



The dark sides of AI

Xusen Cheng¹ · Xiao Lin² · Xiao-Liang Shen³ · Alex Zarifis⁴ · Jian Mou⁵

© The Author(s), under exclusive licence to Institute of Applied Informatics at University of Leipzig 2022

This preface introduces the special issue on The Dark Sides of AI. The special issue offers six papers that focus on challenges of AI technology. In the twenty-first century, artificial intelligence (AI) is an extremely disruptive innovation that has attracted considerable attention from practitioners and academics. AI provides extensive, and unprecedented, opportunities for fundamental changes and extensive upgrades across many industries. This disruptive technology makes incredible things possible, such as autonomous vehicles, facial recognition payment, guidance robots, etc.

More specifically, AI injects fresh vitality into digital business, and it facilitates the development of smart services and promotes digital transformation. Currently, AI is considered as one of the top five emerging technologies when enterprises attempt to implement the digital first strategy. It is predicted that 70% of organizations will develop AI architectures owing to the increasing maturity and availability of AI technology (Goasduff, 2021). The age of AI is coming without doubt. Recently, AI has attracted many academics

attention in their corresponding fields. Many scholars have investigated the application of AI in different contexts such as information systems (e.g., Gursoy et al., 2019), tourism and hospitality (e.g., Li et al., 2019), marketing (e.g., Syam & Sharma, 2018), and financial management (e.g., Culkin & Das, 2017). Research results indicate that AI has the potential to change the way of interaction between organizations and customers, and brings higher business benefits (e.g., increasing efficiency, increasing effectiveness and decreasing cost).

Despite these great benefits for a business, Tarafdar et al. (2013) have warned of the dark sides of information technology. AI technology is no exception. It is admitted that AI has potential to induce potential risks at the individual level, organization level and social level (Alt, 2018). However, extensive attention is paid to positive aspects of AI, while the dark sides of AI are receiving relatively little attention, especially from the academic community. Considering the importance and universality of AI, the significant negative consequences brought by AI to individuals, organizations and society deserve further exploration. Given the limited research on the dark sides of AI, we thus provide this special issue and encourage researchers to explore the important issues in the AI realm especially in the electronic market context, such as electronic commerce, social media, emergent digital platforms, etc.

✉ Jian Mou
jian.mou@pusan.ac.kr

Xusen Cheng
xusen.cheng@ruc.edu.cn

Xiao Lin
xiaolin@nuaa.edu.cn

Xiao-Liang Shen
xlshen@whu.edu.cn

Alex Zarifis
zarifis.a@unic.ac.cy

- ¹ Renmin University of China, Beijing, China
- ² Nanjing University of Aeronautics and Astronautics, Nanjing, China
- ³ Wuhan University, Wuhan, China
- ⁴ University of Nicosia, Nicosia, Cyprus
- ⁵ School of Business, Pusan National University, 2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241, South Korea

AI-enabled applications

AI technology has been widely applied in various industries to promote business benefits. The application to each context can have some differences. In the e-Commerce contexts, online vendors adopt chatbots to provide their consumers 24/7 services (Luo et al., 2019). In the financial industry, facial recognition payment is becoming increasingly popular in online marketplaces (Liu et al., 2021). In brick-and-mortar retail, guidance robots are used to help consumers select products. In the medical health industry, AI-enabled

diagnostic systems are utilized in the medical imaging and histopathology context. AI shows great potential in medical diagnosis and treatment planning (Ploug & Holm, 2020). In the entertainment industry, the success of DeepMind's AlphaGo Zero has caused heated controversy (Chao et al., 2018). In the field of transportation, autonomous vehicles have attracted extensive attention from businesses (Chatopadhyay et al., 2020). In addition, other AI technologies such as speech recognition, language translation, image processing and deep learning are making various aspects of our lives more convenient. We also find AI-driven smart home products, such as XiaoIceJIN, Xiaodu, etc. Admittedly, AI does facilitate in problem solving and encourages technology innovation in various contexts. As mentioned above, AI contributes to high efficiency, effectiveness and low cost. However, AI is causing increasing controversy in practices where it is used. Firstly, AI is not performing as intelligently as some people expect. For instance, Facebook's Project M reported that about 70% of interactions between humans and AI failed (Griffith & Simonite, 2018). Furthermore, about 80% of consumers are unwilling to communicate with chatbots in e-commerce websites since chatbots cannot understand their needs (Forbes, 2019). Therefore, AI has also spurred dark side effects we will explore further in the next section.

