ZESZYT PRAWNICZY UAM
NR 15/2025
https://doi.org/10.14746/zpuam.2025.15.8

ISSN (ONLINE) 2956-2767
ISSN (PRINT) 2720-2526

JAKUB SAWICKI

Adam Mickiewicz University, Poznań, Poland
https://orcid.org/0009-0000-6591-4979
jaksaw2@st.amu.edu.pl

The Role of Decision-Making Algorithms in Public Administration—Legal and Ethical Challenges

Abstract: The increasing reliance on decision-making algorithms in public administration raises significant legal and ethical challenges. This article examines the key issues associated with algorithmic governance, including transparency, accountability, and potential biases in automated decision-making processes. Using a legal-analytical method, I evaluate whether machine-learning algorithms can comply with existing legal principles while enhancing efficiency in governance. My findings suggest that while algorithms can improve decision-making speed and accuracy, their nature complicates compliance with legal transparency and due process requirements. I argue that algorithmic accountability mechanisms, including explainability frameworks and regulatory oversight, are essential in order to ensure fairness and legality in automated administrative decisions.

Keywords: algorithmic governance, public administration, transparency, accountability, artificial intelligence

Introduction

The integration of decision-making algorithms into public administration has significantly transformed governmental processes by enhancing operational efficiency, streamlining bureaucratic workflows, and improving the accuracy of predictive analytics.[1] These technologies enable authorities to rapidly analyse vast amounts of data, allowing for more informed policy-making, optimized resource allocation, and improved service delivery.[2] However, while these advancements present numerous benefits, their adoption also raises critical concerns regarding transparency, fairness, and legal accountability, particularly in the context of democratic governance.[3] One of the most pressing challenges is the "black-box" nature of machine-learning models, which often operate through complex, opaque processes that make it difficult for both citizens and oversight bodies to scrutinize administrative decisions.[4] This

[1] Cary Coglianese and David Lehr, "Transparency and Algorithmic Governance," *Administrative Law Review* 71, no. 1 (2019): 6–8.

[2] Michael Veale and Lilian Edwards, "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling," *Computer Law & Security Review* 34, no. 2 (2018): 398–99, https://doi.org/10.1016/j.clsr.2017.12.002.

[3] Joshua A. Kroll et al., "Accountable Algorithms," *University of Pennsylvania Law Review* 165, no. 3 (2017): 638–39.

[4] Tal Zarsky, "The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making," *Science, Technology, & Human Values* 41, no. 1 (2016): 121–22, https://doi.org/10.1177/0162243915605575.

lack of interpretability can undermine the principles of due process, as affected individuals may struggle to understand the rationale behind automated decisions that impact their rights, benefits, or obligations. Without adequate explainability, these systems risk eroding public trust in governmental institutions and exacerbating concerns about algorithmic bias and discrimination.[5] Furthermore, the deployment of algorithmic decision-making at scale within public institutions introduces broader implications for democratic legitimacy and social acceptance. The reliance on data-driven governance models necessitates critical examination of whether these systems align with constitutional principles, including equal protection under the law, non-discrimination, and procedural fairness. Issues such as algorithmic bias, the disparate impact on marginalized communities, and the reinforcement of existing social inequalities must be carefully addressed to ensure that these technologies do not perpetuate or exacerbate systemic injustices.[6] Additionally, concerns extend beyond legal and ethical considerations to include the fundamental question of human oversight in automated decision-making. The risk of "automation bias", where public officials defer excessively to algorithmic outputs without critically assessing their validity, further complicates the landscape of responsible AI implementation in public administration.[7] Effective regulatory interventions, such as independent audits, algorithmic impact assessments, and mechanisms for contestability and redress, must be established to ensure accountability and compliance with democratic values. Given these complexities, policymakers, legal scholars, and technologists must engage in an interdisciplinary dialogue to develop robust frameworks that balance efficiency gains with the protection of fundamental rights. Addressing these concerns requires a multi-faceted approach, including legislative measures, ethical AI guidelines, and participatory mechanisms that allow stakeholders, including the public, to be involved in the governance of algorithmic systems. Only through comprehensive regulation and governance can the full potential of algorithmic decision-making be harnessed in a manner that upholds principles of transparency, fairness, and democratic integrity.[8]

---

[5] Danielle Keats Citron and Frank Pasquale, "The Scored Society: Due Process for Automated Predictions," *Washington Law Review* 89, no. 1 (2014): 24–25.

