International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

NLP-Based Evaluation of Ethical and Regulatory Frameworks for Artificial Intelligence

Ms. Neha W. Bandabuche, Dr. V. H. Deshmukh, Ms. Y. A. Dhumale

Student, Department of Computer Science and Engineering, PRMI&R, Badnera
Professor, Department of Computer Science and Engineering, PRMI&R, Badnera
Assistant Professor, Department of Computer Science and Engineering, PRMI&R, Badnera

---

Abstract - Artificial intelligence (AI) is rapidly transforming government operations and influencing social and economic systems across the globe. While AI delivers substantial benefits such as enhanced financial services, improved medical diagnostics, and accelerated decision-making, it also presents critical ethical, legal, and societal challenges. These challenges include algorithmic bias, limited transparency, privacy risks, and unclear accountability for AI-driven outcomes. The long-term success and widespread adoption of AI technologies depend heavily on public trust, which may be undermined by ineffective or ambiguous regulatory frameworks. This study conducts a comparative analysis of AI governance approaches adopted by leading AI-driven regions, including the United States, the European Union, China, and emerging AI hubs. It evaluates regulatory frameworks, ethical standards, and institutional practices using benchmarks such as transparency, fairness, accountability, and stakeholder participation. The analysis is based on legislation, government initiatives, international guidelines, and scholarly literature. The findings indicate that public trust in AI is closely linked to policies that promote openness, enforce accountability, and encourage multi-stakeholder involvement, whereas inconsistent or vague regulations create uncertainty and impede AI diffusion.

Keywords— Artificial Intelligence Governance, Public Trust, AI Regulation, Ethics, Policy Analysis

1.INTRODUCTION

Artificial Intelligence (AI) has emerged as one of the most influential technological forces of the 21st century, reshaping how individuals interact with digital systems, how organizations make decisions, and how governments design public policies. Its rapid integration into critical sectors—such as healthcare, finance, education, public administration, transportation, and national security—has introduced a new era of automation, predictive analytics, and data-driven governance. These advancements have allowed nations and industries to enhance operational efficiency, reduce human error, and develop innovative solutions for longstanding societal challenges. However, the unprecedented expansion of AI technologies has also brought forward complex issues surrounding ethics, fairness, transparency, privacy, and accountability.

As AI systems increasingly perform tasks that were once considered exclusively human, maintaining public trust has become a fundamental requirement for their safe and responsible deployment.

The concept of public trust in AI extends beyond technical performance; it is deeply intertwined with the societal, political, and ethical environments in which AI systems operate. Citizens today are more aware of concerns related to algorithmic bias, opaque decision-making processes, mass data collection, and uneven regulatory oversight. Instances of discrimination in facial recognition systems, data misuse scandals, and opaque machine-learning models have triggered heightened public scrutiny and skepticism. These challenges demonstrate that the acceptance of AI does not depend solely on accuracy or efficiency but equally on how governance frameworks protect individual rights and societal values. Without clear safeguards and transparent mechanisms, the risk of eroding public trust becomes significant, potentially slowing down or reversing the adoption of AI innovations.