The dark sides of AI

Regardless of the numerous potential benefits, there are undoubtedly many negative consequences of AI. We cannot merely look at the promising potential of AI (Floridi et al., 2018). Alt (2018) emphasizes that AI may lead to enormous risks at the individual level, organization level and society level. Importantly, these three aspects are considered as the most important elements for digitalization (Alt, 2018). Considering the background of our special issue, we mainly discuss the dark sides of AI in electronic markets.

From the individual's perspective, the detrimental effects of AI are mainly reflected in privacy concern, content recommendation and product recommendation in electronic markets. Firstly, AI has the ability to gain deep insights into consumers, which may exacerbate privacy concern (Grewal et al., 2021). For instance, Dickson (2019) argues that voice assistants (e.g., Alexa) can predict the moment the consumer's current relationship will end by analyzing the consumer's voice with AI technology. In addition, payments with facial recognition also introduce privacy risks since the human face carries a lot of personal information including the appearance, age, gender, etc. (Dibeklioglu et al., 2015; Dantcheva & Brémond, 2016; Liu et al., 2021). A recent study states that personalized recommendations lead to perceived privacy concern and perceived information

narrowing; as a result, people are reluctant to accept related technologies (Li et al., 2021).

From an organizational perspective, the introduction of AI-enabled products is likely to influence the reputation and final profit of companies. For instance, the performance of AI-enabled chatbots has an influence on consumers' satisfaction (Ashfaq et al., 2020; Eren, 2021). If chatbots do not perform as well as consumers expect, consumers will distrust them. As a result, they will no longer trust the sellers, or companies that use them (Yen & Chiang, 2020). In addition, since AI is a relatively emergent technology, organizations also face the great challenge of successfully implementing AI strategies and often fail to address the issue of how this impacts the human workforce.

From a societal perspective, AI also yields dark effects that vary from issues of data security to workforce replacement and ethical problems (Boyd & Wilson, 2017; Wang & Siau, 2018). The use of AI has raised widespread concerns for the human workforce (Danaher, 2019). In fact, the belief that AI may render us unemployed does not only exist within e-commerce, but it exists in all walks of life. Other problems include AI law and regulation, and AI ethics like moral dilemmas, AI discrimination and AI fairness (Wirtz et al., 2020). These potential problems thus pose challenges for AI governance at the society level (Wirtz et al., 2020). In addition, AI also play an important role in the COVID-19 influenced work and life environment, and the issues raised by using AI also changes risk perceptions and coping behavior.

Articles of the present issue

The present special issue on the dark side of AI is the first in Electronic Markets on this topic and includes six research papers. These papers shift people's attention from the bright sides of AI to the challenges of AI adoption in various domains. This enriches our understanding of the application of AI technology from different perspectives and helps us avoid the cycle of unrealistic expectations followed by disappointment. The following paragraphs discuss the articles in this special issue.

The first paper is authored by Sun et al. (2021b) with the title "The dark sides of AI personal assistants: effects of service failure on user continuance intention". In the last few years, smart devices are expanding their capabilities extensively. Smart devices can help companies provide a better service to customers. They can also help individuals to automatically perform some tasks such as checking information on the weather. In this paper, the authors use the context of a smart voice device, namely AI personal assistants (AIPAs). The authors argue that studies have scarcely uncovered the underlying mechanism

through which those dark sides of AIPAs impact users' intention to continue. From the perspective of technostress, this study proposes a theoretical model for consumers to cope with service failure pressure sources. Some interesting results emerge from this study. Furthermore, new avenues for research are opened to explore the mechanism of how the service failures of these AIPAs affect consumers' continuance intention through the perspective of technostress.

