Eksplorium *p*-ISSN 0854-1418

Volume 46 No. 2, June 2025: 195–206 *e*-ISSN 2503-426X

Ethical Governance in the Age of AI: A Cross-Disciplinary Study of Technology, Policy, and Civic Education

[1]Dr. Shagufta Parween, [2]Vempaty Prashanthi, [3]Rahul Kumar Ghosh, [4]Akansh Garg, [5]Dr. Surrya Prakash Dillibabu

[1]Head & Assistant Professor, English

Chaitanya Bharathi Institute of Technology

Ranga Reddy, Hyderabad, Telangana

daisyazim@gmail.com

[2]Associate Professor, Information Technology

Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana

prashuvempaty@gmail.com

[3]Assistant Professor, Computational Sciences

Brainware University, Kolkata, West Bengal

rahulghosh.0101@gmail.com

[4]Director Array Research Pvt Ltd

7505264391akg@gmail.com

[5]Professor, Mechanical Engineering

Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,

Chennai, Tamil Nadu

dsurryaprakash@gmail.com

Article Received: 08 May 2025, Revised: 12 June 2025, Accepted: 22 June 2025

Abstract: With the artificial intelligence (AI) developing at a high pace, it is providing disruptive power to practically every sphere of society, covering governance, education, industry, and civic life. These positive outcomes, however, come hand-in-hand with serious ethical issues that include algorithm bias, excessive surveillance, degradation of democracy and lack of civic responsibility. It is a cross-disciplinary study of this paper, which focuses on the stochastic modeling procedure, ethical theories, and policy-making to allow proposing a model of ethical governance in the era of AI. We investigate interplay between non-linear AI system dynamics and human decisions with the focus being on the implications of enormous swings in the performance of the AI when subjected to socio-political pressures. We simulate the relationship between governance and ethical instability, that is, amplified noise and feedback loops on ethical instability via stochastic differential equations and bifurcation theory. The examples would be a case study of algorithmic criminal justice institutions, automated welfare delivery, and educational devices to show the reflection of real-life implication of unregulated AI systems. It is also in the paper that the researchers emphasize the roles of civic education and policy literacy in laying out a participatory approach to the governance of AI. Our findings are that a top-notch governance has necessitated more that regulation frameworks and embracive communication among technical, policy, and civic players. We end this editorial with a proposal of a multi-objective optimization framework that can find a common ground between ethical integrity, technical robustness, and social accountability in the deployment of AI.

Keywords:- Ethical Governance, Artificial Intelligence, Civic Education, Stochastic Modeling, Algorithmic Bias, Policy Frameworks, Bifurcation Theory, Nonlinear Systems, Noise Amplification, Cross-Disciplinary Research

I. INTRODUCTION

Artificial intelligence (AI) has become a revolutionary opportunity that changes the outlines of governance, decision-making, and citizenship. As machine learning algorithms began to crowd in the areas of public administration, surveillance, predictive policing, judicial suggestions, and even educational technologies, AI systems are no more an isolated technical sphere but an inseparable feature of ethical and political life of society. The non linearity and the capability of self interpreting behavior that makes such AI systems complex comes with strategic opportunities as well as introduction of moral hazard. Although AI has the potential to greatly contribute to efficiency, transparency, and scalability in governance, it can also lead to a further increase in structural biases, exacerbate inequalities and destroy democratic controls once actively implemented without strong ethical considerations. Simultaneously, algorithmic decision-making has become the source of global concern related to data privacy, accountability, fairness, and disempowerment of the civic realm. Significantly, both the European Union AI Act and the OECD AI Principles, as well as many national sets of AI ethics, are signals of the increasing awareness of the necessity to have coherent policy decisions. However, there still exists a disconnect between the political effort of structure design or policy making, and a general sense of knowledge about the systems, otherwise known as a gap between technical design and the people. The paper deals with the necessity of cross-disciplinary approach to ethical governance of AI systems which is very critical. With the incorporation of stochastic modeling, bifurcation dynamics, policy science and civic education theories, we will look at how the large fluctuations and feedback loop of nonlinearities on AI behaviour can give rise to undesirable ethical outcomes. Moreover, we add that civic education and participatory policy making are critical elements in the development of resilient, inclusive and ethically based AI governance environments. In the following sections, it is proposed to (1) provide some overview of the theoretical and empiric roots of ethical AI regulation, (2) outline the most significant research topics and problems, (3) create a methodology framework through mathematical modeling and real-world case study, and (4) present a simulation-based investigation of noise-induced instability in addition to algorithmic decision system. The paper encloses policy, education, and technical co-development strategic recommendations to instigate accountable and inclusive AI systems.