[6] Karen Hao, "This Is How AI Bias Really Happens—and Why It's So Hard to Fix," MIT Technology Review, published 4 February 2019, https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/.

[7] James Vincent, "The Problem with AI Ethics," The Verge, published 3 April 2019, https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech.

[8] Sandra Wachter et al., "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation," *International Data Privacy Law* 7, no. 2 (2017): 76–99, https://doi.org/10.1093/idpl/ipx005.

1. Legal Principles Governing Algorithmic Decision-Making

1.1. Transparency and Explainability

Transparency in government decision-making is a fundamental principle enshrined in democratic governance, ensuring that public institutions operate with openness and accountability.[9] Transparent governance fosters trust between citizens and the state, reinforcing the legitimacy of administrative actions. It allows individuals to understand how decisions that affect their rights, obligations, and access to services are made. The legal concept of "fishbowl transparency" requires that government decisions be subject to public scrutiny, making all relevant information available to ensure accountability.[10] This means that the reasoning, data, and methodologies used in administrative decision-making must be accessible to the public, stakeholders, and oversight bodies. On the other hand, "reasoned transparency" necessitates providing explicit justifications for administrative actions, ensuring that decisions are not made arbitrarily or without proper explanation.[11] In essence, reasoned transparency demands that government agencies articulate the basis for their decisions, whether they result in policy implementation, law enforcement, or the allocation of public resources. However, the widespread adoption of machine-learning algorithms in public administration poses significant challenges to these transparency principles. Many of these algorithms function as "black boxes", meaning that even their developers may struggle to explain how specific decisions are reached. The complexity of artificial intelligence (AI) models, particularly deep learning systems, creates barriers to public understanding and legal compliance.[12] Without adequate transparency, there is a risk that automated decisions may lead to unfair, biased, or erroneous outcomes without affected individuals being able to challenge or understand them. To address these challenges, governments must develop robust explainability frameworks which ensure that algorithms used in public administration can be scrutinized and understood. One approach involves implementing model interpretability techniques, which enable the dissection of AI-driven decision-making processes. These techniques include rule-based explanations, feature importance analysis, and counterfactual reasoning, which help illustrate how different factors influence algorithmic outputs. Another approach is open auditing of algorithmic decisions, allowing external experts, policymakers, and civil society organizations to

[9] Coglianese and Lehr, "Transparency and Algorithmic Governance," 20–21.

[10] Gardner Susan Marie, "Accountability, Transparency, and Living in the Fishbowl," Municipal World, published March 2019, https://www.municipalworld.com/feature-story/life-in-fishbowl/.

[11] Lilian Edwards and Michael Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For," *Duke Law & Technology Review* 16, no. 1 (2017): 38–40.

[12] Zarsky, "The Trouble with Algorithmic Decisions," 126–28.

assess the fairness, accuracy, and reliability of AI-driven decision-making systems.[13] Explainability is particularly crucial in high-stakes applications, where algorithmic decisions have significant consequences for individuals and society. For example, in social welfare benefits allocation, automated systems determine eligibility and distribute resources to citizens in need. Lack of transparency in these systems can lead to unjustified denials or reductions of benefits, disproportionately affecting vulnerable populations. Similarly, in the criminal justice system, risk assessment algorithms are often used to predict recidivism rates or recommend sentencing outcomes. Without clear explanations of how these decisions are made, individuals may face wrongful classifications that impact their freedom and future opportunities. Ensuring transparency and explainability in government decision-making requires a multi-faceted approach. Governments must not only implement technical solutions to enhance AI interpretability but also establish legal frameworks mandating disclosure and justification of algorithmic decisions. Public officials should receive specialized training to critically assess algorithmic outputs and ensure that decisions are subject to human oversight. Furthermore, citizen engagement and participatory governance mechanisms should be promoted to allow the public to contribute to discussions on the ethical deployment of AI in administration. By prioritizing transparency and explainability, governments can maintain public trust, uphold democratic values, and mitigate the risks associated with opaque decision-making processes.[14] A clear regulatory framework combined with technological advancements in AI interpretability can help ensure that decision-making algorithms serve the public interest while safeguarding individual rights and freedoms.

