principal fairness for human and algorithmic decision-making. *arXiv preprint arXiv:2005.10400*.

Jain, S., Basu, S., Dwivedi, Y. K., & Kaur, S. (2022). Interactive voice assistants–does brand credibility assuage privacy risks? *Journal of Business Research*, *139*, 701–717.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, *37*(3), 101493.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399.

Kautish, P., & Khare, A. (2022). Investigating the moderating role of AI-enabled services on flow and awe experience. *International Journal of Information Management*, *66*, 102519. https://doi.org/10.1016/j.ijinfomgt.2022.102519.

Keegan, B. J., Dennehy, D., & Naudé, P. (2022). Implementing artificial intelligence in traditional B2B marketing practices: an activity theory perspective. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-022-10294-1.

Kumar, P., Dwivedi, Y. K., & Anand, A. (2021). Responsible artificial intelligence (AI) for value formation and market performance in healthcare: the mediating role of patient's cognitive engagement. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-021-10136-6.

Koniakou, V. (2023). From the "rush to ethics" to the "race for governance" in Artificial Intelligence. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10300-6.

Mason, R. (1986). Four ethical issues of the information age. *MIS Quarterly*, *10*(1), 5–12.

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*, *58*(3), 103434.

Minkkinen, M., Zimmer, M. P., & Mäntymäki, M. (2023). Co-shaping an ecosystem for responsible AI: five types of expectation work in response to a Technological Frame. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10269-2.

Minkkinen, M., Laine, J., & Mäntymäki, M. (2022). Continuous auditing of Artificial Intelligence: A Conceptualization and Assessment of Tools and Frameworks. *Digital Society*, *1*(3), 1–27.

Mingers, J., & Walsham, G. (2010). Toward ethical information systems: the contribution of discourse ethics. *MIS Quarterly*, *34*(4), 833–854. https://doi.org/10.2307/25750707.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022a). Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance.arXiv preprintarXiv:2206.00335.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022b). Defining organizational AI governance.AI and Ethics,1–7.

Majchrzak, A., Markus, M. L., & Wareham, J. (2016). Designing for digital transformation

Lessons for information systems research from the study of ICT and societal challenges.MIS Quarterly, 40(2),267–277.

Niederman, F. (2021). Project management: openings for disruption from AI and advanced analytics. *Information Technology & People*, *34*(6), 15701599. https://doi.org/10.1108/ITP-09-2020-0639.

Niederman, F., & Baker, E. W. (2023). Ethics and AI issues: Old Container with New Wine? *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10305-1.

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: identifying best Practices and potential barriers and outcomes. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10251-y.

Pappas, I. O., & Giannakos, M. N. (2021, April). Rethinking learning design in IT education during a pandemic. *Frontiers in Education*, *6*, 652856.

Polyviou, A., & Zamani, E. D. (2023). Are we nearly there yet? A desires & realities Framework for Europe's AI strategy. *Information Systems Frontiers*, *25*(1), https://doi.org/10.1007/s10796-022-10285-2.

Ransbotham, S., Fichman, R. G., Gopal, R., & Gupta, A. (2016). Special section introduction-ubiquitous IT and digital vulnerabilities. Information Systems Research, *27*(4), 834–847.

Rai, A., Constantinides, P., & Sarker, S. (Eds.). (2019). Editor's comments: Next-Generation Digital Platforms: Toward Human-AI Hybrids, *MIS Quarterly,*. *43*(1), iii–ix.

Rai, A. (2020). Explainable AI: from black box to glass box. ï»¿*Journal of the Academy of Marketing Science*, *48*(1), 137–114.

Rakova, B., Yang, J., Cramer, H., & Chowdhury, R. (2021). Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–23.

Ramos, C., Juan, P., & Augusto, C. (2008). *Ambient Intelligence — the Next Step for Artificial Intelligence*. 15–18.

Rezende, I. N. (2022). Facial Recognition for Preventive Purposes: The Human Rights Implications of Detecting Emotions in Public Spaces. *Investigating and Preventing Crime in the Digital Era: New Safeguards, New Rights*, *7*, 67.

