

---

# 2025 International Conference on Economics, Management and Education Technology (ICEMET 2025)

Article

## Research on Trust Mechanism Construction and Risk Governance in Public Policy Communication under the Perspective of Human-Machine Collaboration

Yantong Lin <sup>1</sup> and Yulin Zhao <sup>2,\*</sup>

<sup>1</sup> Faculty of Humanities and Arts, Macau University of Science and Technology, Guangzhou, China

<sup>2</sup> School of Journalism and Communication, Northwest University of Political and Law, Xi'an, China

\* Correspondence: Yulin Zhao, School of Journalism and Communication, Northwest University of Political and Law, Xi'an, China

**Abstract:** With the deep integration of artificial intelligence into the field of public policy communication, the paradigm of human-machine collaboration is profoundly reshaping traditional trust building and risk governance mechanisms. This study aims to explore how collaborative interaction between human actors and intelligent systems affects the credibility, transparency, and public acceptance of policy information. Drawing on theories of communication, governance, and science and technology studies, the paper constructs a theoretical framework that integrates trust mechanism design with risk governance strategies in digitally mediated policy processes. Within this framework, it analyzes optimization paths for enhancing information authenticity, improving algorithmic transparency, and broadening meaningful public participation. Particular attention is paid to how algorithmic decision support, automated content generation, and data-driven targeting influence perceptions of fairness, accountability, and reliability in policy communication. The study further discusses potential risks, including information distortion, opacity of algorithmic logic, and the amplification of bias, and proposes corresponding governance tools such as disclosure norms, oversight mechanisms, and participatory feedback channels. The research findings are expected to provide theoretical support and practical implications for enhancing the efficiency, responsiveness, and legitimacy of policy communication, thereby contributing to the modernization of governance capacity in the intelligent era.

**Keywords:** artificial intelligence; public policy; communication; trust; risk governance; transparency

Received: 09 February 2026

Revised: 23 March 2026

Accepted: 07 April 2026

Published: 11 April 2026



**Copyright:** © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

---

### 1. Introduction

#### 1.1. Research Background

With the rapid advancement of artificial intelligence, machine learning, and big data analytics, intelligent technologies have become deeply integrated into the domain of public administration and policy communication [1, 2]. In this evolving human-machine collaborative environment, the dissemination of public policies has shifted away from the traditional linear model of government-to-public information release. Instead, it has transformed into a dynamic and interactive process involving multiple agents, including humans, intelligent systems, and diverse social platforms. This transformation significantly enhances the efficiency, accuracy, and timeliness of information transmission, while simultaneously expanding the avenues for citizen engagement and feedback. Nevertheless, the integration of intelligent technologies also brings forth a variety of potential risks and challenges. For instance, information distortion may occur during algorithmic processing, and inherent biases in machine learning models can lead

to unintended consequences. Additionally, unequal access to intelligent systems may exacerbate existing disparities, while growing public skepticism regarding the authenticity and fairness of digital communication poses further challenges. These issues underscore the need for robust frameworks and strategies to ensure that policy communication in the digital era remains transparent, equitable, and trustworthy.

### *1.2. Research Significance*

Investigating the trust mechanism and risk governance of public policy communication under a human-machine collaborative perspective holds both theoretical and practical significance [3]. From a theoretical standpoint, this research enriches the understanding of intelligent governance by exploring how trust can be cultivated in digital environments and how collaborative frameworks between humans and machines can be optimized. It also introduces new analytical dimensions for examining the transformative role of technology in governance processes, particularly in the context of modern digital ecosystems. From a practical perspective, the study provides actionable insights for improving the credibility and transparency of government information dissemination [3]. By addressing the challenges of building public confidence in policy decisions, it contributes to the development of governance models that effectively balance operational efficiency with principles of fairness and inclusivity. Additionally, the findings can guide policymakers in designing targeted strategies for risk prevention and mitigation, ultimately fostering more resilient and trustworthy systems of policy communication in the era of digital transformation.

