December 2025 | Volume 1 | Issue 3 | Article No. 25411

Journal of Collective Sciences and Sustainability

Research Article | *Open Access* |

Responsible AI in the Age of Generative Media: A Comparative Study of Ethical and Transparent AI Frameworks in India and the USA

Bhavya S. Tripathi,[1,*] Shrawani C. Gaonkar[1] and Anagha Dhavalikar[2]

[1] *Department of Artificial Intelligence and Data Science, Vasantdada Patil Pratishthan's College of Engineering & Visual Arts, Mumbai, Maharashtra, 400022, India*
[2] *Department of Electronics and Computer Science, Vasantdada Patil Pratishthan's College of Engineering & Visual Arts, Mumbai, Maharashtra, 400022, India*
**Email:* bhavyatripathi02@gmail.com (B. S. Tripathi)

Abstract

Advances in generative artificial intelligence, such as Google Gemini, ChatGPT and DALL·E, are opening new possibilities for creativity on digital media but also raising pressing concerns about misinformation, deepfakes and declining trust in media outlets. This paper explores Responsible Artificial Intelligence (RAI) efforts at the policy level in India and the United States, comparing their unique and shared approaches to addressing generative media - where policies are driven toward balancing innovation with transparency, accountability and fairness. India's official AI for All strategy speaks to inclusivity and social development, though concrete enforcement mechanisms and safeguards against generative media misuse are currently lacking. The US relies on the National Institute of Standards and Technology (NIST) AI Risk Management Framework, emphasizing risk assessment, technical robustness and accountability but lacks a robust regulatory mechanism that aligns its varying state and sector-specific initiatives. Alongside policy analysis, we designed and ran a pilot survey of college students and working professionals through September 2025 to capture awareness, trust and concerns related to Responsible AI. At a preliminary level, our findings showed low relative exposure to national policy frameworks on RAI but high expressed concern related to potential misuse of generative AI around deepfakes, manipulated images and the lack of checks on content authenticity. Respondents expressed particularly high endorsement for content labels that would mandate labeling artificial intelligence generated media. We conclude that India and the USA display parallel and diverging paths on RAI for generative media but both experiences are marked by a gap between policy aspirations and public understanding. Moving ahead, we see a need for greater policy clarity, cross-border coordination and public outreach to foster transparency and responsibility in adoption of AI for media.

***Keywords*:** Generative artificial intelligence; Responsible AI; Ethical AI adoption; AI transparency and accountability; Policy frameworks: India and USA; Deepfakes and media trust.

Received: 13 November 2025; Revised: 25 December 2025; Accepted: 25 December 2025; Published Online: 27 December 2025.

1. Introduction

Generative AI is transforming how machines create content by producing text, visuals, and videos at scale.[1] Systems such as ChatGPT, Gemini, and DALL·E now shape a significant share of digital output; approximately 12% of false online material includes machine-generated media.[2] While access to creative tools has expanded, trust in information has weakened due to growing authenticity risks

DOI: https://doi.org/10.64189/css.25411

and ethical concerns.[3,4] Ethical oversight frameworks based on fairness, accountability, transparency, and explainability (FATE) aim to align technological progress with societal values.[5] Organizations such as UNESCO and IEEE advocate people-centric approaches to AI;[1] however, country-level implementations vary. India's "AI for All" initiative (launched in 2018) emphasizes inclusivity and accessibility,[6] whereas the United States' National Institute of Standards and Technology (NIST) AI Risk Management Framework 1.0 (2023) prioritizes accountability and structured risk management.[7] Despite these efforts, public awareness of such frameworks remains limited, and practical trust in AI systems is correspondingly low. Research examining public perceptions of responsible AI use in contexts such as synthetic images and voice cloning remains scarce. This study combines policy analysis with findings from a survey of 40 individuals in India, including students and working professionals, to explore awareness, trust, and value perceptions related to generative AI. It examines alignment between policy objectives and public understanding. Unlike prior studies that focus primarily on theoretical ethics, this work employs empirical evidence to identify gaps between policy intentions and user experience and proposes approaches to strengthen transparency and responsibility in AI systems across both India and the United States.