Schmager., S., Grøder, C., Parmiggiani, E., Pappas, I. O., & Vassilakopoulou, P. (2023). What do citizens think of AI adoption in public services? Exploratory research on citizen attitudes through a social contract lens. *56th Hawaii International Conference on System Sciences (HICSS)*, Maui, Hawaii.

Schneider, S., & Leyer, M. (2019). Me or information technology? Adoption of artificial intelligence in the delegation of personal strategic decisions. *Managerial and Decision Economics*, *40*(3), 223–231. https://doi.org/10.1002/mde.2982.

Seppälä, A., Birkstedt, T., & Mäntymäki, M. (2021). From ethical AI principles to governed AI. *In Proceedings of the 42nd International Conference on Information Systems (ICIS2021)*.

Shneiderman, B. (2021). Responsible AI: bridging from ethics to practice. *Communications of the ACM*, *64*(8), 32–35.

Shokeen, J., & Rana, C. (2020). A study on features of social recommender systems. *Artificial Intelligence Review*, *53*(2), 965–988.

Stoodley, I., Bruce, C., & Edwards, S. (2010). Expanding ethical vistas of IT professionals. *Information Systems Frontiers*, *12*(4), 379–387. https://doi.org/10.1007/s10796-009-9207-7.

Smith, H. J., & Hasnas, J. (1999). Ethics and information systems: the corporate domain. *MIS Quarterly*, *23*(1), 109–127. https://doi.org/10.2307/249412.

Vassilakopoulou, P., Haug, A., Salvesen, L. M., & Pappas, O., I (2022). Developing human/AI interactions for chat-based customer services: lessons learned from the norwegian government. *European Journal of Information Systems*, 1–13. https://doi.org/10.1080/0960085X.2022.2096490.

Vyas, D. A., Eisenstein, L. G., & Jones, D. S. (2020). Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. *New England Journal of Medicine*, *383*(9), 874–882.

Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, *164*, 120482. https://doi.org/10.1016/j.techfore.2020.120482.

Zamani, E. D., & Pouloudi, N. (2020). Generative mechanisms of Workarounds, Discontinuance and Reframing: a study of negative disconfirmation with consumerised IT. *Information Systems Journal*, *31*(3), 384–428. https://doi.org/10.1111/isj.12315.

Zamani, E. D., Pouloudi, N., Giaglis, G. M., & Wareham, J. (2020). Appropriating Information Technology Artefacts through Trial and Error: the case of the tablet. *Information Systems Frontiers*, *24*, 97–119. https://doi.org/10.1007/s10796-020-10067-8.

Zamani, E., Smyth, C., Gupta, S., & Dennehy, D. (2022). Artificial Intelligence and Big Data Analytics for Supply Chain Resilience: A Systematic Literature Review, *Annals of Operations Research, 1–28*

Zimmer, M. P., Minkkinen, M., & Mäntymäki, M. (2022). Responsible Artificial Intelligence Systems: Critical Considerations for Business Model Design, *Scandinavian Journal of Information Systems* (forthcoming).

Dr. Denis Dennehy is Associate Professor of Business Analytics and School Research Lead at the School of Management, Swansea University, Wales, UK. His research focuses on the mediating role of technologies and analytics, and their implications for people, organisations, and society. This research has been published in leading journals including *International Journal of Operations & Production Management, European Journal of Operational Research, Information Systems Frontiers, Information & Management, Information Technology & People, Government Information Quarterly, Annals of Operations Research, and IEEE*. He is a Senior Editor of *Information Technology & People*, and he has edited many special issues related to his field. His research is informed through extensive engagement with organizations, including Dell Technologies, Intel, Fexco, Leading Edge Group, Texuna, and Kepak Group.

Anastasia Griva is an Assistant Professor in Business Information Systems at the University of Galway (Galway), Ireland, and Programme Director at the MSc in Business Analytics. She has worked on research commercialization via establishing two AI and analytics start-ups which have raised more than 1.3M from VC funds. Her research interests lie in the areas of BA, AI, Innovation, and Customer Behaviour. She has published in edited books, academic journals (e.g., Information Systems Frontiers, IT & People, International Journal of Information Management, Journal of Retailing and Consumer Services, Expert Systems with Applications, IEEE Software, Journal of Decision Systems), and proceedings of international conferences (e.g., ECIS, MCIS, EURO). She has served as a Guest Editor for various journals (e.g., ISF, IJIM, JDS). For the last 2 years, she serves as track chair for the AI track at ECIS. Anastasia was also a Marie Skłodowska-Curie Postdoctoral Fellow recipient.