## **2. Theoretical Foundation of Public Policy Communication and Human-Machine Collaboration**

### *2.1. Concept and Characteristics of Public Policy Communication*

Public policy communication refers to the systematic process through which government entities, administrative bodies, and affiliated institutions convey policy-related information to the public, social organizations, and other stakeholders using diverse channels such as official documents, media platforms, and digital technologies [4, 5]. This process is not limited to a one-way dissemination of information; it also serves as a vital mechanism for fostering dialogue and building mutual understanding between the government and society. The defining characteristics of public policy communication include authority, publicity, and interactivity. Authority is derived from the legitimacy of government institutions as the primary source of policy creation and dissemination, ensuring that the information provided is authentic and reliable. Publicity emphasizes the openness and accessibility of policy information, aiming to reach a broad spectrum of social groups and enhance transparency in governance [3]. Interactivity highlights the importance of two-way communication between the government and the public, facilitating feedback, consultation, and active participation. This interactive exchange contributes to refining decision-making processes and strengthening trust between the government and its citizens. By integrating these elements, public policy communication becomes a cornerstone for effective governance and societal cohesion, ensuring that policies are both understood and supported by the communities they impact.

### *2.2. Core Features of Human-Machine Collaboration*

Human-machine collaboration focuses on the seamless integration of human cognitive abilities with the computational strengths of machines [6, 7]. This approach moves beyond traditional manual communication methods by utilizing advanced technologies such as machine learning, artificial intelligence, and big data analytics. These systems enable rapid processing of vast amounts of information, facilitating real-time decision-making and enhancing the efficiency of communication processes. Scalability is another critical feature, allowing communication efforts to extend across diverse platforms and reach larger audiences through digital innovations. Furthermore, adaptive

intelligence plays a pivotal role by enabling systems to dynamically adjust strategies based on user behaviors, preferences, and evolving social contexts. This adaptability ensures that communication becomes increasingly personalized and responsive, fostering more effective engagement. By combining these features, human-machine collaboration offers a transformative approach to optimizing communication processes, making them more efficient, scalable, and tailored to the needs of various stakeholders.

### *2.3. Application of Human-Machine Collaboration in Public Policy Communication*

The integration of human-machine collaboration into public policy communication has experienced significant advancements, fundamentally transforming both the content and methods of interaction. Intelligent recommendation systems, for instance, are capable of analyzing the preferences of citizens to deliver tailored policy information to specific groups, thereby enhancing both efficiency and precision in communication [8]. Additionally, AI-powered sentiment analysis tools enable governments to monitor public attitudes toward policies in real time, facilitating timely adjustments and the identification of emerging risks. Interactive Q&A platforms, which leverage technologies such as chatbots and natural language processing, provide immediate responses to public inquiries, reducing communication costs and improving accessibility for diverse populations. Despite these benefits, the adoption of such technologies introduces new governance challenges. Issues such as data privacy concerns, algorithmic bias, excessive dependence on automated systems, and the potential erosion of governmental authority could undermine the fairness and credibility of policy communication if left unaddressed. Consequently, it is imperative to develop robust governance frameworks that strike a balance between leveraging technological advancements and adhering to ethical, legal, and social principles. By doing so, governments can ensure that these innovations contribute positively to public policy communication while safeguarding public trust and equity.

## **3. Optimization Strategies for Trust Mechanism Construction**

### *3.1. Enhancing Information Authenticity and Transparency*

Ensuring the authenticity and transparency of information forms the cornerstone of fostering trust in public policy communication. Within the framework of human-machine collaboration, this involves not only the precise dissemination of policy content but also the implementation of comprehensive verification mechanisms and standardized disclosure practices. Governments should prioritize enhancing consistency across various platforms to prevent distortions that may arise during secondary dissemination. Employing advanced digital tools, such as blockchain technology, can further secure the integrity of data sources by enabling traceability and accountability. Additionally, refining disclosure standards is essential. This includes clearly defining the timeliness, completeness, and accessibility of released information to minimize information asymmetry. By doing so, citizens are better equipped with reliable and transparent references, enabling them to make informed decisions [6, 9]. Such measures collectively contribute to a more robust and trustworthy communication framework in the public policy domain.