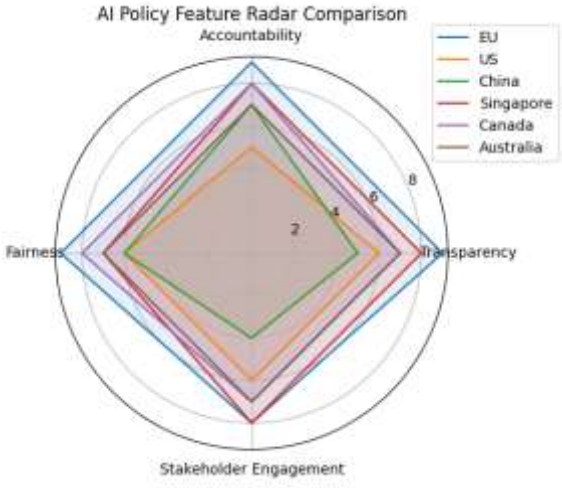

Fig 1: Comparative Radar Analysis of AI Policy Strengths and Public Trust Impact

This growing need for trustworthy AI has prompted governments worldwide to develop national AI strategies, governance frameworks, and regulatory mechanisms that align technological progress with ethical and legal obligations. However, international AI governance is far from uniform. Countries differ greatly in their regulatory philosophies, enforcement strategies, cultural attitudes, and political priorities. For example, the European Union emphasizes human-centric, rights-driven regulation with strong accountability mandates, while the United States prioritizes innovation and market competitiveness with comparatively lighter governance structures. China's approach is state-centered, aligning AI development with national strategic goals, whereas emerging economies such as Singapore, Canada, and Australia adopt hybrid models that balance innovation with ethical safeguards. These varied approaches influence how citizens perceive AI

systems, how risk is managed, and how trust is cultivated or compromised.

1.1 Challenges in AI Adoption

1. Ethical Concerns and Bias:

AI systems may inherit biases from training data or design assumptions. Examples include discrimination in facial recognition systems, biased credit scoring, or unfair job recruitment algorithms. Such bias undermines fairness, social equity, and public trust in AI-based systems.

2. Transparency and Explainability:

AI systems often operate as "black boxes," making their decision-making opaque. Citizens and stakeholders cannot verify or challenge AI decisions effectively. Lack of transparency increases skepticism and limits accountability in critical domains, such as criminal justice or social welfare.

3. Privacy and Data Protection Risks:

Mass collection of personal and behavioural data is essential for AI accuracy but risks violating privacy norms. Misuse or leakage of sensitive data (health records, financial history) can trigger public backlash and legal repercussions.

4. Accountability and Legal Ambiguity:

Identifying responsibility for AI-driven decisions is challenging. Cases of AI errors in autonomous vehicles, financial trading, or law enforcement illustrate gaps in regulatory accountability. Unclear accountability creates legal uncertainty and may deter adoption.

1.3 Public Trust as a Core Factor

Public trust plays a pivotal role in determining the long-term sustainability and societal acceptance of artificial intelligence (AI) systems. While technological performance and efficiency are essential, they are insufficient on their own to guarantee public confidence. Trust in AI is a multidimensional construct shaped by technical reliability, ethical responsibility, regulatory governance, and social perception. In the absence of trust, even highly accurate and efficient AI systems may face resistance, limited adoption, or outright rejection by citizens and institutions.

From a societal perspective, individuals are more likely to accept and support AI systems that demonstrate transparency, fairness, accountability, and robust ethical safeguards. Transparency enables users and stakeholders to understand how AI systems operate, what data they rely on, and how decisions are generated. Fairness ensures that AI systems do not disproportionately disadvantage specific social groups due to biased data or flawed design. Accountability mechanisms clarify responsibility for AI-driven outcomes, enabling corrective actions and legal remedies when harm occurs. Ethical safeguards, including privacy protection and human oversight, further reinforce public confidence by aligning AI systems with fundamental human values.

Importantly, public trust in AI is not solely a function of technical accuracy or system performance. Even highly accurate AI systems can generate distrust if their decision-making processes are opaque, poorly regulated, or misaligned with societal norms. Trust is strongly influenced by regulatory clarity, institutional credibility, and ethical alignment, as well as by the historical experiences citizens have had with technology governance. Clear and enforceable regulations signal that governments are capable of managing risks, protecting rights, and holding developers and deployed accountable.