Nancy Pouloudi is Professor of Information Systems Management in the Department of Management Science and Technology at the Athens University of Economics and Business (AUEB), Greece, where she served as Department Chair. She is President-Elect of the Association for Information Systems (AIS) and will be serving as AIS President in 2023-2024. She was Region 2 Representative on AIS Council (2010-2013) and regularly serves on the AIS Region 2 Board and the Executive Committee of the Mediterranean Conference on Information Systems (MCIS). Her research focuses on organizational and social issues in IS adoption and implementation. Her recent work studies user post-adoption behavior (SIG Grounded Theory Methodology 2021 Best Paper Award Runner-Up) and the impact of cloud computing and the pandemic in the work and personal contexts of IS use. She has worked in a number of European Union funded projects on e-business, e-government, digital health and e-society, and is currently studying children's digital maturity (digymatex.eu) and gender equality in business schools (targeted-mpi.eu). Her service has been recognized by the 2016 AIS Sandra Slaughter Service Award and the 2020 AIS Technology Challenge Award.

Yogesh K. Dwivedi is a Professor of Digital Marketing and Innovation and Founding Director of the Digital Futures for Sustainable Business & Society Research Group at the School of Management, Swansea University, Wales, UK. In addition, he holds a Distinguished Research Professorship at the Symbiosis Institute of Business Management (SIBM), Pune, India. Professor Dwivedi is also currently leading the International Journal of Information Management as its Editor-in-Chief. His research interests are at the interface of Information Systems (IS) and Marketing, focusing on issues related to consumer adoption and diffusion of emerging digital innovations, digital government, and digital and social media marketing particularly in the context of emerging markets. Professor Dwivedi has published more than 500 articles in a range of leading academic journals and conferences that are widely cited (more than 47 thousand times as per Google Scholar). He has been named on the annual Highly Cited Researchers™ 2020, 2021 and 2022 lists from Clarivate Analytics. Professor Dwivedi is an Associate Editor of the Journal of Business Research, European Journal of Marketing, Government Information Quarterly and International Journal of Electronic Government Research, and Senior Editor of the Journal of Electronic Commerce Research.

Matti Mäntymäki is an Associate Professor of Information Systems Science (Tenure Track) at University of Turku, Finland and the head of the Digital Economy and Society (des.utu.fi) research platform. His research interests cover a broad range of psychosocial, organizational, and business implications of digitalization. He has authored more than 100 peer-reviewed papers, published in outlets such as *Information Systems Journal, Technological Forecasting & Social Change, International Journal of Information Management, Journal of Business Research, Information Systems Frontiers, Information Technology & People, Computers in Human Behavior, Journal of Systems & Software, and Communications of the Association for Information Systems*, among others.

Ilias Pappas is a Professor of Information Systems at the Department of Information Systems, University of Agder (UiA), Norway. His current research activities are within the area of Human-Centered AI (HCAI). He has been actively working in the areas of data science and digital transformation, social innovation and social change, user experience in different contexts, as well as digital marketing, e-services, and information technology adoption. He has published over 100 articles in peer reviewed journals and conferences including the *European Journal of Information Systems, Journal of Business Research, European Journal of Marketing, Information & Management, Psychology & Marketing, International Journal of Information Management, Journal of Systems and Software*. Pappas has been a Guest Editor for various journals (e.g., ISF, IJIM, IT&P, I&M, TFSC) and serves as an Associate Editor for several IS journals. He is or has been a track chair on AI as well as on Big Data Analytics at the *European Conference on Information Systems (ECIS)* and the *Americas Conference on Information Systems (AMCIS)*, among others. Pappas is a recipient of ERCIM and Marie Skłodowska-Curie fellowships.