INTERNATIONAL JOURNAL OF COMPUTER AUDITING

Editorial

Responsible Use of Artificial Intelligence

Tawei (David) Wang

School of Accountancy & MIS, College of Business, DePaul University, United States

Reference Format: Wang, T., 2022, Responsible Use of Artificial Intelligence, International Journal of Computer Auditing, Vol.4, No.1, pp.1-3.
https://doi.org/10.53106/256299802022120401001

Artificial intelligence (AI) has once again attracted the public's attention in the past decade. Riding on the wave of big data revolution, the development of AI is much more promising than that 50 years ago. One example is ChatGPT (https://openai.com/blog/chatgpt/), which has quickly become a hot topic in the past several months and has brought back all kinds of discussion about AI and decision making. In this editorial, I would like to highlight several perspectives that may help us rethink about the implications of using AI for decision making especially for audit professionals.

Accuracy and Transparency

Some forms of machine learning are now everywhere in our daily lives, from YouTube and Facebook recommendations to auto-driving cars. It does bring a lot of benefits, but it is not uncommon to see reported incidents related to machine driven decisions. Some of them can be a matter of life and death, such as Uber's self-driving car incident (Davies, 2016) and Tesla's auto pilot accident (NPR, 2022). Some other incidents were more hilarious. For instance, Oremus (2017) reported that, in early 2017, if someone asked Google Home, "Is Obama planning a coup?" and Google Home will say he's "in bed with the communist Chinese." All these examples highlight the accuracy concern and the lack of transparency of the AI algorithm. For instance, Lazer et al. (2014), using how Google trends can be used to predict flu, suggest the aspects readers should consider when interacting with machine learning, which includes discussion of accuracy, transparency, and algorithm. This kind of discussion has already triggered the regulators' attention. One recent development in the European Union, the proposed AI Act, is an example of the regulator's responses to the increased use and focus on AI (European Union, 2022). The AI Act not only emphasizes on the notification requirement but also brings in risk management procedures regarding data, documentation, transparency, human oversight, and accuracy as well as security issues.

Risk and Responsibility

The proposed AI Act categorizes AI applications based on the risk level from unacceptable risk, high risk, to minimal risk, which suggest another major concern of AI: risks and responsibility. The risks imposed can be from a micro or individual level, such as the use of chatbots, to bigger risks, such as the impact on education, the change in employment decisions, or the change in infrastructure. For instance, the systematic issue from the history of health disparities is reflected in the dataset we use for model training, so the predictive model used by health care providers also faces the same systematic issue (Gandhi, 2020). The risk also comes from the combination of different algorithms and the learning of the machine. When the rationale behind the decision becomes so complicated that no human beings can understand, it imposes another level of risks (Smith, 2018). Though there are discussion about whether AI can be trained to explain itself (Kuang, 2017), the concept of responsible AI has also emerged. Specifically, researchers have started to focus more on the risks, privacy/security issues, and fairness of AI from the machine learning life cycle with the hope to mitigate the negative impacts (Gandhi, 2020). From the legal or ethical perspectives, the discussion of who should be responsible for AI decisions has never stopped in the past few years and it remains a datable topic (Henz, 2021).

Biases and Manipulation

One of the risks that can come from the whole machine learning life cycle is biases (Suresh and Guttag, 2021). For instance, the digital camera may believe a Taiwanese-American blogger is blinking (Rose, 2010) simply because of how the camera recognizes facial characteristics. It becomes even more concerning when the public believes that correlations may somehow reflect causation and the data patterns demonstrate an objective truth (Crawford, 2013). For example, the more privileged group of Internet users may accidentally change what we observe simply because of their postings on social media is “louder” compared to other groups of users. The bias may also come from how the “reward” system in the algorithm is determined. For example, Facebook’s newsfeed needs to be clicky and upbeat (Tufekci, 2016), which may change how the newsfeed is selected. Though it may not be purposely determined, it demonstrates the possibility of a manipulated reality or data pattern. Another example is that ChatGPT has human labelers to rank its outputs from best to worst (Wang, 2023). It definitely makes it more human-like, but what are the implications of this level of human intervention?

The development of AI is not reversible. It is also exciting to see how AI may change every aspect of our lives and profession. Though different stakeholders in this ecosystem have been working hard to address concerns raised by the public, as an audit professional, it is inevitable for us to consider potential implications before relying on AI for decision making.

References

1. Crawford, K. (2013). *The hidden biases in big data*. Harvard Business Review. https://hbr.org/2013/04/the-hidden-biases-in-big-data
2. Davis. (2016). *As Uber launches self-driving in SF, regulators shut it down*. Wired. https://www.wired.com/2016/12/ubers-self-driving-car-ran-red-light-san-francisco/
3. European Union. (2022). *The AI Act*. https://artificialintelligenceact.eu/the-act/.
4. Henz, P. (2021). Ethical and legal responsibility for artificial intelligence. *Discover Artificial Intelligence*, *1*(1). https://doi.org/10.1007/s44163-021-00002-4
5. Gandhi, T. (2022). *Introducing the responsible AI in practice series*. https://blog.dataiku.com/introducing-the-responsible-ai-in-practice-series-and-healthcare-use-case
6. Kuang, C. (2017). *Can AI be taught to explain itself?* The New York Times. https://www.nytimes.com/2017/11/21/magazine/can-ai-be-taught-to-explain-itself.html

7. Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of google flu: Traps in big data analysis. *Science*, *343*(6176), 1203–1205.
https://doi.org/10.1126/science.1248506

8. Press, T. A. (2022). *A Tesla driver is charged in a crash involving autopilot that killed 2 people.* NPR.
https://www.npr.org/2022/01/18/1073857310/tesla-autopilot-crash-charges.

9. Oremus, W. (2017). *Google's fake facts and Uber's self-driving slip-ups are symptoms of a deeper problem.* Slate Magazine.
https://slate.com/technology/2017/03/googles-featured-snippets-and-ubers-self-driving-slip-ups-point-to-a-problem-for-the-future-of-a-i.html

10. Rose, A. (2010). *Are face-detection cameras racist?* Time.
https://content.time.com/time/business/article/0,8599,1954643,00.html

11. Smith, A. (2018). *Franken-algorithms: The deadly consequences of unpredictable code.* The Guardian.
https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger

12. Suresh, H., & Guttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. *Equity and Access in Algorithms, Mechanisms, and Optimization.* 1-9.
https://doi.org/10.1145/3465416.3483305

13. Tufekci, Z. (2016). *The real bias built in at facebook.* The New York Times.
https://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html

14. Wang, B. (2023). *CHATGPT has a human team train it to be a lot better.* NextBigFuture.com.
https://www.nextbigfuture.com/2023/02/chatgpt-has-a-human-team-train-it-to-be-a-lot-better.html.