### *3.2. Increasing Algorithm Transparency and Interpretability*

As AI-driven recommendation and analysis tools become increasingly integral to policy communication, concerns surrounding algorithm transparency and interpretability have grown significantly. The risks associated with "black box" systems—where the underlying logic behind algorithmic decisions remains obscure—can erode public trust and foster doubts about potential bias or manipulation. To mitigate these challenges, it is essential for governments and technology providers to proactively disclose the foundational principles guiding algorithmic operations, clearly define the criteria used for filtering information, and present results in a manner that is accessible and

comprehensible to the general public. Furthermore, the establishment of third-party audits and robust accountability frameworks can ensure that algorithms function equitably and align with broader governance objectives. These measures not only enhance algorithm accountability but also reinforce their legitimacy, fostering greater confidence in their application within critical decision-making processes.

### *3.3. Promoting Multi-Agent Collaborative Participation*

Effective policy communication in the digital age requires a shift from unilateral government action to a more inclusive and collaborative approach [10]. This involves the active participation of various stakeholders, such as governmental agencies, media organizations, technology companies, and ordinary citizens. By fostering a collaborative governance model, these diverse actors can leverage their unique strengths to achieve common goals. Governments can serve as the primary source of authoritative information, ensuring accuracy and credibility. Media organizations play a crucial role in disseminating information widely and shaping public narratives, while technology companies provide the platforms and tools necessary for efficient communication. Citizens, on the other hand, contribute valuable feedback and grassroots perspectives, ensuring that policies are responsive to societal needs. This multi-agent collaboration not only broadens the scope and reach of policy communication but also promotes inclusiveness and mitigates the risks associated with information monopolies. A well-balanced communication network built on co-construction, co-governance, and mutual trust ultimately strengthens societal cohesion and enhances the effectiveness of policy implementation.

### *3.4. Improving Feedback and Iterative Mechanisms*

In addition to dissemination, effective policy communication relies heavily on the establishment of robust mechanisms for feedback and continuous optimization. Governments should prioritize the development of accessible and user-friendly feedback channels, such as online surveys, interactive digital platforms, and advanced AI-driven sentiment analysis tools, to gather public opinions in real time. These channels must be designed to ensure inclusivity and ease of use, enabling diverse demographic groups to participate actively. However, collecting feedback is only the first step; it is essential that this input is systematically analyzed and integrated into iterative improvements of communication strategies. By doing so, governments can adapt to the dynamic nature of societal needs and technological advancements. This approach fosters a more resilient, inclusive, and responsive communication framework, ultimately strengthening public trust and enhancing the overall effectiveness of governance.

## **4. Risk Governance Strategies in Public Policy Communication**

### *4.1. Governance of Information Distortion Risks*

Information distortion represents a significant and immediate challenge in public policy communication within the context of human-machine collaboration. The rapid proliferation of AI-driven platforms, social media algorithms, and automated content dissemination tools has dramatically increased the speed and reach of information, creating an environment where rumors, misinformation, and manipulated content can spread swiftly and widely. Such distortions pose a serious threat to the credibility and legitimacy of governmental communication, potentially eroding public trust and confidence. To address this issue, it is imperative to establish robust cross-verification and fact-checking systems that leverage the strengths of both machine intelligence and human expertise. For example, AI technologies can be employed to detect irregular patterns in information flow, flag potentially harmful or suspicious content, and trace the origins of misinformation. Meanwhile, human reviewers can apply contextual understanding and ethical judgment to ensure the accuracy of flagged content and minimize false positives. Furthermore, governments should actively collaborate with media organizations,

technology platforms, and independent fact-checking institutions to implement multi-layered verification mechanisms. Clear accountability frameworks must also be introduced to delineate the responsibilities of all stakeholders, ensuring that platforms are held accountable for the dissemination of harmful content. By integrating advanced technological tools with institutional safeguards and legal measures, the risk of information distortion can be significantly mitigated, thereby enhancing the integrity and reliability of public policy communication [11, 12].