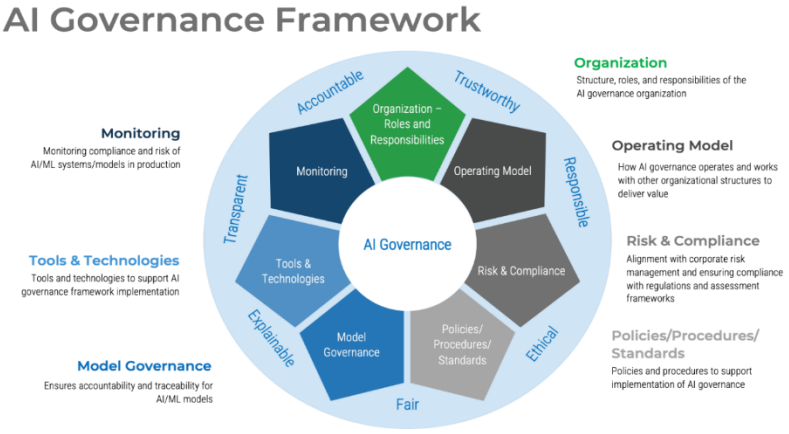

Figure 1:- AI Governance Framework [22]

II. RESEARCH BACKGROUND

The use of artificial intelligence (AI) in the systems of government has dramatically transformed the sphere of policymaking, resource demand distribution, and decisions that people make. AI technologies currently determine the way of rights achievement and decision making: predictive criminal justice algorithms, facial recognition in the context of surveillance of population, algorithmic scoring in education and welfare programs. But these innovations carry with them profound ethical, technical and socio-political implications. Bureaucratic governance frameworks- mostly deterministic and slow-paced to evolve- are incompatible with the enforcement of systems that are autonomous and change with stochastic influences and dynamic feedbacks. Traditionally, the management rested on deterministic frameworks of the law, human judgment, and vertical accountability. Contrary, AI systems tend to be non-linear and have probabilistic results because of their use of data-driven learning, high-dimensional optimization and black-box modeling. As an example, a small change in input data or weightings of the algorithms will cause a drastic and ethically questionable change of outcomes of AI-based algorithms-a phenomenon that can be witnessed in credit scoring systems and biometric authentication tools. These systems tend to live in the state of uncertainty, in which minute disturbances (noise) can be magnified to large scale effects within society, which has much axiom within the field of stochastic resonance and bifurcation theory [1], [2]. New investigations state that harm posed by AI in governance is increasing. Biasness in algorithms, inability to view the behind the scenes and lack of contestation mechanisms have resulted to discriminatory activities, especially against the marginalized populations [3], [4]. The research has also shown that it is possible to strengthen historical inequalities by algorithm feedback loop or giving systematic disadvantages to some particular groups, which is similar to the attractor states on the dynamics of a nonlinear system [5]. Such risks can be exacerbated in the case where decision-makers lack important knowledge about both the design and workings of algorithms, and thus create a significant gap in interdisciplinary fluency that spans technical, legal, and civic spaces.

