INTEGRATING ETHICAL AI IN PUBLIC ADMINISTRATION: A PATHWAY TO ACHIEVING SUSTAINABLE DEVELOPMENT GOALS

Perga Iurii

Igor Sikorsky Kyiv Polytechnic Institute
yperga@gmail.com

Summary

This research explores the transformative potential of ethical artificial intelligence (AI) in public administration to improve services, increase transparency, and promote social equity while aligning with Sustainable Development Goals (SDGs). Key ethical principles such as transparency, accountability, fairness, inclusivity, and privacy are essential for trustworthy AI deployment in government. Ethical AI can support SDGs by reducing inequalities (SDG 10), enhancing justice and institutional trust (SDG 16), and fostering sustainable innovation (SDG 9). Challenges include lack of universal ethical standards, resource constraints, data privacy concerns, and biases embedded in historical data. To address these, the article recommends developing comprehensive ethics guidelines, promoting interagency collaboration, investing in AI education, engaging diverse stakeholders, and implementing accountability mechanisms. The integration of ethical AI in public administration offers a pathway toward inclusive, transparent governance and sustainable development, provided that ethical considerations guide its adoption and use.

Key words:

ethical AI, public administration, transparency, accountability, Sustainable Development Goals

The adoption of artificial intelligence in public administration offers transformative opportunities to improve public services, streamline operations, and address complex social issues (Mikhaylov, Esteve, & Campion, 2018). However, these benefits are accompanied by ethical concerns that challenge governments to align AI usage with principles of transparency, fairness, accountability, and human rights. Ethical AI in public administration requires adhering to several key principles. Transparency is crucial, as AI systems must be open to scrutiny; public institutions should disclose how AI algorithms function, their intended purpose, and potential biases (Floridi et al., 2018). Transparency allows citizens to understand AI-driven decisions, fostering trust in public institutions. Accountability is also

essential in public sector AI implementations, requiring mechanisms that ensure any harmful impact from AI decisions can be addressed through legal or regulatory channels (Mittelstadt et al., 2016). Clear accountability policies are necessary for protecting citizens' rights and maintaining institutional integrity. Moreover, fairness and inclusivity are central to ethical AI, as these systems can inadvertently amplify biases, leading to unfair outcomes. To prevent this, governments should ensure that AI models are designed to minimize discrimination, particularly when dealing with vulnerable populations (Binns, 2018). Privacy is another critical ethical consideration since AI systems often rely on large volumes of data. Governments must adhere to stringent data protection standards, ensuring that citizens' personal information is safeguarded against misuse (Crawford & Calo, 2016).

Ethical AI in public administration can directly impact several SDGs, supporting efforts to build inclusive, equitable, and sustainable societies. Ethical AI can help reduce inequalities (SDG 10) by ensuring fair access to public services. For instance, AI systems can be designed to eliminate biases in healthcare, education, and welfare, allowing underserved communities to receive equal treatment (O'Neil, 2016). Public institutions can leverage AI-driven insights to allocate resources more equitably, addressing social disparities. AI also supports SDG 16 (Peace, Justice, and Strong Institutions) by promoting transparency and accountability through data-driven decision-making. Using AI in law enforcement, for example, can enhance the accuracy of crime predictions without racial bias if designed ethically, thus contributing to institutional trust and justice (Richardson, Schultz, & Crawford, 2019). Additionally, SDG 9 (Industry, Innovation, and Infrastructure) benefits from ethical AI integration as it promotes sustainable innovation in public services. Ethical guidelines encourage responsible experimentation with AI while minimizing societal risks, thereby contributing to resilient infrastructure that benefits all citizens (Brynjolfsson & McAfee, 2014).

However, implementing ethical AI in public administration presents challenges. A significant issue is the lack of clear standards, as defining ethical standards for AI in public institutions is complex due to varied cultural, social, and legal perspectives (Floridi et al., 2018). Additionally, resource constraints are a barrier; ethical AI development requires significant financial and technical resources, which may be limited in some regions (Whittaker et al., 2018). Data and privacy concerns also arise as governments need extensive data to develop effective AI solutions, yet balancing data use with privacy concerns is challenging, particularly in areas with strict data protection laws (Crawford & Calo, 2016). Furthermore, AI algorithms are often trained on historical data, which may contain inherent biases, posing risks for fairness and equality in public services (O'Neil, 2016).

To effectively integrate ethical AI, several recommendations are proposed. First, governments should develop comprehensive AI ethics guidelines that outline ethical standards for AI use, emphasizing fairness, accountability, transparency, and data protection (Jobin et al., 2019). Promoting interagency collaboration is also essential, as effective AI governance requires collaboration across government sectors to share resources, expertise, and best practices, creating a unified approach to ethical AI (Mittelstadt et al., 2016). Investing in AI education and training for public sector employees is crucial, as it enables them to understand AI technologies and make informed decisions (Whittaker et al., 2018). Engaging with stakeholders, including civil society, academia, and the private sector, can help governments better understand public concerns regarding AI, ensuring that policies reflect diverse perspectives (Floridi et al., 2018). Finally, implementing accountability mechanisms, such as audits and oversight bodies, can ensure that AI systems adhere to ethical standards, with regular assessments identifying potential risks and ensuring transparency (Richardson et al., 2019).

In conclusion, integrating ethical AI into public administration has the potential to advance SDGs by fostering inclusivity, accountability, and transparency. By adopting ethical principles, public institutions can build citizen trust and promote sustainable development. Although achieving this vision requires overcoming challenges related to resources, standards, and bias mitigation, strategic planning, capacity-building, and stakeholder

engagement can enable public administrators to harness ethical AI responsibly. This paper emphasizes the need for ethical AI to support SDGs in public administration, advocating a balanced approach that combines technological innovation with societal responsibility.

References

1. Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency, 149–159.
2. Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: W. W. Norton & Company.
3. Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. Nature News, 538(7625), 311.
4. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689–707.
5. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399.
6. Mikhaylov, S. J., Esteve, M., & Campion, A. (2018). Artificial intelligence for the public sector: Opportunities and challenges of cross-sector collaboration. Philosophy & Technology, 31(4), 1–15.
7. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.
8. O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.
9. Perga, I. (2024). The role of AI (large language models) in managerial decision-making: Benefits and challenges. Наукові перспективи, 10(52), 32–49. https://doi.org/10.52058/2708-7530-2024-10(52)-32-49
10. Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. NYU Law Review Online, 94, 192–233.
11. Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., … & Schwartz, O. (2018). AI now report 2018. AI Now Institute at New York University.