2. Literature review

Responsible Artificial Intelligence (AI) rests on fairness, accountability, transparency, and explainability,[5] shaping how emerging technologies are guided ethically. These principles aim to protect individual rights, reduce discriminatory outcomes, and maintain societal trust in automated systems. Global initiatives, including the Organisation for Economic Co-operation and Development (OECD) AI Principles (2019), UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021), and the European Union's Artificial Intelligence Act (2024), promote a human-centric approach to AI governance that prioritizes fundamental rights and social well-being. However, the rapid growth of generative AI, which can produce realistic synthetic text, images, and speech, poses new ethical challenges by obscuring content authenticity and source credibility. To address these risks, technical mechanisms such as the Coalition for Content Provenance and Authenticity (C2PA) framework have been proposed to embed verifiable provenance metadata into AI-generated content.[8,9]

2.1 Global ethical frameworks

More than 40 published AI ethics guidelines emphasize FATE principles, namely fairness, accountability, transparency, and explainability, alongside human rights protection.[1,3] Among these, five frameworks are particularly influential at the global level: the OECD AI Principles (2019), UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021), the European Union AI Act (2024), IEEE's Ethically Aligned Design (2021), and the NIST AI Risk Management Framework (AI RMF) (2023). While all five frameworks align broadly with FATE principles, they differ in enforcement strength and operational focus. Some emphasize legally binding compliance mechanisms, whereas others function primarily as voluntary guidelines or technical standards. Recent governance-focused surveys note that, despite broad ethical consensus, gaps remain in translating high-level principles into enforceable and measurable practices, particularly for fast-evolving generative AI systems.[1,10]

These frameworks collectively establish ethical baselines for AI governance, yet their scope, legal authority, and implementation mechanisms vary considerably across regions.

2.2 National frameworks: India vs. United States

India's *AI for All* strategy (2018), together with the NITI Frontier Tech Hub initiative (2024), emphasizes inclusivity, skill development, and socially beneficial AI deployment.[6,13] However, these initiatives function primarily as advisory frameworks and lack binding regulatory authority. In contrast, the United States' NIST AI RMF 1.0 (2023) focuses on structured risk identification, documentation, and accountability across the AI lifecycle.[7] Although comprehensive in technical guidance, its adoption remains voluntary and fragmented across federal agencies and private sectors. Comparative policy studies indicate that while both countries share core ethical values, their governance models differ substantially in enforcement mechanisms and institutional coordination.[1,3] Recent cross-national analyses further highlight that generative AI governance remains uneven, with limited standardized mechanisms for content verification and accountability across jurisdictions.[4,10]

Table 1: Comparative overview of major global responsible AI frameworks.

Framework	Year	Primary Focus	Key Limitation
IEEE Ethically Aligned Design[1]	2021	Engineering ethics	Complex implementation
NIST AI RMF 1.0[7]	2023	Technical accountability	Fragmented U.S. adoption
EU AI Act[10]	2024	Legal compliance tiers	High burden for SMEs
OECD AI Principles[11]	2019	Voluntary global guidelines	Non-binding; lack enforcement
UNESCO Ethics Recommendation[12]	2021	Socio-cultural governance	Limited auditability

2.3 Research gap

Despite extensive policy and ethical guideline development, three key gaps persist in the existing literature. First, an empirical gap remains, as many studies focus predominantly on conceptual or policy-level analysis without measuring public awareness or perceptions of Responsible AI in practice.[1,4] Second, although generative media risks such as deepfakes and manipulated content are widely acknowledged, research has largely emphasized content creation rather than user-level validation, detection, and trust mechanisms.[8,14] Third, cross-cultural and cross-national comparisons remain limited, particularly studies that connect national AI governance frameworks with public understanding and trust outcomes.[15] Existing studies focus primarily on policy design rather than public awareness and trust in generative media, leaving a disconnect between regulatory intent and user experience. This research addresses these gaps by combining policy analysis with survey-based empirical data to examine awareness, trust, and ethical expectations surrounding generative AI.

3. Methodology

This study uses both quantitative and qualitative methods, connecting survey responses with official policy documents. This mixed-methods approach enables comparison across governance systems while exploring individual perceptions of transparency and ethics.