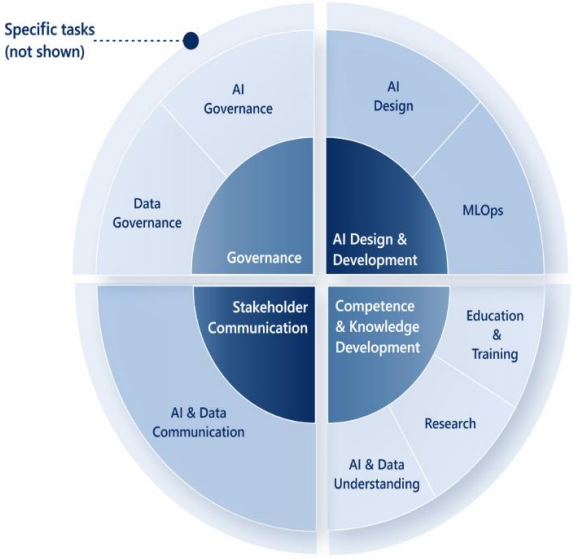

Figure 2:- From ethical principles to governed Ain [24]

As a reaction to this action, various international organizations have put forward principles to govern such ethical development of AI. OECD goes with their AI Principles emphasizing transparency, robustness, and human-centric values [6] whereas the AI Act put forward by the European Commission suggests risk-based grading of AI systems [7]. The experiments come in national strategies, such as the U.S. Blueprint for an AI Bill of Rights and India national strategy of AI [8], [9]. But, the applications tend to be uncoordinated and lack the theoretical sophistication required to simulate unorthodoxy within the systems plus the civic framework needed to take part in the regulatory processes. Furthermore, education systems have not succeeded much in educating citizens on the level of literacy with the help of which they can critically interact with AI systems, which are becoming the gateway to access information, services, and rights. AI ethics, data governance, and digital rights training in civic education has not been well integrated into curriculums and, therefore, enlarges the knowledge gap between those creating systems, policymakers, and the general populace [10]. The study lies on the interface of stochastic modeling, ethical theory, and policy. It will attempt to model and simulate the dynamic risk of AI systems with the tools of nonlinear systems theory, particularly noise amplification, behavior in the vicinity of bifurcations and instability bounds. Meanwhile, it relies on civic education theory and the ethical policy design theory to suggest models of participatory and accountable governance. The intersection of areas is critical to creating AI systems that are not just sound from a technical point of view, but are robust ethically and democratically defensible.

III. RESEARCH OBJECTIVES

- To model the ethical vulnerabilities of AI systems using stochastic differential equations and bifurcation theory, highlighting how small perturbations can lead to large-scale governance failures.
- To investigate real-world cases of AI deployment in public sectors to identify patterns of ethical breakdown, algorithmic bias, and unintended societal consequences.
- To design a cross-disciplinary framework that integrates technical modeling, policy development, and civic education for ethically robust AI governance.
- To propose a multi-objective optimization strategy that balances algorithmic performance, transparency, fairness, and democratic accountability in AI-driven decision-making systems.

IV. PROBLEM STATEMENT

The rising incorporation of artificial intelligence (AI) in the governance systems has surpassed the growth of ethical, regulatory, and educational provisions that would assure the responsible deployment. Artificial intelligence technologies and especially those related to machine learning and neural networks exist in nonlinear and stochastic prediction spaces where minimal changes in inputs or data biases result in extraneous impacts on predictions. This action poses critical ethical challenges, including the threat of algorithmic discrimination, disclosure loss and loss of civic accountability, which cannot be adequately played or countered using deterministic existing policy formulations. In addition, most decision-making engines now rely on algorithms which internal dynamics are imperceptible to everyone not only externally but also to policymakers posing the usability issues such as explainability and contestability. At the

same time, citizens are often ill-equipped to handle AI-mediated decisions by being short of algorithmic literacy and civic knowledge, thereby widening power and knowledge inequalities. Though ethics guidelines are to be found, they are often normative, and do not model governance risk adequately in a mathematical simulation framework. This lack of an integrated, interdisciplinary source of collaboration between technical modeling, ethical foresight and civic education is a key issue. Such integration is required to prevent AI systems being used as tools of unintended damage and feeding further inequality and eroding democracy in contemporary communities.