1.2. Due Process and Accountability

Administrative law mandates that individuals subject to governmental decisions must receive an explanation and an opportunity to contest the outcome. This principle is deeply rooted in democratic governance and ensures that state authorities remain accountable for their actions. It serves as a safeguard against arbitrary decision-making, reinforcing the rule of law and protecting citizens from unjustified administrative actions.[15] As public institutions increasingly rely on automated decision-making systems, the challenge lies in ensuring that these fundamental principles remain upheld, even in cases involving AI and machine-learning algorithms. The European General Data Protection Regulation (GDPR) explicitly addresses the risks associated with automated decision-making, particularly in cases where such

[13] Wachter et al., "Why a Right to Explanation of Automated Decision-Making Does Not Exist," 76–99.

[14] Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review* 104, no. 3 (2016): 729–31, http://dx.doi.org/10.15779/Z38BG31.

[15] Citron and Pasquale, "The Scored Society," 13–14.

decisions may have significant legal or personal consequences.[16] Article 22 of the GDPR limits fully automated decision-making processes that result in legal effects or similarly significant impacts. However, the applicability of Article 22(3)—which mandates safeguards such as human intervention, the right to express views, and the right to contest decisions—is limited to cases where the automated decision is either necessary for the performance of a contract or based on the explicit consent of the data subject (Article 22(2)(a) and (c) GDPR). These conditions are rarely met in public administration contexts, where decisions are typically based on legal mandates (Article 22(2)(b)). In such cases, the regulation merely requires "appropriate safeguards", without detailing specific mechanisms.[17]

Moreover, Article 22 must be interpreted in conjunction with other provisions of the GDPR, including Article 13(2)(f), Article 14(2)(g), and Article 15(1)(h), which establish transparency obligations concerning the logic, significance, and consequences of automated decisions.[18] These complementary provisions form an integrated framework for ensuring data subject awareness and agency. Beyond the GDPR, the recently adopted Artificial Intelligence Act (AI Act, August 2024) provides additional requirements, particularly Article 86, which imposes duties of transparency and explainability for high-risk AI systems.[19] This includes the obligation to provide understandable information about how AI systems function and how decisions are made. Notably, Article 86(3) stresses that these obligations must be read in light of existing Union and Member State law, creating a layered framework of interrelated norms.

Therefore, a comprehensive legal assessment must go beyond Article 22 GDPR and include the AI Act as well as relevant national provisions, such as Germany's *Verwaltungsverfahrensgesetz* (Administrative Procedure Act), which already incorporates safeguards for automated decision-making processes.[20] The failure to acknowledge this broader legal landscape significantly undermines the robustness of legal analysis. While the article correctly calls for "comprehensive legal frameworks that go beyond the basic protections outlined in GDPR," it overlooks existing

[16] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (OJ L 119, 4.5.2016, pp. 1–88).

[17] Andrew D. Selbst and Julia Powles, "Meaningful Information and the Right to Explanation," *International Data Privacy Law* 7, no. 4 (2017): 233–42, https://doi.org/10.1093/idpl/ipx022.

[18] Wachter et al., "Why a Right to Explanation of Automated Decision-Making Does Not Exist," 76–99.

[19] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (OJ L, 2024/1689, 12.7.2024).

[20] Karen Yeung, "Algorithmic Regulation: A Critical Interrogation," *Regulation & Governance* 12, no. 4 (2018): 520–23, https://doi.org/10.1111/rego.12158.

instruments that already begin to address these challenges. A complete evaluation must critically assess whether the current legal regime—including the GDPR, AI Act, and national administrative law—is sufficient or requires enhancement.

Another critical issue relates to the lack of human oversight in algorithmic decision-making. While GDPR requires meaningful human intervention in high-stakes automated decisions, the practical implementation of this safeguard remains inconsistent across jurisdictions. In many cases, human oversight is either superficial or non-existent, leading to what has been described as automation bias—a phenomenon where human reviewers over-rely on algorithmic outputs without critically assessing their validity.[21] This is particularly problematic in areas such as social welfare allocation, immigration decisions, predictive policing, and risk assessment in the criminal justice system, where algorithmic errors can result in wrongful denials of rights, unjustified penalties, or disproportionate targeting of marginalized communities. To address these challenges, there is an urgent need for comprehensive legal frameworks that go beyond the basic protections outlined in GDPR.