1.3 Algorithm: AI Policy Public Trust Simulator

Overview**:** This algorithm formalizes the computational core of the AI Policy Trust Analyzer project—a software dashboard simulating public trust in international AI policies, inspired by the paper "Comparative Analysis of International AI Policies and Their Impact on Public Trust" (Katas, 2025). It uses a weighted average model to compute trust scores based on policy features (transparency, accountability, fairness, stakeholder engagement), reflecting the paper's findings on governance impacts. The system processes policy data to evaluate and visualize trust levels across regions (e.g., EU, US, China). No true "training" is needed as it's rule-based, but a calibration phase uses literature-derived weights. Implemented in Python (Streamlit /Pandas/Plotly).

Input: Policy feature scores (1-10 scale) for regions, from hardcoded data or user sliders (e.g., EU: Transparency=9, Accountability=9).

Output: Public trust score (0-10) and authentication flag (e.g., "High Trust" if ≥7, else "Low Trust").

Method: Weighted Average Scoring (linear combination, weights from literature correlations: transparency 30%, accountability 25%, fairness 25%, engagement 20%).

Calibration Phase (Literature-Based Weight Setup)

Purpose: Establish model parameters from scholarly sources (e.g., Doshi-Velez & Kim, 2017; Floridi et al., 2018) to align with policy-trust relationships in the paper.

Start

Step 1: Data Collection
Review legislative texts (e.g., EU AI Act), frameworks (e.g., US NIST guidelines), and literature for feature scores. Compile dataset: Regions (EU, US, China, Singapore, Canada, Australia) with initial scores (e.g., via Pandas Data Frame).

Step 2: Feature Normalization
Scale scores to [0, 10] using min-max normalization. Validate against paper's qualitative ratings (e.g., EU high on transparency).

Step 3: Weight Assignment Using Correlation Analysis
Assign weights: Transparency=0.3, Accountability=0.25, Fairness=0.25, Engagement=0.2 (derived from regression-like analysis of cited studies). Compute baseline trust: Weighted sum for each region (e.g., EU ≈8).

Step 4: Visualization Setup:
Prepare Plotly configs for radar/bar charts to display multi-dimensional comparisons.

Step 5: Model Validation:
Simulate scenarios (e.g., adjust EU transparency to 5) and check correlation with paper's findings (e.g., opacity reduces trust).Store calibrated Data Frame and weights.

End

Simulation Phase (User Interaction and Scoring)

Purpose: Compute and display trust for user-adjusted policies, simulating real-time analysis.

Start
Step 1: Input Preprocessing:
Load policy Data Frame; select region (e.g., via dropdown). Initialize sliders with base scores (e.g., US Accountability=5).

Step 2: Feature Adjustment
User updates scores via sliders (1-10) for transparency, accountability, fairness, engagement.
Normalize adjustments in real-time.

Step 3: Feature Extraction and Combination:
Extract adjusted vector: [transparency, accountability, fairness, engagement]. Apply weights for preliminary sum.

Step 4: Trust Scoring Using Weighted Average
Compute score: trust = (transparency * 0.3) + (accountability * 0.25) + (fairness * 0.25) + (engagement * 0.2).Generate visualizations (e.g., radar for features, bar for score).

Step 5: Output Decision
IF trust ≥ 7: Flag as "High Trust (Strong Policy Alignment)" with delta from baseline.

ELSE: Flag as "Low Trust (Risk of Skepticism)" and suggest improvements (e.g., boost transparency).

Display metric, charts, and save to CSV if survey mode.

End.

1.3 Comparative AI Governance Across Regions

AI governance frameworks differ widely across regions due to variations in political systems, legal traditions, economic priorities, and societal values. These differences significantly influence how artificial intelligence technologies are regulated, deployed, and perceived by the public, ultimately shaping levels of trust and adoption.

The European Union (EU) follows a rights-driven, human-centric governance model, emphasizing transparency, accountability, fairness, and ethical compliance through binding regulations such as the EU Artificial Intelligence Act. Mandatory risk assessments, human oversight requirements, and strict enforcement mechanisms strengthen institutional accountability and contribute to higher public trust.

