JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Ethics of using AI in Public Policy Decisions

Dhanashree R. Kulthe, Dhanashree Gund,Ayush Bhosale,Aniket Sapkal, H.R.Kulkarni, Rajashri Gaikwad*
* Author for Correspondence, Email: rajeshriwalunj@gmail.com
G H Raisoni College of Arts, Commerce & Science Pune, Maharashtra India.

Abstract:
Artificial Intelligence (AI) is increasingly used to inform, automate, or support public policy decisions. While the potential benefits-such as enhanced efficiency, consistency, and data-driven decision-making-are widely touted, ethical concerns around fairness, accountability,
transparency, privacy, algorithmic bias, and public trust are growing. This paper synthesizes findings from 25 recent research studies on these ethical dimensions to identify major ethical dilemmas, propose a framework for responsible Al in public policy, and suggest policy-
governance mechanisms that could mitigate risks. We use a mixed-methods approach combining systematic literature review, expert interviews, and comparative case studies. The results
highlight that transparency and accountability are often under-implemented, that ethical
governance tends to be reactive rather than proactive, and that citizens' involvement is limited. We propose a multi-stakeholder governance model with clearer roles, periodic audits, and legally binding frameworks. We conclude with directions for building ethical Al policy regimes,
especially in contexts with weaker governance infrastructures.
Overview One of the 21st century's most significant technologies is artificial intelligence.
Globally, it is becoming more and more incorporated into public governance and policymaking. AI is being used by governments to automate administrative processes, manage healthcare systems, identify crime-prone areas, predict traffic patterns, and distribute welfare funds. These applications seek to improve the effectiveness, data-drivenness, and societal responsiveness of
policies. But efficiency is only one aspect of public policy; other aspects include equality, justice, fairness, and the defense of citizens' rights. The ethical application of AI becomes crucial since
decisions about public policy have a direct impact on the lives of millions of people. AI's design, data, and results are not value-neutral.

Keywords: Artificial Intelligence, Ethics, Public Policy, Governance, Decision-Making.

I. Introduction:
Governments worldwide are adopting Al technologies to manage complex policy problems — from welfare fraud detection to predictive policing and pandemic forecasting. However, the delegation of sensitive policy decisions to algorithms raises questions about legitimacy, fairness, and accountability.Unlike private sector applications, public policy contexts involve high-stakes consequences, democratic legitimacy, and legal accountability. As such, ethical considerations in Al for public policy require stricter scrutiny [5]. This paper

investigates the ethical implications of Al in public policy decisions by reviewing 25 academic and policy contributions, analyzing
recurring principles and governance gaps, and proposing a framework for ethical Al governance in the public sector.

In recent years, governments around the world have begun integrating Artificial Intelligence into public policy and public administration. These applications range from predictive modeling for public health to automated decision systems in welfare, law enforcement, and infrastructure planning. While these technologies hold significant promise—improved policy targeting, better resource allocation, and faster, data-informed actions—they also carry substantial ethical risks.

Ethical concerns include algorithmic bias, lack of transparency, threats to privacy, risks of discrimination, undermining public trust, and the potential for reducing human autonomy. The deployment of AI in public policy raises fundamental questions: Who is responsible when an AI decision harms someone? How can citizens understand or challenge algorithmic decisions? How are ethical trade-offs made between efficiency and fairness? This paper aims to bring together
recent research on the ethics of AI in public policy decisions to map the major issues, assess how current policies and governance structures address them, and propose a framework for more
ethically robust AI deployment in public decision-making.

Ⅱ. Literature Review:

1. Ethical Principles in AI:

Most Al ethics frameworks emphasize principles such as fairness, accountability, transparency, privacy, and human oversight [1]. The UNESCO Recommendation on the Ethics of Artificial
Intelligence stresses inclusiveness, justice, and sustainability [3].Comparative analyses of
national Al policies reveal recurring clusters of principles including governance, data protection, and oversight [4].

2. Challenges in Public Policy Contexts:

Public sector adoption of Al presents unigue challenges not seen in private applications.
Algorithmic nudging can undermine citizen autonomy [7]. Al in welfare systems risks
reinforcing discrimination against vulnerable groups [8] Predictive policing has been criticized for perpetuating racial bias [9].Moreover, opacity in Al models raises concerns about
explainability and accountability [10]. 3.Governance Frameworks and Oversight:
High-level ethical principles require Translation into institutional mechanisms.Algorithmic Impact Assessments, third-party audits, and independent ethics boards are among the most discussed solutions [11;12]. Responsible Al governance requires balancing transparency with Security and efficiency with legitimacy [5].Practical tools such as fairness metrics, explainability Technigues, and human-in-the-loop design approaches have been proposed [13;14].