Governments should adopt strict transparency requirements mandating that public institutions disclose the logic, criteria, and datasets underlying algorithmic decisions. Additionally, independent oversight mechanisms—such as external audits, algorithmic impact assessments, and specialized regulatory agencies—should be established to monitor the fairness, accuracy, and legality of AI-driven decision-making processes. Such oversight must also include mechanisms for redress, allowing individuals to effectively appeal and rectify erroneous or unjust algorithmic decisions. Furthermore, ensuring due process in algorithmic governance requires greater public participation in the design and deployment of AI-driven systems. Civil society organizations, legal scholars, and technical experts should be actively involved in assessing the potential risks of these technologies before they are implemented in critical areas of public administration. Democratic accountability also demands that policymakers engage in open discussions with affected communities to better understand their concerns and ensure that AI systems are aligned with human rights principles.

2. Ethical Challenges in Algorithmic Public Administration

2.1. Bias and Discrimination

Algorithmic decision-making, while often perceived as objective and data-driven, can inadvertently reinforce and even exacerbate biases present in historical data. These biases may stem from systemic inequalities embedded in past decision-making processes, social prejudices, or institutional discrimination that have historically disad-

[21] Yeung, "Algorithmic Regulation," 505–23.

vantaged certain demographic groups.[22] When machine-learning models are trained on biased datasets, they tend to learn and replicate those biases, leading to algorithmic outcomes that disproportionately affect marginalized communities.[23] One of the most prominent examples of this phenomenon is the use of algorithmic risk assessments in the criminal justice system. Predictive models are frequently employed to assess the likelihood of recidivism, determine bail eligibility, or guide sentencing decisions.[24] However, studies have shown that these algorithms can systematically classify individuals from minority backgrounds as higher-risk compared to their white counterparts, even when controlling for similar offence histories.[25] Such discriminatory outcomes undermine the principles of fairness and equal treatment under the law, leading to concerns about due process violations and systemic racial bias. This issue is further exacerbated when training data reflect historical prejudices, as they encode patterns of past discrimination into automated decision-making. For example, if a hiring algorithm is trained on historical employment data from industries with a long history of gender discrimination, the algorithm may systematically disadvantage female applicants by replicating past hiring biases.[26] Similarly, predictive policing algorithms trained on crime data from neighbourhoods that have been over-policed in the past may perpetuate a cycle of excessive law enforcement targeting those communities, reinforcing existing inequalities. Beyond technical solutions, policy interventions and legal frameworks play a crucial role in mitigating algorithmic bias. Governments and regulatory bodies should enforce transparency and accountability measures, requiring institutions that use algorithmic decision-making to conduct regular bias audits and report on fairness metrics. Additionally, public participation and oversight mechanisms should be implemented to allow affected communities to voice concerns and provide feedback on the fairness of automated decision-making processes.[27]

2.2. The Risk of "Automation Bias"

As public administration increasingly integrates artificial intelligence and machine-learning models into decision-making processes, concerns regarding au-

[22] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown, 2016), 151–53.

[23] Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press, 2018), 80–82.

[24] Rashida Richardson et al., "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice," *New York University Law Review Online* 94, 2019: 193–95.

[25] Eubanks, *Automating Inequality*, 90.

[26] Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of Machine Learning Research* 81, 2018: 84–85.

[27] United Nations Human Rights Office, "B-Tech Project Report: Human Rights Risks in AI-Based Decision Making," 2021, https://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf, accessed 27 February 2025.

tomation bias have become more pronounced.[28] Automation bias refers to the cognitive tendency of human decision-makers, including public officials, to place excessive trust in algorithmic recommendations, often without critically assessing their validity, accuracy, or potential biases. This overreliance can lead to the unquestioned acceptance or rubber-stamping of flawed algorithmic outputs, reducing the role of meaningful human oversight in governance.[29] One of the major risks of automation bias is that it can create a false perception of objectivity in decision-making. Algorithms, though often perceived as neutral and data-driven, may still inherit biases from historical data or flawed training methodologies.[30] When government officials assume that algorithmic outputs are inherently more accurate than human judgments, they may fail to recognise instances where the system produces errors or reinforces structural inequalities. This is particularly concerning in high-stakes scenarios, such as social welfare allocation, law enforcement and criminal justice, and public sector hiring.[31] The consequences of unchecked automation bias extend beyond individual cases of injustice. Over time, its systemic effects can erode public trust in government institutions, particularly when citizens perceive that algorithmic decisions lack transparency, fairness, or accountability.[32] If people believe that AI-driven decisions cannot be challenged or that human officials merely serve as passive enforcers of algorithmic outputs, faith in democratic governance may weaken. To address these challenges, governments must implement comprehensive training programs for public officials, ensuring that they develop the skills necessary to critically evaluate and question algorithmic recommendations.[33] Additionally, government agencies should establish accountability mechanisms, such as mandatory human review of AI-driven decisions, independent audits of algorithmic systems, and accessible appeal processes for citizens affected by automated outcomes. By proactively addressing automation bias, public administrations can harness the benefits of AI while ensuring that algorithmic governance remains aligned with the principles of fairness, transparency, and human-centred decision-making.[34]