#### *4.2. Governance of Algorithm Bias Risks*

Algorithmic bias poses a significant challenge in the realm of intelligent policy communication. Machine learning models, when trained on datasets that are incomplete, unrepresentative, or skewed, can unintentionally reinforce stereotypes, exclude vulnerable groups, or suppress diverse viewpoints. These biases undermine the fairness of information dissemination and erode public trust in the transparency of policy processes. To mitigate these risks, it is essential to institutionalize regular audits of algorithmic systems, with a focus on identifying biases in data collection, model training, and recommendation mechanisms. Effective corrective strategies should include the utilization of diverse and representative datasets, adherence to inclusive design principles, and ongoing monitoring to detect and address emerging biases. Interdisciplinary collaboration is also crucial, bringing together experts from fields such as law, ethics, social sciences, and technology to evaluate algorithmic designs and their societal impacts. This ensures that applications are ethical, equitable, and socially responsible. Additionally, fostering public education about algorithmic principles is vital, enabling citizens to critically assess AI-driven outputs and engage with them in an informed manner. Addressing algorithmic bias is not merely a technical challenge but a governance priority, requiring alignment between technological advancements and societal values to uphold fairness, accountability, and public trust.

#### *4.3. Governance of Privacy and Security Risks*

The widespread application of artificial intelligence in policy communication requires the collection, analysis, and storage of substantial amounts of personal and behavioral data. While this data facilitates personalized content delivery, targeted recommendations, and improved service efficiency, it also introduces significant concerns regarding privacy and security. Individuals may worry about excessive data collection, unauthorized usage, and the possibility of data breaches, which can undermine trust in public institutions. To address these challenges, governments must prioritize the development and enforcement of robust data protection frameworks. These frameworks should include stringent regulations governing data access, usage, storage, and sharing practices. Incorporating privacy-by-design principles into system development ensures that measures such as data minimization, anonymization, encryption, and secure storage are embedded from the beginning. Clear and transparent policies on data usage further enable citizens to understand how their information is processed, fostering greater trust and acceptance. Additionally, international collaboration is becoming increasingly essential to manage cross-border data flows and harmonize regulatory standards, given the interconnected nature of digital governance. By carefully balancing the benefits of intelligent technologies with comprehensive safeguards for privacy and security, governments can strengthen their legitimacy while minimizing systemic risks and enhancing public confidence in digital systems.

#### *4.4. Governance of Social Trust Risks*

Beyond technical and operational challenges, a more profound issue lies in the potential erosion of social trust. Excessive reliance on machines for information dissemination, decision-making support, and public interaction may inadvertently reduce human oversight and weaken the perceived authority of institutions. This could lead to

public uncertainty about whether decisions are being made by accountable policymakers or by opaque algorithms, creating a gap in legitimacy and potentially alienating citizens. To address this, it is essential to establish robust institutional trust frameworks that balance technological efficiency with human responsibility. Human agents must retain the ultimate authority in decision-making processes, with artificial intelligence serving as a supportive tool rather than a replacement. Clear transparency in defining the roles and responsibilities of humans versus machines is critical. Furthermore, the establishment of ethical review boards to oversee the deployment of AI in governance contexts can provide an additional layer of accountability. Promoting digital literacy and civic education is equally important, as it equips citizens with the knowledge to understand, interpret, and critically engage with human-machine interactions. This enables rational trust and informed participation. By embedding human accountability into governance mechanisms, technological advancements can enhance rather than undermine democratic legitimacy, fostering long-term trust and ensuring sustainable communication of policies.