4. Research Gaps:

While the literature is rich in principles, empirical studies of Al implementation in government are scarce. Few Works address how principles are Operationalized across diverse contexts [12]. There is also limited Discussion of institutional capacity in Resource-constrained governments. This Creates a gap between ethical aspirations and real-world governance.

5. Fairness, Bias, Discrimination:

AI systems often reflect and amplify historical and systemic biases present in training data. For example, bias was identified as one of the primary ethical risks in public deployment by machine learning researchers [4]. Studies on AI in policing and welfare similarly show that biased
datasets can lead to unjust outcomes for marginalized groups [25].

6. Transparency and Explainability:

The opacity of AI algorithms is a recurrent concern. Decision-makers and citizens often lack

visibility into how automated policy tools function. Transparency is one of the most cited ethical principles in governance guidelines [2]. Proposals for public administration emphasize explainability to sustain trust and legitimacy [28].

7. Accountability and Responsibility:

Questions of liability remain unresolved: who is accountable when AI-driven public policies cause harm? Research indicates that existing legal frameworks lag behind AI applications in public administration [4]. Scholars argue that public agencies often rely on non-binding guidelines instead of enforceable accountability mechanisms[29].

8. Privacy, Data Protection, and Surveillance:

AI use in governance often depends on large-scale personal data, raising risks of surveillance and privacy infringement. Public surveys highlight strong concerns about data misuse, especially in healthcare settings [16]. Similarly, AI systems in law enforcement pose risks of overreach and discriminatory targeting[25].

9. Human Autonomy and Dignity:

Ethical challenges extend to autonomy and dignity, as AI systems may reduce human oversight in decisions that deeply affect individual lives. Research in China shows that citizens worry about AI undermining human freedom and dignity, even when efficiency gains are promised [25].

10. Trust, Legitimacy, and Public Engagement:

Public trust is essential for AI in governance to succeed. A scoping review found that while publics recognize the innovative potential of AI, they remain concerned about legitimacy and ethical oversight[30]. Research suggests that governments often engage citizens too late in the design process, undermining trust [28].

11. Legal and Regulatory Gaps:

While many countries publish national AI strategies, specific regulation for AI in public decision-making remains rare. A systematic review of responsible AI governance found that most frameworks emphasize principles but lack enforceable instruments[29].

12. Challenges in Low-Resource or Developing Contexts:

Ethical concerns are especially acute in low-resource contexts, where institutional capacities are weaker. In India, AI governance frameworks exist, but enforcement and oversight remain inconsistent[30]

III. Methodology:

This research employed a systematic literature review approach combined with qualitative thematic analysis to investigate the ethical implications of using Artificial Intelligence in public policy decision-making. The methodology followed established guidelines for systematic reviews in social sciences to ensure transparency, replicability, and rigor.

1. Research Design:

The study was designed as a qualitative, exploratory synthesis of existing peer-reviewed literature and scholarly reports. The focus was on identifying recurring ethical issues, governance mechanisms, and proposed solutions concerning AI use in public policy.

2. Data Collection:

Databases Searched: Web of Science, Google Scholar,etc.

Search Keywords: “AI ethics AND public policy,” “artificial intelligence AND governance,” “responsible AI AND public administration,” “AI accountability AND government decision- making,” “algorithmic bias AND public sector.”

Inclusion Criteria:

Peer-reviewed journal articles, conference papers, or reputable institutional reports published between 2020 and 2025.

Articles explicitly addressing AI ethics in the context of public policy or public administration. Papers written in English.

Exclusion Criteria:Technical AI papers without ethical or governance focus.Articles dealing only with private sector AI applications.

This process yielded an initial pool of 112 papers. After applying inclusion/exclusion criteria, screening abstracts, and removing duplicates, a final sample of 25 papers was selected for detailed analysis.

3. Data Extraction and Coding:

A structured data extraction sheet was developed to capture the following for each paper: Author(s), year, country/context.

Domain of application (healthcare, policing, welfare, governance, etc.).

Ethical issues highlighted (fairness, accountability, transparency, privacy, autonomy, etc.). Governance approaches or policy frameworks discussed.

Methodological approach (case study, survey, conceptual, review).

The extracted data was then coded using thematic analysis. Codes were clustered into key themes such as fairness, accountability, transparency, privacy, public trust, and regulatory gaps.

4. Data Analysis:

The analysis proceeded in three stages:

1. Descriptive Analysis – Mapping the distribution of publications by year, geography, and sector.

2. Thematic Synthesis – Grouping findings into major ethical themes and subthemes, highlighting consensus and divergence among scholars.