---

[28] O'Neil, *Weapons of Math Destruction*, 100.

[29] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015), 216–17.

[30] Yeung, "Algorithmic Regulation," 520–23.

[31] Richardson et al., "Dirty Data, Bad Predictions," 201–05.

[32] Julia Angwin et al., "Machine Bias," ProPublica, published 23 May 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[33] Brent Mittelstadt et al., "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society* 3, no. 2 (2016): 8–10, https://doi.org/10.1177/2053951716679679.

[34] Ann Florini, *The Right to Know: Transparency for an Open World* (Columbia University Press, 2007), 210–12.

3. Proposals for Regulatory and Policy Reforms

As public administration increasingly adopts algorithmic decision-making systems, ensuring algorithmic transparency becomes a fundamental priority for regulatory bodies and policymakers. Transparency in this context means that government agencies must not only disclose the existence of AI-driven decision-making tools but also provide meaningful insights into their underlying logic, data sources, and performance metrics.[35] Without such transparency, citizens and oversight bodies may struggle to assess the fairness, accountability, and reliability of automated administrative decisions. A multilayered transparency approach should be implemented to address different aspects of AI-driven decision-making.

Government institutions should be required to publish clear explanations of how algorithms generate decisions, including the key factors influencing the outcomes. This would help prevent situations where individuals are adversely affected by opaque automated rulings without any recourse to understanding why.[36] The adoption of explainable AI (XAI) techniques, such as interpretable machine learning models, counterfactual explanations, and rule-based decision-making processes, should be prioritized. These methods help bridge the gap between algorithmic efficiency and the need for accountability by making automated decisions more understandable to both policymakers and affected individuals.[37]

Public institutions using AI should, to the extent legally permissible, disclose relevant characteristics of the datasets that train their models. Full disclosure may be constrained by data protection regulations, such as the GDPR, or by confidentiality obligations and trade secrets.[38] The AI Act also introduces provisions restricting the use and disclosure of personal data in high-risk systems, particularly when data subjects cannot be sufficiently anonymized.[39] Therefore, transparency obligations must be balanced against privacy and intellectual property considerations, requiring nuanced implementation strategies. Public scrutiny of training data can help ensure fairness and reduce discriminatory outcomes.[40] To build public trust in AI-driven governance, regulatory frameworks should require institutions to document and publicly disclose validation metrics used to assess algorithmic fairness and reliability. These metrics should include assessments of bias detection, predictive accuracy across different population groups, and evaluations of unintended discriminatory effects.[41] Algorithmic transparency should not be limited to retrospective analysis;

[35] Pasquale, *The Black Box Society*, 215–16.

[36] Selbst and Powles, "Meaningful Information and the Right to Explanation," 233–42.

[37] Yeung, "Algorithmic Regulation," 505–23.

[38] Pasquale, *The Black Box Society*, 219–20.

[39] AI Act.

[40] Reuben Binns, , "Fairness in Machine Learning: Lessons from Political Philosophy," *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency* 81, 2018: 157–58.

[41] Eubanks, *Automating Inequality*, 220–21.

there should also be mechanisms for ongoing audits and real-time monitoring of automated decision-making processes. Additionally, government agencies must implement clear appeal procedures that allow individuals to challenge decisions made by algorithms, thereby ensuring that automated rulings are not final and unreviewable.[42]

Regulatory bodies should enact legislation mandating transparency standards for AI deployment in public administration. Such laws should specify the types of information that must be disclosed, establish mandatory oversight mechanisms, and outline penalties for non-compliance. Moreover, governments should consider establishing independent AI ethics commissions tasked with evaluating the societal impacts of algorithmic decision-making and ensuring compliance with transparency requirements.[43] By institutionalizing these transparency measures, governments can foster greater public trust in AI-driven decision-making, uphold democratic principles, and mitigate risks associated with opaque, unaccountable, and potentially biased automated systems.