3.1 Policy comparison

Key policy references include India's AI for All strategy (2018),[6] the NITI Frontier Tech Hub initiative (2024),[13] and the U.S. NIST AI RMF 1.0 (2023).[7] Policy documents were qualitatively coded using a deductive framework based on five Responsible AI dimensions: fairness, accountability, transparency, inclusivity, and enforceability. Each policy was independently assessed against these dimensions and compared for scope, legal authority, and governance mechanisms.

3.2 Survey design

A structured online survey using Google Forms was conducted in September 2025, involving 40 voluntary respondents. The survey contained 12 items divided into four thematic sections:

1. Awareness & Exposure – familiarity with AI tools and RAI concepts;
2. Reliability and clarity - perceptions of trustworthiness and comprehensibility;
3. Ethical issues - concerns about bias, data misuse, and synthetic media;
4. Policy Focus – opinions on labeling, monitoring, and oversight policies.

Students and early-career professionals were selected due to their high exposure to generative AI tools and their emerging role as primary adopters. Convenience sampling was employed due to the exploratory nature of the study. Sample questions included:

- "Have you heard of the term Responsible AI before this survey?" (Yes/No)
- "How important is transparency when using AI tools?" (1–5 scale)

Participation was anonymous, restricted to adults (18+), conducted in accordance with institutional ethical standards.

3.3 Sampling and demographics

Respondents represented both students (85%) and professionals/developers (15%), all based in India. Given the exploratory sample size (n = 40), the findings are descriptive and intended to identify patterns rather than support inferential generalizations.

Table 2: Comparative overview of selected responsible ai policies.

Policy	Year	Country	Fairness	Accountability	Transparency	Inclusivity	Enforceability	Remarks
AI for All	2018	India	✓	✓	✓	Partial	Advisory	Focuses on access and skill development; non-binding
NITI Frontier Tech Hub	2024	India	✓	✓	✓	✓	Advisory	Emphasizes responsible deployment and collaboration; non-binding
NIST AI Risk Management Framework 1.0	2023	USA	✓	✓	✓	Partial	Voluntary	Provides structured risk management guidance; adoption fragmented across agencies

Table 3: Demographic profile of respondents (n = 40).

Category	Sub-group	%
Country	India	100
Primary Role	Students	85
	Professionals/Developers	15
AI Tool Usage	Daily	75
	Few times/week	22.5
	Rarely	2.5
Survey Reliability	Cronbach's α = 0.84	High consistency

3.4 Analytical procedure

Descriptive statistics and cross-tabulation analyses were conducted using Google Forms outputs and manual aggregation. Relationships between awareness levels, trust perceptions, and policy preferences were examined through percentage comparisons and thematic interpretation.

Expanded descriptive cross-tabulation analyses were included to illustrate patterns such as:

- Awareness vs. Trust
- Awareness vs. Support for Labeling

Due to the exploratory sample size (n = 40), inferential statistical tests, such as correlation coefficients or p-values, were not applied. This approach allows identification of trends and patterns across respondents while linking findings to the policy frameworks reviewed.

4. Results and analysis

Findings from the public survey (n = 40) are organized around four themes: awareness, trust, policy orientation, and cross-national perception.

4.1 Awareness of responsible AI and generative tools

The results indicate a high level of engagement with AI tools, with 75% of respondents reporting daily usage, 22.5% using AI tools a few times per week, and 2.5% reporting infrequent use. Despite widespread usage, awareness of Responsible AI concepts remained limited. Only 30% of respondents reported a clear understanding of the term, while 32.5% indicated partial familiarity. Notably, 37.5% encountered the concept of Responsible AI for the first time through this survey. This gap between usage intensity and conceptual understanding highlights the need for improved awareness initiatives alongside expanding AI adoption.

4.2 Trust in AI-generated media and transparency expectations

Trust in AI-generated content was moderate to low among respondents. Only 25% reported high levels of trust, while a substantial majority expressed uncertainty or skepticism, primarily due to concerns regarding misinformation and synthetic media manipulation. Transparency emerged as a critical factor influencing trust, with more than half of respondents rating it as "very important" or higher when using AI systems. Respondents who reported familiarity with verification mechanisms, such as digital provenance tags or content markers, demonstrated comparatively higher trust in AI-generated outputs than those without such awareness.