3. Comparative Analysis – Comparing governance approaches across contexts (e.g., EU, India, China, US) to identify patterns in ethical concerns and regulatory strategies.

5. Reliability and Validity:

Triangulation: Findings were cross-checked across multiple sources (academic papers, policy reports, and surveys). Peer Debriefing: Preliminary coding and themes were validated by two independent researchers to reduce subjective bias.

Transparency: The full list of included papers is provided in the References to enable replication.

IV. Result and Discussion:

1. Result:

Empirical evidence shows repeated instances of algorithmic bias and unequal outcomes when AI systems are used in public services (health, social care, policing, benefits). Several field studies and audits report that models trained on administrative or historical data tend to reproduce or amplify socioeconomic, racial and gender disparities. Governance assessments find large variation in readiness across governments: while many jurisdictions have ethics principles or guidance, relatively few have operational requirements (mandatory impact assessments, independent audits, clear procurement rules, inventories of models) and even fewer enforceable accountability mechanisms. This gap exposes citizens to harms despite widespread formal commitments to “transparency” or “fairness.” Mixed — but measurable — benefits to productivity and service delivery have been reported when AI is used for routine administrative tasks (e.g., document drafting, case triage). However, benefits are accompanied by error rates and cases of harmful outputs; studies stress that productivity gains do not automatically imply ethical acceptability. Reviews of policy documents show convergence on high-level ethical principles (human rights, transparency, fairness, accountability, privacy), but divergence on definitions and implementation pathways. Cross-country analyses reveal substantial heterogeneity in how those principles translate into concrete rules and oversight mechanisms. 2.Discussion:

1. Bias, data provenance and legitimacy. When public policy decisions rely on historical

administrative data, models can encode structural injustices (e.g., policing or welfare decisions). This raises normative questions about legitimacy: whose values are encoded, who benefits, and how redress is provided. Several authors argue that relying solely on post-hoc model explainers is insufficient — upstream decisions about data collection and labeling matter most.

2. Accountability and democratic oversight. A recurring theme is that automated decision-making shifts responsibility across actors (vendors, contractors, civil servants) and often reduces visibility for affected citizens. The literature stresses institutional solutions — model inventories, mandatory algorithmic impact assessments, independent audits, and public disclosure — as necessary conditions to restore democratic oversight. However, the effectiveness of such measures depends on legal teeth and resourcing.

3. Trade-offs: efficiency vs. rights protection. Practical deployments highlight a tension: AI can speed processing and free staff time, but blind reliance risks errors that disproportionately harm vulnerable groups. Policymakers face trade-offs between short-term efficiency and long-term trust; several works argue for phased pilots, rigorous evaluation, and human-in-the-loop safeguards to manage the trade-off.

4. Policy fragmentation & need for interoperability. Global scans show many overlapping or inconsistent frameworks. While international standards (e.g., UNESCO's Recommendation) provide high-level anchors, local operational norms and procurement rules vary — creating regulatory fragmentation that complicates cross-jurisdictional learning and vendor accountability. Harmonization of terminology and minimal compliance baselines is repeatedly recommended.

Conclusion and Future work:

- **Conclusion:**

The literature converges on a clear conclusion: AI can bring real service improvements in the public sector, but without robust governance and accountability instruments the risks to fairness, privacy and public trust are substantial. Ethical principles alone are insufficient — they must be converted into operational requirements (impact assessments, audits, transparency, redress).

Short-term productivity benefits observed in trials do not eliminate the need for sector-specific safeguards; governments must treat AI systems used in public policy as high-risk socio-technical systems that require continuous oversight and public contestability. Al in public policy offers both opportunities and dangers. Ethical risks such as bias, opacity, and accountability gaps threaten public trust and democratic legitimacy. Our review finds convergence on ethical principles but highlights the urgent need for operational governance frameworks.

- Future work:

2. Standardized empirical evaluations and benchmarks. Develop publicly auditable benchmarks and testbeds for domain-specific harms (e.g., fairness metrics for social care, bias stress tests for policing tools). Empirical studies should publish methodology and raw (anonymized) evaluation data where possible.

3. Comparative studies of governance instruments. Comparative, multi-country research on which accountability tools (e.g., impact assessments, independent auditing bodies, procurement clauses) actually reduce harms in practice — including cost/benefit and political feasibility analyses.

4. Participatory design & contestability mechanisms. Research and pilot programs that embed affected communities in the design, pre-deployment review, and remediation of systems (co- design, public hearings, appeal mechanisms). Measure how participation alters outcomes and trust.