It is important to note that since the drafting of this article, the European Union has adopted the Artificial Intelligence Act (August 2024), which addresses several of the concerns raised herein. In particular, the AI Act codifies obligations regarding transparency, documentation, and oversight for high-risk AI systems used in public administration.[44] These include requirements for explainability, data governance, and human oversight, thereby partially responding to demands for regulatory clarity and accountability mechanisms. Nonetheless, gaps remain, particularly in relation to participatory governance, appeal rights, and the detailed procedural safeguards required for ensuring due process in administrative contexts.[45] Therefore, while the AI Act represents a significant step forward, it does not fully obviate the need for continued legal development and empirical evaluation of its implementation.

4. Independent Oversight and Audits

As governments increasingly integrate AI and machine learning algorithms into public administration, the need for independent oversight and auditing mechanisms has never been more critical. Without robust evaluation frameworks, algorithmic decision-making systems may operate with unchecked bias, opacity, and a lack of accountability, leading to potential injustices and public distrust in automated governance.[46] To address these concerns, comprehensive external audits and oversight

[42] United Nations Human Rights Office, "B-Tech Project Report," 10.
[43] Binns, "Fairness in Machine Learning," 159.
[44] AI Act.
[45] Mittelstadt et al., "The Ethics of Algorithms," 7–8.
[46] Yeung, "Algorithmic Regulation," 505–23.

mechanisms must be established to ensure that AI-driven decision-making adheres to legal, ethical, and democratic principles. A multi-layered independent oversight approach should cover several areas.

Regulatory agencies should create independent institutions specifically tasked with auditing government-deployed AI systems. These institutions should have the authority to assess the fairness, accuracy, and transparency of automated decision-making models, ensuring they do not discriminate against vulnerable populations or reinforce systemic biases.[47] Judicial oversight should be integrated into the deployment of AI in public administration. Courts—particularly administrative courts—must be empowered to review whether algorithmic decisions comply with constitutional rights, due process, and administrative law principles.[48] In jurisdictions where administrative decisions are reviewed by specialized courts, these bodies should be equipped to conduct substantive review of algorithmic reasoning and procedural fairness. This may necessitate the adaptation of evidentiary rules to allow for the presentation and critical assessment of algorithmic outputs. In some cases, enabling limited discovery or expert testimony may be necessary to ensure effective judicial oversight.[49] This includes evaluating whether affected individuals have access to meaningful recourse and ensuring that government agencies uphold transparency requirements.[50]

Algorithmic systems should be subject to periodic external evaluations conducted by independent researchers, civil society organisations, and academic institutions. These evaluations should analyse algorithmic fairness, predictive accuracy, and unintended consequences, ensuring that AI tools are aligned with ethical and legal standards.[51] Governments should consider establishing specialized regulatory agencies dedicated to overseeing the deployment of AI in public administration. These agencies should provide domain-specific expertise in algorithmic governance, data ethics, and human rights law, ensuring that AI implementations are equitable, accountable, and socially beneficial.

Beyond technical and legal evaluations, public participation should also play a crucial role in AI oversight. Government agencies should establish transparent feedback channels, allowing affected citizens and advocacy groups to report concerns, provide input on AI policies, and contribute to shaping AI governance frameworks. By involving diverse stakeholders in algorithmic oversight, governments can foster greater trust and inclusivity in AI-driven decision-making.[52] To institutionalize

[47] Mittelstadt et al., "The Ethics of Algorithms," 9–10.

[48] Kroll, "Accountable Algorithms," 633–705.

[49] Coglianese and Lehr, "Transparency and Algorithmic Governance," 23.

[50] "B-Tech Project Report."

[51] Brent Mittelstadt, "Principles Alone Cannot Guarantee Ethical AI," *Nature Machine Intelligence* 1, no. 11 (2019): 501–07, https://doi.org/10.1038/s42256-019-0114-4.