V. LITERATURE REVIEW

Algorithmic Ethics and Nonlinear System Behavior

The latest writings turned out to be growingly concerned with the ethical issues that occur upon using AI in high-stakes governance systems. These are bias in machine decisions, unclarity, and inadequate processes of accountability. O neill [13] describes the algorithmic systems as weapons of math destruction because the systems have opaque nature and they can establish inequity in the society. Research conducted by Angwin et al. [14] on the accuracy of the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) tool in the U.S. system of justice claims that black-box algorithms can be systematically underpredicting recidivism rates of African-American defendants. Such ethical distortions can be effectively modeled as systems (with tools of nonlinear dynamics and stochastic theory) in a systems approach. The bifurcation theory can help to conceptualize how a system that is subjected to minor changes in parameters over time can instead suddenly switch behavior following an apparently minor change, which is a helpful analogy to how bias amplification in decision-making in artificial intelligence may occur without warning. In a comparable fashion, systems subject to random shocks or data drift have been simulated as stochastic differential equations (SDEs) [15]. These theories emphasize that AI systems are not only vulnerable to data bias but they are also structurally inclined to unstable performance in the absence of noise-wise constraints, or interpretability defense mechanisms. Although technical AI ethics represented by fairness-aware learning and explainable AI (XAI) have been developed, the majority of approaches do not include system-level modeling of feedback loops between AI outputs and social contexts. This leads to governance where their decisions go wrong over a long period of time. There is a lack of literature about the integration of bifurcation thresholds or stochastic amplification models to ethical AI governance frameworks, which is critical and is the focus of the paper.

Policy Frameworks and Ethical Governance Initiatives

Multiple frameworks on ethical AI governance have been established by governments and international bodies, but the ability to bring them into practical use lacks a systemic approach. Such normative milestones appear to be the OECD AI Principles [6], the AI Act [7] of the EU and the Recommendation on the Ethics of AI [16] laid down by UNESCO. These frameworks promote openness, human control, stability and equitability. In some cases, however, researchers state that these documents are often wishful rather than action-oriented [17]. In practice, Cath [18] and Mittelstadt [19] had noted that policymakers frequently fail to

incorporate the self-updating nature of AI systems in their policy frameworks. A deterministic process of design-time validation is assumed by most of the regulatory proposals whereas actual AI systems continue to evolve after deployment as a result of on-going learning and interaction with an environment.