## **5. Evaluation and Prospects of Trust and Risk Governance Mechanisms**

### *5.1. Evaluation Indicators*

The effectiveness of public policy communication in a human-machine collaborative environment can be evaluated using a comprehensive set of multidimensional indicators. Firstly, information credibility is fundamental, as it ensures that the policy content disseminated is accurate, consistent, and resistant to distortion throughout the communication process. This aspect is critical for fostering trust and reliability in the information shared. Secondly, public satisfaction serves as a vital measure, reflecting the degree to which citizens perceive the communication of policies as transparent, timely, and responsive to their concerns and expectations. This indicator emphasizes the emotional and perceptual dimensions of trust, which are essential for maintaining public confidence. Thirdly, the participation rate is an important metric, capturing the extent of citizen involvement in interactive communication platforms, such as online consultations, feedback mechanisms, and deliberative forums. This measure highlights the inclusiveness and accessibility of the communication process, ensuring that diverse voices are heard [13–15]. Lastly, the effectiveness of risk prevention is a critical indicator, assessing the system's capacity to foresee, identify, and address potential risks, including misinformation, algorithmic biases, and breaches of privacy. Collectively, these indicators provide a robust framework for analyzing both the technical efficiency and societal impact of policy communication strategies in a collaborative human-machine context.

### *5.2. Evaluation Methods*

To operationalize these indicators, a combination of qualitative and quantitative methods should be employed. Public opinion surveys remain a traditional yet essential tool, enabling policymakers to collect direct perceptions from citizens regarding authenticity, clarity, and satisfaction. These surveys can be further enhanced through the integration of digital technologies, which streamline data collection processes and improve representativeness by reaching broader demographics. Additionally, big data monitoring utilizes advanced analytics powered by artificial intelligence to track online discourse, analyze sentiment trends, and identify patterns in the dissemination of information across social platforms. This approach provides governments with the capability to conduct real-time assessments and swiftly detect emerging risks or shifts in public sentiment. Experimental simulations of communication scenarios offer controlled environments for testing the effectiveness of various strategies, such as algorithmic recommendations, message framing techniques, or feedback mechanisms, in influencing public reception and fostering trust. By combining surveys, big data analytics, and experimental research, evaluations can achieve a multidimensional perspective, ensuring that findings are robust, evidence-based, and actionable for policy development.

### 5.3. Future Prospects

Looking ahead, the future development of human-machine collaboration in public policy communication is expected to progress along three interconnected dimensions [16–18]. First, the integration of intelligent technologies with governance processes will deepen, transitioning from basic information dissemination to more advanced applications such as predictive governance, tailored citizen engagement, and adaptive policy frameworks. This evolution will enable governments to anticipate societal needs more effectively and respond with precision. Second, transparency will emerge as a cornerstone for maintaining public trust, necessitating ongoing advancements in algorithmic interpretability, comprehensive information sharing, and robust accountability structures. These measures will ensure that technological systems operate in a manner that is both understandable and trustworthy to the public. Third, the adaptability of governance strategies will gain prominence, driven by the need to address swiftly changing societal expectations, technological breakthroughs, and complex global challenges. To achieve this, trust-building mechanisms will require continuous enhancement, focusing on institutional credibility, fairness in algorithmic processes, and stringent privacy protections. Simultaneously, risk management systems must evolve from reactive approaches to proactive strategies, embedding resilience and foresight throughout the communication lifecycle. Ultimately, the trajectory points toward a harmonized model where efficiency, inclusiveness, and legitimacy are balanced, enabling human-machine collaboration to significantly improve the effectiveness and credibility of public policy communication.

### 6. Conclusion

This study systematically analyzed the construction of trust mechanisms and the development of risk governance strategies in public policy communication from the perspective of human-machine collaboration. The findings underscore the importance of establishing a comprehensive trust framework by prioritizing the authenticity of information, ensuring greater transparency in algorithmic processes, and fostering active participation from diverse stakeholders. Effective risk governance must address critical challenges such as mitigating information distortion, reducing algorithmic biases, safeguarding privacy, and strengthening social trust. These efforts are essential to navigating the complexities of the intelligent era. Furthermore, achieving a balance between technological innovation and governance responsibility is pivotal for enhancing the credibility, inclusiveness, and sustainability of public policy communication. Future research should explore adaptive governance models, investigate the ethical implications of emerging technologies, and assess the long-term societal impacts of human-machine collaboration in policy communication frameworks.