[52] Yeung, "Algorithmic Regulation," 505–23.

these oversight mechanisms, legislative mandates should be introduced to require public-sector AI systems to undergo independent audits before deployment and at regular intervals post-implementation. Non-compliance with these requirements should result in penalties or restrictions on the use of automated decision-making tools, reinforcing the necessity for ethical AI governance. By implementing independent oversight and audit frameworks, governments can mitigate the risks of algorithmic bias, enhance transparency, and safeguard public trust in AI-driven governance. A well-regulated AI ecosystem will ensure that technology serves democratic values, human rights, and social equity, rather than reinforcing existing inequalities or operating in an unaccountable manner.[53]

Summary

The adoption of decision-making algorithms in public administration presents both opportunities and challenges. While AI can enhance efficiency, streamline bureaucratic processes, and improve policy implementation, it also raises concerns regarding bias, transparency, accountability, and due process. Without adequate safeguards, algorithmic decision-making may reinforce systemic biases, undermine public trust, and erode democratic oversight. To address these risks, governments must implement strong transparency measures, including explainability frameworks, open data disclosure, and fairness assessments. The risk of automation bias underscores the importance of human oversight and specialized training for public officials to critically engage with AI-generated recommendations. Moreover, independent audits and regulatory oversight are crucial to ensuring that AI-driven governance complies with constitutional principles and ethical standards. Moving forward, a multi-faceted regulatory approach that includes mandatory audits, judicial review mechanisms, and participatory governance is necessary to align AI implementation with fundamental rights and social justice. By embedding transparency, fairness, and accountability into AI governance, policymakers can maximize the benefits of automation while safeguarding public trust and democratic values.

BIBLIOGRAPHY

Barocas, Solon, and Andrew D. Selbst. "Big Data's Disparate Impact." *California Law Review* 104, no. 3 (2016): 671–732. http://dx.doi.org/10.15779/Z38BG31.

Binns, Reuben. "Fairness in Machine Learning: Lessons from Political Philosophy." *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency* 81, 2018: 149–59.

[53] "B-Tech Project Report."

Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." *Proceedings of Machine Learning Research* 81, 2018: 77–91.

Citron, Danielle Keats, and Frank Pasquale. "The Scored Society: Due Process for Automated Predictions." *Washington Law Review* 89, no. 1 (2014): 1–33.

Coglianese, Cary, and David Lehr. "Transparency and Algorithmic Governance." *Administrative Law Review* 71, no. 1 (2019): 1–56.

Edwards, Lilian, and Michael Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For." *Duke Law & Technology Review* 16, no. 1 (2017): 18–84.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, 2018. https://doi.org/10.5204/lthj.v1i0.1386.

Florini, Ann. *The Right to Know: Transparency for an Open World*. Columbia University Press, 2007.

Kroll, Joshua A., Joanna Huey, Solon Barocas et al. "Accountable Algorithms." *University of Pennsylvania Law Review* 165, no. 3 (2017): 633–705.

Mittelstadt, Brent. "Principles Alone Cannot Guarantee Ethical AI." *Nature Machine Intelligence* 1, no. 11 (2019): 501–07. https://doi.org/10.1038/s42256-019-0114-4.

Mittelstadt, Brent, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* 3, no. 2 (2016): 1–21. https://doi.org/10.1177/2053951716679679.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, 2015.

Richardson, Rashida, Jason Schultz, and Kate Crawford. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice." *New York University Law Review Online* 94, 2019: 192–233.

Selbst, Andrew D., and Julia Powles. "Meaningful Information and the Right to Explanation." *International Data Privacy Law* 7, no. 4 (2017): 233–42. https://doi.org/10.1093/idpl/ipx022.

Veale, Michael, and Lilian Edwards. "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling." *Computer Law & Security Review* 34, no. 2 (2018): 398–404. https://doi.org/10.1016/j.clsr.2017.12.002.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation." *International Data Privacy Law* 7, no. 2 (2017): 76–99. https://doi.org/10.1093/idpl/ipx005.

Yeung, Karen. "Algorithmic Regulation: A Critical Interrogation." *Regulation & Governance* 12, no. 4 (2018): 505–23. https://doi.org/10.1111/rego.12158.

Zarsky, Tal. "The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making." *Science, Technology, & Human Values* 41, no. 1 (2016): 118–32. https://doi.org/10.1177/0162243915605575.