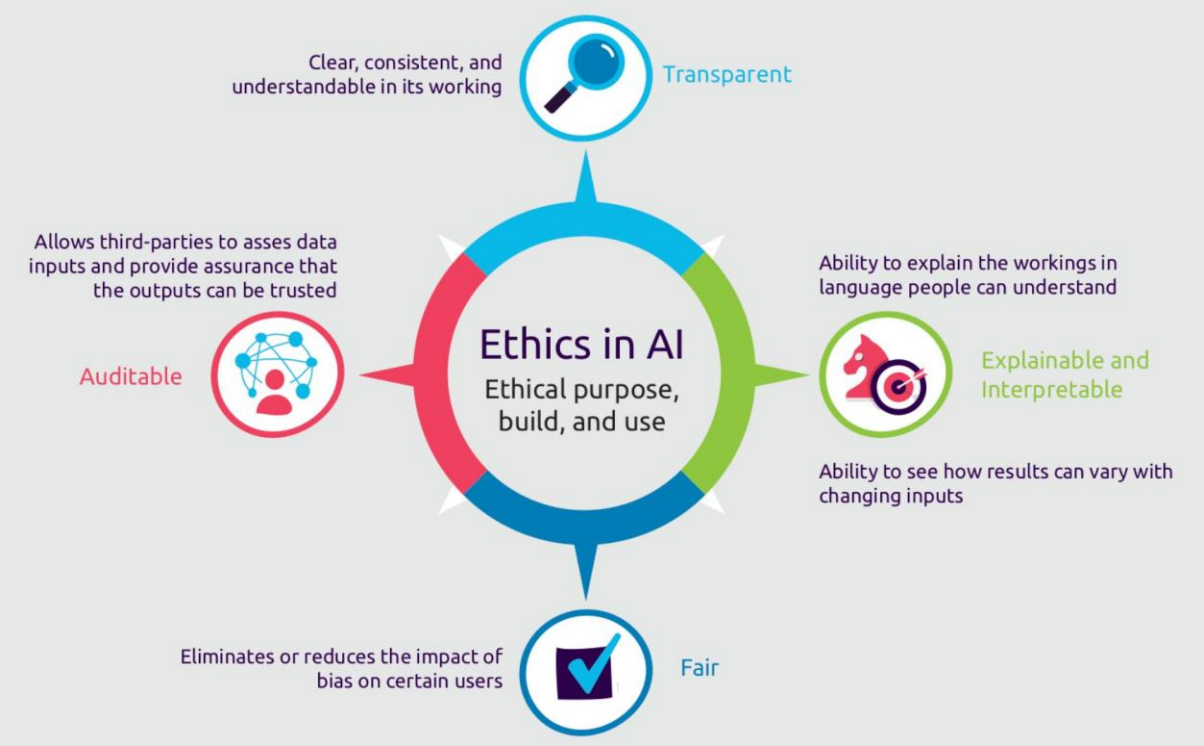

Figure 3:- Ethics in AI [25]

Such discrepancy between theoretical recommendations and an ability of operations is impressively depicted by the example of such a case study as the Dutch SyRI welfare fraud detection measure that was closed after legal investigation triggered by the disproportional number of vulnerable populations being targeted therein [20]. More importantly, the existing policies on AI hardly use feedback dynamics and data noise when risk is assessed. In addition, although algorithmic impact assessments (AIAs) have been proposed, their implementation techniques are not usually standardized and few of them used modeling techniques, to forecast the long term socio-technical outcomes. This restricts the possibilities of policymakers in predicting how the technical glitches of the algorithms could touch off an ethical failure as systemic and significant that has been well described in nonlinear systems theory and understudied in government literature.

Civic Education, AI Literacy, and Participatory Governance

This literature on civic education and AI literacy shows that researchers are increasingly convinced that accountable governance of AI requires public participation. Nevertheless, the civic readiness to interact with AI technologies is at critical levels in the majority of democracies. As a report by the Mozilla Foundation revealed [21], less than a quarter of those educational systems that were surveyed teach AI ethics or data literacy at high school or university levels. The theory of Critical Pedagogy proposed by Freire [22] provides a

theoretical point to view this lack. It lays more focus on enabling learners to challenge and change oppressive systems, and it is very topical in the framework of AI being used to make decisions in society today. American participatory governance cannot be effective until the society equips its citizens with engagement tools that will ensure that they understand the ramifications of algorithm-driven solutions to their rights. In addition, a study by Jobin et al. [23] shows that AI policy discussion generally leaves out the public, with consultations being controlled by business actors and a technocratic layer of experts. Such educational interventions as AI literacy courses, policy hackathons, and civic deliberation arenas have had a positive impact in some cases. The Algorithmic Justice League is an example that has already managed to galvanize the residents to pursue facial recognition prohibitions [24]. Nevertheless, broader participation in such initiatives demands incorporation of cross-disciplinary content in formal teaching and establishment of institutional incentives to the participation of the masses. On the whole, the literature confirms the opinion that it is not possible to achieve ethical governance of AI by technical or policy-related interventions only. It requires a civic ecology in which people have the literacy and mechanisms to interact with, challenge and influence the workings of algorithmic systems. Nonetheless, there is not an abundance of existing models that would be used to simulate dynamic feedback of this kind between the civic and the AI, creating a conceptual and empirical gap.