The second paper is authored by Ma et al. (2021) entitled "Understanding users' negative responses to recommendation algorithms in short-video platforms: a perspective based on the Stressor-Strain-Outcome (SSO) framework". Because of their ability to customize and enhance social media and e-commerce, recommendation systems have been widely employed. They can promote products and services to the consumers with a better fit to users' needs. This paper focuses on the role of AI-based recommendation algorithms. Although AI-based recommendation systems can bring benefits to both the users and the platform, not much is known regarding the dark side, especially users' negative responses. From the perspective of recommendation features and information characteristics, this study aims to uncover users' negative responses to AI-based recommendation algorithms in the algorithm-driven context of short-video platforms. Drawing on the stressor-strain-outcome (SSO) framework, this study identifies information-related stressors and examines their influence on users' negative responses to a recommendation algorithm. This paper also offers some useful insights on studying AI-based recommendation algorithms and expands our knowledge of the dark side of recommendation algorithms.

The third paper is authored by Kai and Zhang (2021) and entitled "Categorization and eccentricity of the AI risks: A comparative study of the global AI guidelines". This study develops a four-dimensional matrix based on the classic theoretical framework of "probability severity" in the field of risk management as a benchmark to analyze the possible risk of AI. Using the framework, a comparative study of the extant guidelines is conducted by coding the 123 guidelines with 1023 articles. The structure consists of four pairs of risks: specific-general, legal-ethical, individual-collective and generational-transgenerational. Results show that the extant guidelines are eccentric, while collective risk and generational risk are largely underestimated by stakeholders. Based on this analysis, gaps and conflicts are outlined for future research.

The fourth paper is authored by Mirbabaie et al. (2021) entitled "The rise of artificial intelligence – understanding the AI identity threat at the workplace". This paper addresses the problem of AI resistance by examining

the antecedents of AI identity threat. Applying a combination of quantitative and qualitative approaches, this study reveals three central predictors for AI identity threat in the workplace: changes to work, loss of status position, and AI identity. The work presented in this paper contributes to the theme of the special issue by clarifying how professional identity and AI identity relate to the negative impacts of AI (i.e., identity threat caused by AI) differently, and further provides useful information for practitioners who wish to introduce AI in the workplace.

The fifth paper is authored by Sun et al. (2021a) entitled "Prick the filter bubble: A novel cross domain recommendation model with adaptive diversity regularization". Recommender systems harness the data we provide with our behaviour to focus sharply on what we want, discarding other alternatives (Deldjoo et al., 2020; Milano et al., 2020). Do we lose something in our personal and professional lives from these narrow-targeted results? The paper attempts to mitigate this. A model is developed that balances accuracy of the recommendation with diversity. It is important for people to have a breadth of information and stimulation, so that they do not lose their own ability to judge the results they receive, and make the final decision.

The sixth paper is authored by Hornung and Smolnik (2021) and entitled "AI invading the workplace: negative emotions towards the organizational use of personal virtual assistants". We increasingly interact with AI through Personal Virtual Assistants (PVA), also referred to as chatbots. These can be convenient, effortless and effective but there are also challenges (Cheng et al., 2021; Zarifis et al., 2021). This paper focuses on the negative emotions towards them. The qualitative methodology of this research enables it to identify the most influential negative emotions as being related to having a loss and fear. These findings can help the design and implementation of PVAs that will have less resistance.

Acknowledgements We would like to thank National Natural Science Foundation of China (Grant NO. 72061147005) and fund for building world-class universities (disciplines) of Renmin University of China (Project No.KYGJA2021004) for providing funding for part of this research.