4.3 Policy support and content labeling

Strong support was observed for regulatory interventions addressing AI-generated content. A majority of respondents (65%) favored mandatory labeling of AI-generated material, while 27.5% supported labeling in critical or high-risk contexts. Only a small minority (2.5%) opposed labeling altogether. Similarly, 62.5% of respondents expressed support for official oversight mechanisms, indicating broad public approval for governance measures emphasizing transparency and accountability.

Cross-tab Observations:

- Awareness vs. Trust: Respondents familiar with Responsible AI concepts exhibited higher levels of trust in AI-generated content compared to respondents without prior awareness.

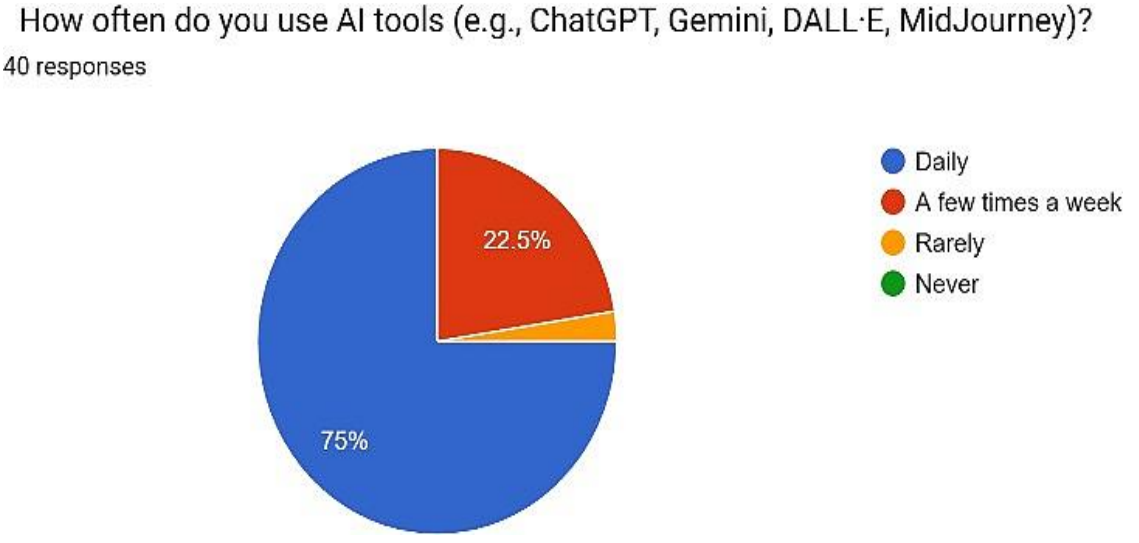

Fig. 1: Respondents' frequency of using AI tools (ChatGPT, Gemini, DALL·E, MidJourney) (n = 40).