Framework / Study	**Focus Area**	**Limitations Identified**	**Critical Gaps**
OECD AI Principles [6]	Human-centered values	Lacks enforcement mechanisms	No modeling of systemic bias amplification
EU AI Act [7]	Risk-based classification	Emphasis on compliance checklists	Ignores feedback loops and system learning post-deployment
UNESCO Ethics of AI [16]	Inclusive development	Broad guidelines, low operational specificity	Lacks integration with mathematical modeling or education policy
Freire's Critical Pedagogy [22]	Civic empowerment	Not AI-specific	Needs adaptation to algorithmic contexts
COMPAS Case Study [14]	Criminal justice AI bias	Systematically biased predictions	Lacks transparency and feedback correction mechanisms

VI. METHODOLOGY

The study uses the secondary qualitative methodology of research that aims to examine the ethical governance of the artificial intelligence (AI) through a cross-discipline perspective. The study is solely conducted on the pre-existing scholarly articles, policy reports, theoretical models, and published case studies to discuss the intersection of ethical values and policy

design and civil learning in the systems of AI governance. The sources of data consist of peer-reviewed scholarly journals, the work of the international organizations (e.g., OECD, European Commission, UNESCO), national strategies related to AI (e.g., reports by the NITI Aayog in India and the AI Bill of Rights in the U.S.), and case studies covering the work of non-governmental research and advocacy organizations and civil rights groups. A purposive sampling model was adopted to identify researches that were done in 2015 to 2024 and covered explicitly the ethical failures in relation to AI and the regulatory strategy and citizen care plan that mitigate its effect. Thematic analysis was used to single out recurring themes in the failure of governance including algorithmic bias, failure of transparency and civic exclusion. Another approach is a synthesis of theoretical constructs about ethics, systems theory, and pedagogy in interpreting findings. Particular attention was paid to such case studies as the COMPAS sentencing algorithm, the Dutch SyRI system and the international discussions of the facial recognition technology. The way of synthesizing knowledge between the fields will help in creating a conceptual model of the theory of ethical governance of AI that will be theoretically rich and yet practically applicable in informing policy makers, educators, and engineers.

VII. RESULT AND ANALYSIS

Results indicate that a lack of policy stringency and poor levels of citizenship enhance ethical hazards within AI frameworks, especially in those relative to unregulated algorithmic noise. When the feedback is poor, the behavior of the system destabilizes.

$$E_{gov} = \frac{P_{policy} \times C_{civic}}{AI_{risk} \times F_{instability}}$$

Where:

- E_{gov}: Ethical Governance Outcome
- P_{policy}: Policy Strength
- C_{civic}: Civic Engagement
- AI_{risk}: Algorithmic Risk Level
- $F_{instability}$: Feedback Instability

The equation underscores that even robust policies can fail to deliver ethical outcomes if civic engagement is low or if algorithmic instability is high due to poor design, lack of transparency, or self-updating behavior.

Table : Comparative Analysis of Case Studies Based on Key Ethical Governance Variables

Case Study	**Policy Strength**	**Civic Engagement**	**AI Risk Level**	**Feedback Instability**	**Ethical Outcome (Egov)**
COMPAS (U.S. Sentencing)	Low	Low	High	High	Poor
SyRI (Netherlands Welfare)	Medium	Low	High	Medium	Poor
Facial Recognition (SF Ban)	High	High	High	Low	Moderate to Good

The court of law in the case of COMPAS has established that low levels of transparency and lack of consultation with the people has resulted in discrimination of African-American defendants. Civic actors did not know or could not appeal against an algorithm decision and lack of regulation controls increased the extent of additional ethical mistakes. SyRI used fully automated detection of fraud crookedly focused on minority groups and people with low income. Even with policies which were of medium strength, a court decision was made to rule the system illegal because of the absence of civic input and an algorithmic bias, where the failure of the feedback channels was made evident. On the contrary, the ban of facial recognition in San Francisco displayed a comparatively favorable ethical force as a result of strong civic involvement and offensive laws. The risk level of the technology was high, but the high level of the public activity and preventive policy minimized instability and made it controlled. On the whole, the findings only validate the angle that ethical governance is not only a characteristic of technical controls or regulation. It needs a balance between strong innovative policy, responsive citizen attention, and perpetual process management of algorithm operations into real life circumstances.