5. Operationalizing transparency without harming security. Work on disclosure standards that provide meaningful transparency (model purpose, data provenance, evaluation results) while protecting privacy and

legitimate security concerns; this includes legal-technical work on redaction, secure audit logs, and standardized reporting formats.

6. Longitudinal impact studies & governance experiments. Fund and run longitudinal studies to observe socio-economic impacts of AI in public services over time and iterate governanceinstruments (regulatory sandboxes, staged procurement with mandatory audits).

V. Reference:

1. Khan, Z., et al. (2021).“Ethics of Artificial Intelligence: A Systematic Literature Review."
2. Corrêa, J.C., et al.(2022).“A Systematic Review of Al Ethics and Governance Frameworks."
3. UNESCO (2021). Recommendation on the Ethics of Artificial Intelligence. 4.Zhang, B., et al. (2023).“Mapping Ethical AI Policy Landscape." PLOS ONE.
5.Zuiderwijk, A., et al.(2021) “Implications of the Use of Al in Public Administration." Government Information Quarterly, 38(3)
6. Yar, M.(2024)."Ethics and Moral Alignment in Al." Information, 15(6)556. 7.Calboli, F.(2025).“Algorithmic Nudging and Public Policy" AsSociety.
8.Eubanks, V.(2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.St. Martin’s Press.
6. Lum, K, s Isaac, W.(2016).“To Predict and Serve?” Significance 13(5) 10.Mittelstadt, B, et al. (2016).“The Ethics of Algorithms.”Big Data sSociety, 3(2).
11. Reisman, D., et al. (2018) “Algorithmic Impact Assessments: A Practical Framework.” AI Now Institute.
12. Batool, F., et al.(2025).”AI Governance Frameworks: A Systematic Review”Als Ethics
13. Morley, J, et al.(2020).“The Ethics of Al: Mapping the Debate.”Nature Machine Intelligence, 2(2)
14. Floridi, L., et al.(2018).“A14People:An Ethical Framework for a Good AI Society.” Minds and Machines24(4)
5. Jobin A, lenca, M., s Vayena, E.(2016)."The Global Landscape of Al Ethics Guidelines." Nature Machine Intelligence, 1(6)
16. Whittlestone, J., et al. (2016).“The Role and Limits of Principles in Al Ethics." Proceedings of AIES.
17. Greene, D., Hoffmann, A., s Stark, L.(2016). “Better, Nicer, Clearer, Fairer:A Critical Assessment of Al Ethics Guidelines." Proceedings of FAT Conference.
18. Winfield, A, sJirotka, M.(2018) "Ethical Governance for Robotics and Al." Philosophical Transactions of the Royal Society A.
16.Veale, M, s Binns, R.(2017).“Fairer Machine Learning in the Public Sector" Philosophy s Technology 20.Selbst, A. D., et al.(2016).“Fairness and Abstraction in Sociotechnical Systems." Proceedings of FAT Conference.
21.Taddeo, M, sFloridi, L.(2018). “How Al Can Be a Force for Good.” Science, 361(6404)
22.Zuboff, S.(2016). The Age of Surveillance Capitalism.PublicAffairs. 23.Crawford, K.(2021). Atlas of Al. Yale University Press.
24. European Commission (2021) Proposal for a Regulation on Artificial Intelligence (Al Act)
25. Rani, M.(2023).Impacts and ethics of using Artificial Intelligence (AI) by the Indian Police. Public Administration and Policy, 26(3), 245-260
26. Zhang, B, Anderljung, M, Kahn, L, Dreksler, N, Horowitz, M.C., sDafoe A. (2021). Ethics and governance of artificial intelligence: Evidence from a Survey of machine learning researchers.
27. Corrêa, N. K. Galvão, C, Santos, J.W. Galvão, L., Pontes Pinto, E., Barbosa, C .sMambrini, R.(2022). Worldwide Al Ethics: A review of 200 guidelines and recommendations for Al governance.
28. Sigfrids, A., Nieminen, M, s Leikas J.(2022). How should public Administrations foster the ethical Development and use of artificial Intelligence? A review of proposals for developing governance of Al.
26. Grigalashvili, V.(2025). Artificial Intelligence in Public Administration:An Ethical Dilemma. International Journal of Innovative Technologies in Social Science, 4(40), 1-6.
30.Floridi, L., Cowls, J, Beltrametti, M,Chatila, R, Chazerand, P, Dignum, V….s Vayena, E.(2023). Publics’ views on ethical challenges of artificial intelligence: A scoping review. Al s Society