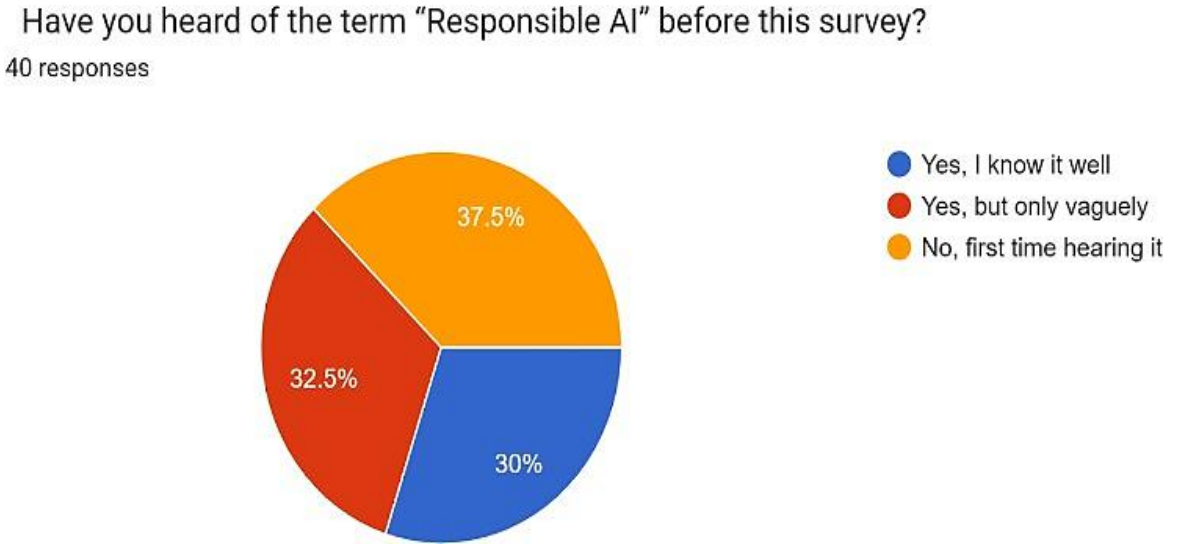

Fig. 2: Prior awareness of the concept of Responsible AI among respondents (n = 40).

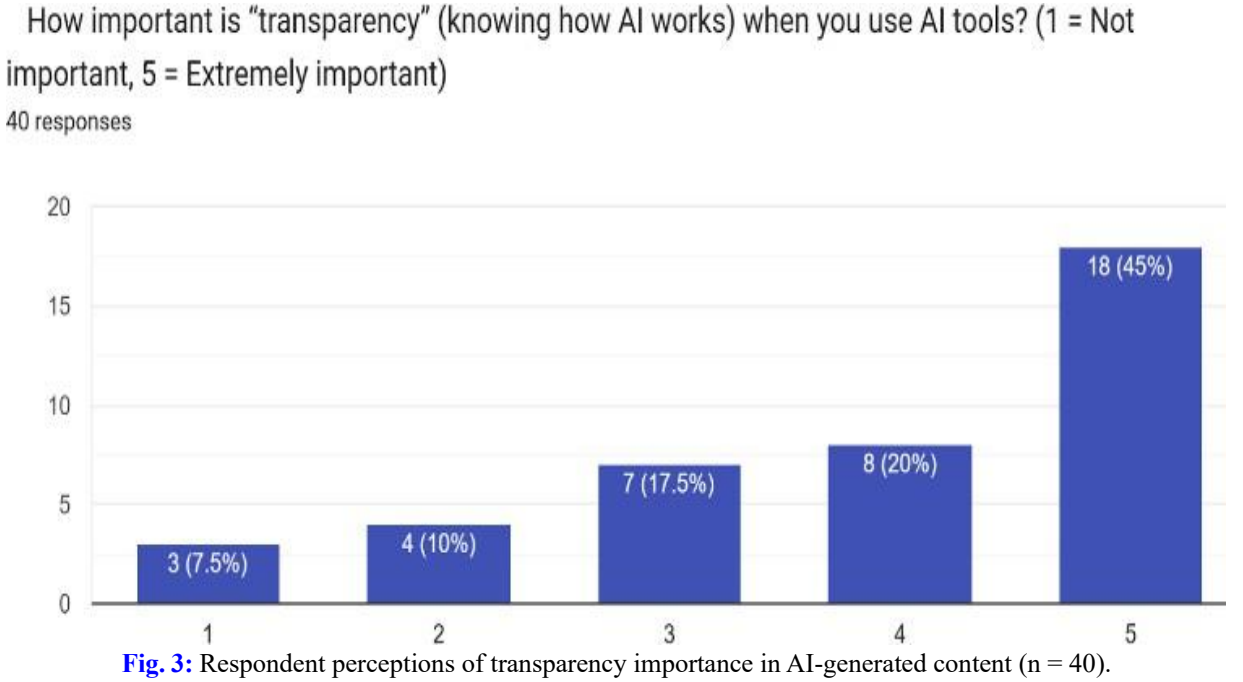

Fig. 3: Respondent perceptions of transparency importance in AI-generated content (n = 40).