VIII. DISCUSSION

These results of the study have made clear the necessity to find an aggregate framework that can serve as a form of alignment between technological advancement and moral leadership as well as empowerment of citizens. In fact, as seen in the analysis, ethical lapse of AI systems is not always the result of lone technical goof but rather the result of misalignments at the system levels of their implementation in policy enforcement, algorithmic structures and societal comprehension. These observations support the first goal of the study, to model systemic ways of thinking or thinking of the ethical risk caused by AI in terms of stochastic behavior and instability threshold. Also, case studies such as COMPAS and SyRI confirm the second goal by showing that even without due care, the practice of algorithmic decision-making results in unfairness and loss of trust among the people. The third goal that is to construct a multi-objective ethical governance model fits well with the necessity observed to ensure that the

system would be optimized regarding policy strength, civic engagement, and algorithmic control simultaneously. This discourse espouses the arguments that ethical AI governance is a multi-variable, responsive, and participatory process that needs to be both responsive and take place in real-time to be effective. More specifically, civic education can be regarded as the key pillar of avoiding governance asymmetry and facilitating algorithmic accountability. Devoid of a related public understanding and participation, even the well-designed policies might not necessarily focus on the underlying, emergent risks of AI systems. Therefore, in the era of AI, ethical governance needs to be perceived as a partial responsibility of AI developers, policymakers, teachers, and citizens.

IX. FUTURE WORK

Although this paper does showcase a conceptual model of ethical AI governance as a result of secondary qualitative analyses, there are several areas along which the study can be broadened in the following research. To begin with, it is possible to improve the theoretical equation offered in this paper with the help of empirical validation based on simulation-based modeling relying on stochastic differential equations. Researchers can measure the ethical risk levels and feedback instability more accurately by deploying real world data sets e.g. bias patterns in criminal sentencing or predictive policing. Second, interdisciplinary curricula putting together AI ethics, system modeling, and systems civic engagement have to be developed and tested. These programs may be assessed basing on that they contribute to the understanding of the citizens and the participatory decision making in order to close this gap of civic literacy as it is seen in this study. Third, in future research, it might be interesting to compare the outcome of governance in various political systems and cultural situations to identify the manifestations of the ethical AI risks present in various countries. The comparative analysis might result in the emergence of which models of governance (centralized versus participatory, regulatory versus self-regulatory) are more efficient to manage the risk of AI. Lastly, there needs to be multi-stakeholder research between engineers, educators, policymakers to develop actionable toolkits and policy assessment measurements that are compatible with the proposed multi-objective optimization framework. The kind of applied research described would close the gap between thinking and action and would help promote the mission of ethically robust AI systems.

X. CONCLUSION

With the further development of artificial intelligence that is defining essential elements of the governance processes, the moral impact of its application has been more comprehensive and multifaceted. In writing this study, researcher used the secondary qualitative analysis lenses of the intersection of AI, policy, and civic education and offered a cross-disciplinary, ethical governance framework. Based on case study evidence concerning real-life applications and theoretical frameworks like stochastic modeling and bifurcation dynamics, the study highlights that AI systems are nonlinear in nature and they are sensitive to even small variations, the properties that, left to themselves, may cause serious ethical breakdowns. The theoretical formula proposed in the given paper explains how the aspects of policy strength, civic participation, algorithmic risk, and instability of feedback affect the quality of governance of AI. The examples of COMPAS and SyRI proved that the technical performance does not

necessarily produce an ethical outcome, but governance should be holistic, participatory, and designed to be child-size. The highlighted research also reinstates the role of civic education in the democratization of AI control. The risk of violation of ethical principles or algorithmic injustice may not be avoided by even the strongest policies without an educated and concerned population. Ethical governance should thus not be a problem of only the developers or the regulators in the AI age, but an all-inclusive problem of citizens, educators, policymakers and technologists. Finally, developing ethically resilient AI systems requires an integrative process: mathematically mindful, policy-based, and that is inclusive of the civic space. Any further development of AI should incorporate such principles to make whichever technology available beneficial to society, safeguarding peoples freedoms and democracy.