### References

1. S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box: Automated decisions and the GDPR," *Harv. JL & Tech.*, vol. 31, p. 841, 2017.
2. T. Mandják and Z. Szántó, "How can economic sociology help business relationship management?," *Journal of Business & Industrial Marketing*, vol. 25, no. 3, pp. 202–208, 2010.
3. D. L. Coleman, A. E. Kazdin, L. A. Miller, J. S. Morrow, and R. Udelsman, "Guidelines for interactions between clinical faculty and the pharmaceutical industry: one medical school's approach," *Academic Medicine*, vol. 81, no. 2, pp. 154–160, 2006.
4. S. Feizi and M. Médard, "A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing," in \*2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)\*, pp. 1048–1054, Sep. 2011.
5. C. S. Representativeness, "National surveys via RDD telephone interviewing versus the internet," *Public Opinion Quarterly*, vol. 73, no. 4, pp. 641–678, 2009.
6. V. Bitzer, J. van Helmsing, A. B. Helmsing, and V. Van Der Linden, "Partnering to facilitate smallholder inclusion in value chains," in *Value Chains, Social Inclusion and Economic Development*, Routledge, pp. 221–246, 2012.
7. W. B. Knox and P. Stone, "Reinforcement learning from human reward: Discounting in episodic tasks," in \*2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication\*, pp. 878–885, Sep. 2012.

8. R. N. Knowles, "Self-Organising Leadership: Transparency and Trust," in *Management Models for Corporate Social Responsibility*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 134–139, 2006.
9. E. A. Stolk, A. De Bont, A. R. Van Halteren, R. J. Bijlmer, and M. J. Poley, "Role of health technology assessment in shaping the benefits package in the Netherlands," *Expert Review of Pharmacoeconomics & Outcomes Research*, vol. 9, no. 1, pp. 85–94, 2009.
10. W. Naudé and N. Dimitri, "The race for an artificial general intelligence: implications for public policy," *AI & Society*, vol. 35, no. 2, pp. 367–379, 2020.
11. D. Valle-Cruz, J. I. Criado, R. Sandoval-Almazán, and E. A. Ruvalcaba-Gomez, "Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation," *Government Information Quarterly*, vol. 37, no. 4, p. 101509, 2020.
12. R. Paul, "Can critical policy studies outsmart AI? Research agenda on artificial intelligence technologies and public policy," *Critical Policy Studies*, vol. 16, no. 4, pp. 497–509, 2022.
13. R. Calo, "Artificial intelligence policy: a primer and roadmap," *UCDL Rev.*, vol. 51, p. 399, 2017.
14. J. M. Sánchez, J. P. Rodríguez, and H. E. Espitia, "Review of artificial intelligence applied in decision-making processes in agricultural public policy," *Processes*, vol. 8, no. 11, p. 1374, 2020.
15. D. Valle-Cruz, E. Alejandro Ruvalcaba-Gomez, R. Sandoval-Almazan, and J. Ignacio Criado, "A review of artificial intelligence in government and its potential from a public policy perspective," in *\*Proceedings of the 20th Annual International Conference on Digital Government Research\**, pp. 91–99, Jun. 2019.
16. S. K. Misra, S. Das, S. Gupta, and S. K. Sharma, "Public policy and regulatory challenges of artificial intelligence (AI)," in *International Working Conference on Transfer and Diffusion of IT*, Cham: Springer International Publishing, pp. 100–111, Dec. 2020.
17. A. Lauterbach, "Artificial intelligence and policy: quo vadis?," *\*Digital Policy, Regulation and Governance\**, vol. 21, no. 3, pp. 238–263, 2019.
18. A. D. Thierer, A. Castillo O'Sullivan, and R. Russell, "Artificial intelligence and public policy," *Mercatus Research Paper*, 2017.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Publisher and/or the editor(s). Publisher and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.