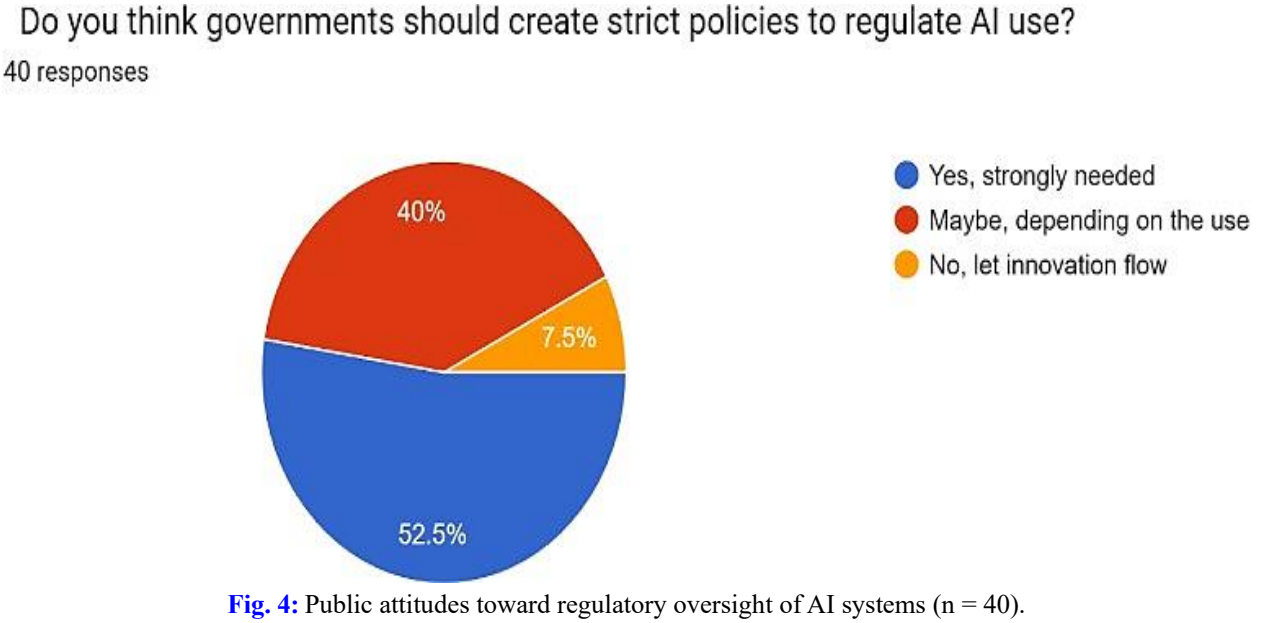

Fig. 4: Public attitudes toward regulatory oversight of AI systems (n = 40).

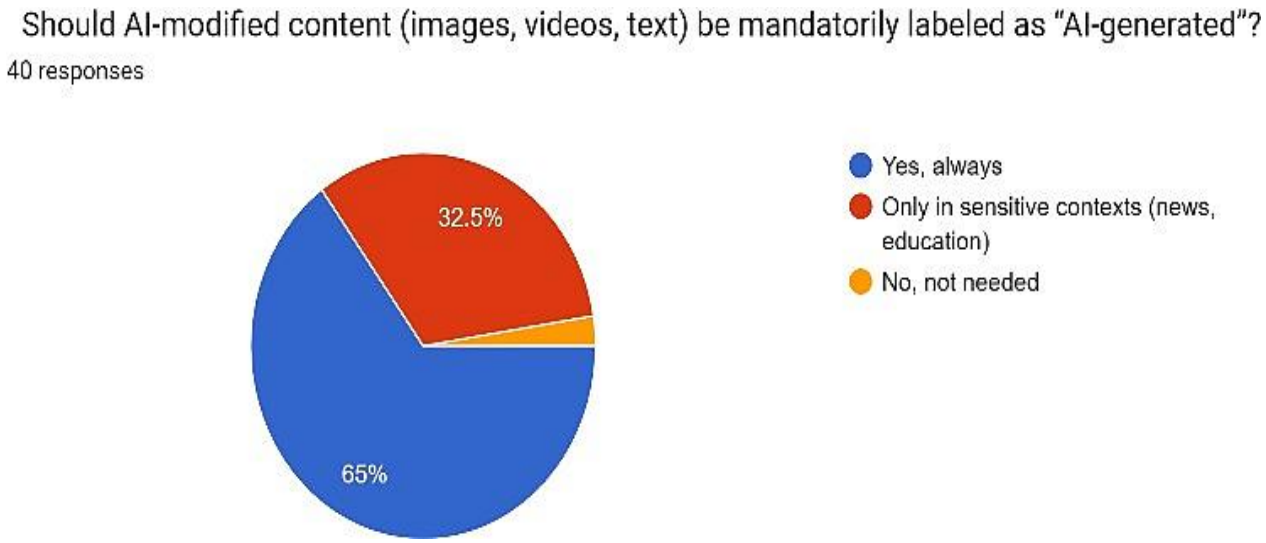

Fig. 5: Respondents' support for mandatory labeling of AI-generated content (n = 40).