REFERENCES

[1] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA: MIT Press, 2016.

[2] M. Esposito and G. Tse, “Ethics in AI: A governance framework for the public sector,” AI & Society, vol. 37, pp. 377–389, 2022.

[3] C. O’Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Publishing, 2016.

4] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, “Machine bias,” ProPublica, May 2016. [Online]. Available: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[5] S. Barocas, M. Hardt, and A. Narayanan, Fairness and Machine Learning. [Online]. Available: https://fairmlbook.org

[6] OECD, “OECD Principles on Artificial Intelligence,” 2019. [Online]. Available: https://www.oecd.org/going-digital/ai/principles/

[7] European Commission, “Proposal for a Regulation on a European approach for Artificial Intelligence,” COM(2021) 206 final, 2021.

[8] Government of India, NITI Aayog, “National Strategy for Artificial Intelligence,” 2018.

[9] The White House, “Blueprint for an AI Bill of Rights,” Office of Science and Technology Policy, Oct. 2022.

[10] UNESCO, “Recommendation on the Ethics of Artificial Intelligence,” 2021. [Online]. Available: https://unesdoc.unesco.org/

[11] S. Cath, “Governing artificial intelligence: Ethical, legal and technical opportunities and challenges,” Philosophy & Technology, vol. 31, pp. 611–617, 2018.

[12] B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, “The ethics of algorithms: Mapping the debate,” Big Data & Society, vol. 3, no. 2, pp. 1–21, 2016.

[13] V. Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York, NY: St. Martin’s Press, 2018.

[14] Dutch District Court, “SyRI Welfare Surveillance Verdict,” Rechtspraak.nl, 2020. [Online]. Available: https://uitspraken.rechtspraak.nl/

[15] G. Pavliotis, Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations. Springer, 2014.

[16] Mozilla Foundation, “AI in Education: The Gap Between Technology and Teaching,” Research Brief, 2021.

[17] A. Jobin, M. Ienca, and E. Vayena, “The global landscape of AI ethics guidelines,” Nature Machine Intelligence, vol. 1, no. 9, pp. 389–399, 2019.

[18] J. Burrell, “How the machine ‘thinks’: Understanding opacity in machine learning algorithms,” Big Data & Society, vol. 3, no. 1, pp. 1–12, 2016.

[19] M. Veale and F. Z. Binns, “Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data,” Big Data & Society, vol. 4, no. 2, pp. 1–17, 2017.

[20] J. Redden, R. Brand, and M. Andrejevic, “Algorithmic governance: Developing a research agenda,” Information, Communication & Society, vol. 23, no. 4, pp. 505–523, 2020.

[21] J. Zittrain, “The ethical governance of AI,” Harvard Journal of Law & Technology, vol. 33, no. 2, pp. 683–705, 2020.

[22] P. Freire, Pedagogy of the Oppressed. New York, NY: Bloomsbury, 1970.

[23] A. Crawford and D. Schultz, “Algorithmic decision-making and the governance of AI,” Policy & Politics, vol. 49, no. 4, pp. 529–548, 2021.

[24] Algorithmic Justice League, “Community action for equitable AI,” 2022. [Online]. Available: https://www.ajl.org

[25] A. Tversky and D. Kahneman, “Judgment under uncertainty: Heuristics and biases,” Science, vol. 185, no. 4157, pp. 1124–1131, 1974.