References

- Alt, R. (2018). Electronic markets and current general research. *Electronic Markets*, 28(2), 123–128. <https://doi.org/10.1007/s12525-018-0299-0>
- Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance

- intention of AI-powered service agents. *Telematics and Informatics*, 54, 101473. <https://doi.org/10.1016/j.tele.2020.101473>
- Boyd, M., & Wilson, N. (2017). Rapid developments in artificial intelligence: How might the New Zealand government respond? *Policy Quarterly*, 13(4), 36–44. <https://doi.org/10.26686/pq.v13i4.4619>
- Chao, Kou, Li, & Peng. (2018). Jie Ke versus AlphaGo: A ranking approach using decision making method for large-scale data with incomplete information. *European Journal of Operational Research*, 265(1), 239–247. <https://doi.org/10.1016/j.ejor.2017.07.030>
- Chattopadhyay, A., Lam, K. Y., & Tavva, Y. (2020). Autonomous Vehicle: Security by Design. *IEEE Transactions on Intelligent Transportation Systems*, 1–15. <https://doi.org/10.1109/TITS.2020.3000797>
- Cheng, X., Bao, Y., Zarifis, A., Gong, W., & Mou, J. (2021). Exploring consumers' response to text-based chatbots in e-commerce: The moderating role of task complexity and chatbot disclosure. *Internet Research*, (21). <https://doi.org/10.1108/INTR-08-2020-0460>
- Culkin, R., & Das, S. R. (2017). Machine learning in finance: The case of deep learning for option pricing. *Journal of Investment Management*, 15(4), 92–100.
- Danaher, J. (2019). The rise of the robots and the crisis of moral patiency. *AI & Society*, 34(1), 129–136. <https://doi.org/10.1007/s00146-017-0773-9>
- Dantcheva, A., & Brémond, F. (2016). Gender estimation based on smile-dynamics. *IEEE Transactions on Information Forensics and Security*, 12(3), 719–729. <https://doi.org/10.1109/TIFS.2016.2632070>
- Deldjoo, Y., Schedl, M., Cremonesi, P., & Pasi, G. (2020). Recommender systems leveraging multimedia content. *ACM Computing Surveys*, 53(5). <https://doi.org/10.1145/3407190>
- Dibeklioğlu, H., Alnajjar, F., Salah, A. A., & Gevers, T. (2015). Combining facial dynamics with appearance for age estimation. *IEEE Transactions on Image Processing*, 24(6), 1928–1943. <https://doi.org/10.1109/TIP.2015.2412377>
- Dickson, E. J. (2019). *Can Alexa and Facebook predict the end of your relationship?*. Retrieved October 1, 2021 from, <https://www.vox.com/the-goods/2019/1/2/18159111/amazon-fa>
- Eren, B. A. (2021). Determinants of customer satisfaction in chatbot use: Evidence from a banking application in Turkey. *International Journal of Bank Marketing*, 39(2), 294–311. <https://doi.org/10.1108/IJBM-02-2020-0056>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). An ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Forbes (2019). *AI stats news: chatbots increase sales by 67% but 87% of consumers prefer humans*. Retrieved October 1, 2021 from <https://www.forbes.com/sites/gilpress/2019/11/25/ai-stats-news-chatbots-increase-sales-by-67-but-87-of-consumers-prefer-humans/?sh=19efe3cf48a3>
- Goasduff, L. (2021). *While advances in machine learning, computer vision, chatbots and edge artificial intelligence (AI) drive adoption, it's these trends that dominate this year's Hype Cycle*. Retrieved October 8, 2021, from <https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021>
- Grewal, D., Guha, A., Saturnino, C. B., & Schweiger, E. B. (2021). Artificial intelligence: The light and the darkness. *Journal of Business Research*, 136, 229–236. <https://doi.org/10.1016/j.jbusres.2021.07.043>
- Griffith, E., & Simonite, T. (2018). *Facebook's virtual assistant M is dead*. Retrieved January 10, 2019 from, <https://www.wired.com/story/facebook-virtual-assistant-m-is-dead-so-are-chatbots/>
- Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, 49, 157–169. <https://doi.org/10.1016/j.ijinfomgt.2019.03.008>