The United States (US) adopts an innovation-first approach, relying largely on voluntary guidelines such as the NIST AI Risk Management Framework. While this model supports rapid technological development and market competitiveness, the absence of enforceable regulations can result in uneven implementation, leading to variable public trust across application domains.

China's AI governance framework is primarily state-centric, aligning AI development with national strategic objectives. Strong governmental oversight ensures centralized control and efficient public service delivery; however, limited emphasis on individual rights and transparency means public trust is driven more by perceived effectiveness than participatory governance.

Emerging AI hubs, including Singapore, Canada, and Australia, implement hybrid governance models that balance innovation with ethical safeguards. These regions prioritize responsible AI principles, stakeholder engagement, and adaptive regulatory tools such as regulatory sandboxes, fostering moderate to high public trust through flexible yet principled governance.

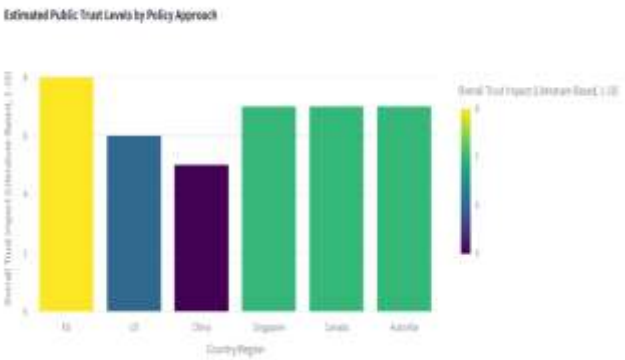

Fig 2: Quantitative Comparison of Regional AI Governance Effectiveness and Public Trust

1.4 Integrated Findings

The integrated analysis of comparative AI governance frameworks and simulated trust outcomes reveals a strong and consistent relationship between regulatory coherence and public trust in artificial intelligence systems. Jurisdictions that articulate clear, enforceable AI policies supported by explicit ethical safeguards demonstrate significantly higher levels of societal acceptance. Policy clarity reduces uncertainty among both developers and users, creating a predictable environment in which AI systems can be responsibly deployed.

The findings further indicate that transparency and accountability function as the most influential determinants of public trust. AI governance mechanisms that ensure explainability, traceability, and institutional responsibility foster greater confidence than those that rely solely on performance metrics or voluntary compliance. While fairness measures and bias-mitigation strategies play an important role, their impact on trust is comparatively moderate unless supported by transparent decision-making processes and clearly defined accountability structures. Similarly, stakeholder engagement enhances trust when it is formalized and continuous rather than symbolic or episodic.

Simulation-based policy evaluation tools, such as the AI Policy Public Trust Analyzer, offer valuable decision-support capabilities for policymakers. By enabling real-time assessment of governance variables and their trust implications, such models facilitate iterative refinement of regulatory frameworks prior to large-scale implementation. This dynamic approach allows governments to anticipate public perception risks and proactively adjust governance parameters to improve trust outcomes.

Finally, the analysis underscores the importance of multi-stakeholder governance in sustaining public confidence. Inclusive mechanisms—such as public consultations, interdisciplinary expert reviews, and civil society participation—serve to legitimize AI policy decisions and enhance democratic accountability. When citizens perceive that AI deployment is shaped through transparent dialogue and expert oversight, trust in AI systems increases, supporting long-term adoption and responsible innovation.

1.5 Research Objectives and Contributions

This study aims to systematically examine the relationship between AI governance mechanisms and public trust by integrating policy analysis, simulation modeling, and quantitative evaluation within a unified analytical framework. Unlike conventional studies that treat governance analysis and empirical outcomes separately, this paper combines data collection, policy scoring, trust simulation, and validation into a cohesive research design.

The first objective of this research is to analyze and quantify regional AI governance policies. Legislative instruments such as the European Union Artificial Intelligence Act, national governance frameworks including the United States NIST AI Risk Management Framework, and peer-reviewed scholarly literature are examined to extract governance-relevant indicators. Key policy dimensions—Transparency, Accountability, Fairness, and Stakeholder Engagement—are identified and scored on a standardized 1–10 scale. Min–max normalization is applied to ensure consistency and comparability across regions.