- Hornung, O., & Smolnik, S. (2021). AI invading the workplace: Negative emotions towards the organizational use of personal virtual assistants. *Electronic Markets*. <https://doi.org/10.1007/s12525-021-00493-0>
- Kai, J., & Zhang, N. (2021). Categorization and eccentricity of AI risks: A comparative study of the global AI guidelines. *Electronic Markets*. <https://doi.org/10.1007/s12525-021-00480-5>
- Li, J., Bonn, M. A., & Ye, B. H. (2019). Hotel employee's artificial intelligence and robotics awareness and its impact on turnover intention: The moderating roles of perceived organizational support and competitive psychological climate. *Tourism Management*, 73, 172–181. <https://doi.org/10.1016/j.tourman.2019.02.006>
- Li J., Zhao H., Hussain S., Ming J., & Wu J. (2021). The Dark Side of Personalization Recommendation in Short-Form Video Applications: An Integrated Model from Information Perspective. In: Toepe K., Yan H., Chu S.K.W. (eds) Diversity, Divergence, Dialogue. iConference 2021. *Lecture Notes in Computer Science*, vol 12646. Springer. https://doi.org/10.1007/978-3-030-71305-8_8
- Liu, Y., Yan, W., & Hu, B. (2021). Resistance to facial recognition payment in China: The influence of privacy-related factors. *Telecommunications Policy*, 45(5), 1021155. <https://doi.org/10.1016/j.telpol.2021.102155>
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6), 937–947. <https://doi.org/10.1287/mksc.2019.1192>
- Ma, X., Sun, Y., Guo, X., Lai, K., & Vogel, D. (2021). Understanding users' negative responses to recommendation algorithms in short-video platforms: A perspective based on the stressor-strain-outcome (SSO) framework. *Electronic Markets*, 2021. <https://doi.org/10.1007/s12525-021-00488-x>
- Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender systems and their ethical challenges. *AI and Society*, 35(4), 957–967. <https://doi.org/10.1007/s00146-020-00950-y>
- Mirbabaie, M., Brünker, F., Möllmann (Frick), N. R. J., & Stieglitz, S. (2021). The rise of artificial intelligence – Understanding the AI identity threat at the workplace. *Electronic Markets*, 2021. <https://doi.org/10.1007/s12525-021-00496-x>
- Ploug, T., & Holm, S. (2020). The four dimensions of contestable AI diagnostics – A patient-centric approach to explainable AI. *Artificial Intelligence in Medicine*, 107, 101901. <https://doi.org/10.1016/j.artmed.2020.101901>
- Sun, J., Song, J., Jiang, Y., Liu, Y., & Li, J. (2021a). Prick the filter bubble: A novel cross domain recommendation model with adaptive diversity regularization. *Electronic Markets*. <https://doi.org/10.1007/s12525-021-00492-1>
- Sun, Y., Li, S., & Yu, L. (2021b). The dark sides of AI personal assistant: Effects of service failure on user continuance intention. *Electronic Markets*. <https://doi.org/10.1007/s12525-021-00483-2>
- Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, 69, 135–146. <https://doi.org/10.1016/j.indmarman.2017.12.019>
- Tarafdar, M., Gupta, A., & Turel, O. (2013). The dark side of information technology use. *Information Systems Journal*, 23(3), 269–275. <https://doi.org/10.1111/isj.12015>
- Wang, W., & Siau, K. (2018). Artificial intelligence: A study on governance, policies, and regulations. *Proceedings of the Midwest Association for Information (MWAIS 2018)*. <https://aisel.aisnet.org/mwaais2018/40>

- Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43(9), 818–829. <https://doi.org/10.1080/01900692.2020.1749851>
- Yen, C., & Chiang, M. C. (2020). Trust me, if you can: A study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments. *Behaviour & Information Technology*, <https://doi.org/10.1080/0144929X.2020.1743362>.
- Zarifis, A., Kawalek, P., & Azadegan, A. (2021). Evaluating if trust and personal information privacy concerns are barriers to using health insurance that explicitly utilizes AI. *Journal of Internet Commerce*, 20(1), 66–83. <https://doi.org/10.1080/15332861.2020.1832817>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.