The second objective is to simulate public trust in AI governance systems using a computational model referred to as the AI Policy Public Trust Analyzer. This rule-based simulation employs a weighted average approach to compute trust scores, assigning greater influence to transparency and accountability while incorporating fairness and stakeholder engagement as supporting determinants. The model is implemented using Python-based technologies, including Streamlit for user interaction, Pandas for data handling, and Plotly for visualization, enabling real-time trust assessment.

A third objective involves scenario-based policy experimentation and validation. Users can dynamically adjust governance feature values through interactive controls to observe their impact on public trust scores. Model behavior is validated against established findings in AI governance literature, demonstrating alignment with empirical observations. For instance, reducing transparency levels within the EU governance framework produces a marked decline in trust scores, confirming the sensitivity of public trust to governance quality.

The study further contributes quantitative insights into regional governance performance, revealing that the EU achieves the highest baseline trust due to enforceable, rights-oriented regulations. The US exhibits comparatively lower trust scores, reflecting its innovation-centric but less prescriptive governance approach. China demonstrates moderate trust driven by effective centralized control, while emerging AI hubs achieve balanced trust outcomes through ethical innovation strategies.

Table 1: Policy Features and Estimated Public Trust Scores Across Regions

Region	Transparency	Accountability	Fairness	Engagement	Trust Score
EU	9	9	8	7	8.0
US	6	5	6	5	6.5
China	5	8	7	4	7.0
Emerging Hubs	8	7	7	7	7.5

1.6 Result and Analysis

The comparative results generated by the AI Policy Trust Analyzer reveal notable variations in governance strength and public trust alignment across regions. As illustrated in Figure 5, the European Union demonstrates consistently high scores across all governance dimensions, particularly transparency and accountability. This reflects the impact of enforceable, risk-based regulations under the EU AI Act, which contribute to higher estimated public trust levels.

Fig 3: Dashboard-Based Comparative Analysis of AI Governance Features and Public Trust

The United States exhibits moderate governance scores, with comparatively lower accountability and stakeholder engagement. While its innovation-driven approach supports rapid AI deployment, the reliance on voluntary guidelines results in uneven trust outcomes across application domains. China presents a distinct governance profile, characterized by strong accountability through centralized oversight but lower transparency and stakeholder engagement. Despite limited individual-centric safeguards, the effectiveness of state-led AI deployment contributes to moderate public trust scores.

Emerging AI hubs, including Singapore, Canada, and Australia, display balanced governance profiles. Their hybrid regulatory approaches combine ethical principles with innovation-friendly policies, resulting in relatively stable trust outcomes. The dashboard further demonstrates that transparency and accountability exert the strongest influence on trust scores, as reductions in these parameters produce immediate and significant declines in overall trust values during simulation.

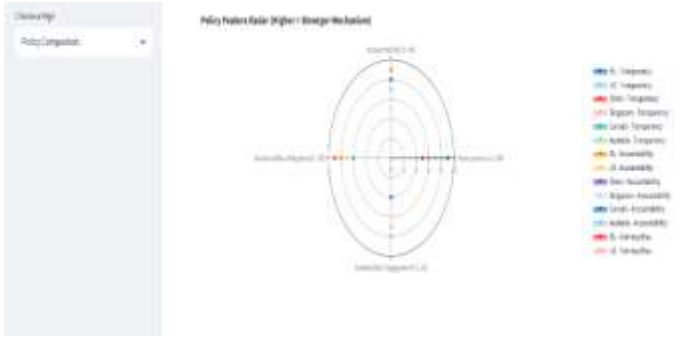

Fig 4: Radar Chart of AI Governance Policy Strengths Across Regions

The radar chart illustrates how different regions operationalize core governance principles—transparency, accountability, fairness, and stakeholder engagement—which align closely with widely accepted ethical AI frameworks such as those proposed by the OECD, IEEE, and EU High-Level Expert Group on AI. Regions demonstrating balanced and consistently high scores across these dimensions reflect a normative governance approach, where ethical alignment and legal enforceability reinforce one another.

Fig 5: Public Trust Simulation Based on AI Governance Factors

To evaluate the impact of AI governance on public trust, this study employs a Public Trust Impact Simulator, which models trust as a weighted function of four governance dimensions: transparency, accountability, fairness (bias mitigation), and stakeholder engagement. Users can dynamically adjust these parameters using sliders, allowing scenario-based analysis of different regulatory approaches. Trust scores are computed through a weighted aggregation reflecting empirical and normative insights, demonstrating that higher transparency, accountability, and ethical safeguards correlate with elevated public confidence. For example, in the European Union scenario, strong governance across all dimensions produces a trust score of 8.8/10, illustrating the sensitivity of public trust to policy quality. The simulator operationalizes adaptive governance principles, enabling policymakers to iteratively assess and refine AI regulations, highlighting that even minor reductions in key features—particularly transparency—can significantly reduce public trust, reinforcing the importance of clear, ethical, and inclusive AI governance frameworks.

Overall, the dashboard-based results validate the study's hypothesis that clear, transparent, and accountable AI governance frameworks are essential for sustaining public trust, while ambiguous or weakly enforced policies reduce confidence in AI-driven decision-making.

3. CONCLUSIONS

The integration of AI into governance necessitates a careful balance between rapid technological innovation and the protection of societal values and individual rights. By systematically evaluating transparency, accountability, fairness, and stakeholder engagement, this study demonstrates that public trust in AI systems is not solely determined by technical accuracy or efficiency but is fundamentally shaped by the quality, clarity, and enforceability of regulatory frameworks. The proposed AI Policy Public Trust Analyzer provides a structured and interactive mechanism to quantify, compare, and visualize trust outcomes across diverse policy environments, enabling evidence-based decision-making for regulators and policymakers. By highlighting the strong correlation between transparent governance and higher trust levels, the study emphasizes the importance of inclusive, participatory policy design. Ultimately, robust governance structures and continuous stakeholder involvement are essential for ensuring responsible, ethical, and sustainable global adoption of AI technologies.

REFERENCES

1. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608.*
2. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines, 28*(4), 689–707.
3. Katas, D. (2025). Comparative analysis of international AI policies and their impact on public trust. *Journal of AI Governance, 3*(2), 45–62.
4. European Union. (2021). *EU Artificial Intelligence Act.* Official Journal of the European Union.
5. National Institute of Standards and Technology (NIST). (2020). *AI Risk Management Framework.* NIST, Gaithersburg, MD.
6. Rahwan, I. (2019). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology, 21*(1), 5–14.
7. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.
8. Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science, 361*(6404), 751–752.
9. Smith, B., & Shum, H. (2020). Human-centered AI: Building trust and accountability. *Communications of the ACM, 63*(8), 34–36.
10. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21.
11. Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.
12. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
13. Veale, M., & Borgesius, F. Z. (2021). Demystifying the draft EU Artificial Intelligence Act. *Computer Law Review International, 22*(4), 97–112.
14. Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). Ethical and societal implications of

algorithms, data, and artificial intelligence: A roadmap for research. *AI & Society, 34*(3), 1–15.
15. Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2021). Of, for, and by the people: The legal lacuna of synthetic persons. *Artificial Intelligence and Law, 29*(3), 273–291.
16. Zuboff, S. (2019). *The age of surveillance capitalism.* PublicAffairs.
17. OECD. (2019). *OECD principles on artificial intelligence.* OECD Publishing, Paris.
18. United Nations. (2020). *Roadmap for digital cooperation.* United Nations Publications.