- Awareness vs. Support for Labeling: Respondents with prior knowledge of Responsible AI principles were more likely to support mandatory labeling of AI-generated content than those encountering the concept for the first time.

4.4 Perceptions of India–U.S. ethical alignment

With respect to cross-national governance, 45% of respondents favored a shared ethical framework between India and the United States, while 35% preferred initially independent national approaches. The remaining respondents expressed no definitive preference. Qualitative responses suggested that U.S. frameworks were perceived as structured and enforcement-oriented, whereas Indian approaches were viewed as inclusive and aspirational. These perceptions indicate complementary strengths across the two governance models rather than direct opposition.

5. Discussion

Results are assessed in relation to the goals of the research - by contrasting country-level systems, measuring public understanding and confidence, also sketching approaches for ethical AI use.

5.1 Governance comparison

India's AI for All along with NITI Tech Hub focus on access and basic skills yet enforcement stays limited.[6,12] The U.S. NIST AI RMF 1.0[7] provides clearer records along with better responsibility tracking but faces scattered supervision. That aligns with Kumari,[15] whose work highlights ongoing variation worldwide when it comes to actual enforcement.

5.2 Knowing and believing

Even though 75% used AI every day, just 30% grasped RAI ideas - similar to findings by Jobin *et al.*,[1] showing poor public awareness of ethics. Confidence in generated content stayed weak at 25%, which aligns with transparency theory.[14] Understanding tools like C2PA or digital marks tied to stronger trust; recognition led to greater belief in AI results.[9]

5.3 Moving toward clear adoption

Respondents supported content labels (65%) along with government supervision (62.5%), showing clear support for responsibility measures. A mixed strategy - using tech protections together with public awareness efforts - aligns with international standards such as human oversight systems, origin tracking data, and openness checks.[4,8,9]

5.4 New ethical issues

Generative AI boosts creative potential; however, it also increases dangers like fake media and data abuse - key issues highlighted in the study and often discussed in AI & Society.[8] For responsible use moving forward, rules alone aren't enough - stronger public understanding of tech, moral awareness, and systems that track digital origins must grow at the same pace.

6. Policy implications and strategic recommendations

6.1 Establishing practical and enforceable guidelines and regulations

Since 65% of respondents supported mandatory labeling of AI-generated content, policymakers in India and the United States should prioritize the adoption of standardized provenance and content authentication mechanisms, such as the Coalition for Content Provenance and Authenticity (C2PA) framework.[8,9] The strong public preference for labeling indicates a clear demand for verifiable indicators that distinguish synthetic content from human-generated media, particularly in high-risk information environments.

6.2 Raising public knowledge while improving skills

Although most respondents reported frequent use of generative AI tools, only 30% demonstrated clear understanding of Responsible AI concepts, highlighting a significant awareness gap. To address this, both countries

should strengthen AI literacy initiatives by integrating Responsible AI, ethics, and media verification concepts into school and university curricula.[6] These efforts should be supported through partnerships between academic institutions, industry stakeholders, and public outreach programs to ensure broader societal understanding of ethical AI use.

6.3 Joint India-U.S. governance efforts

Survey findings show that 45% of respondents favored shared ethical oversight between India and the United States, indicating public openness to cross-national cooperation. Building on this support, a bilateral Responsible AI working group could facilitate coordination on transparency standards, content verification practices, and watermarking approaches. Such collaboration would align with UNESCO's 2021 ethical framework.[12] while connecting the operational guidance of the NIST AI Risk Management Framework with India's inclusive Responsible AI initiatives.

6.4 Economic or social aspects

While the adoption of transparency and verification mechanisms may introduce short-term compliance and implementation costs for digital platforms, such measures have the potential to generate long-term societal benefits. Survey responses indicate strong public support for transparency-oriented governance, suggesting that verified content labeling and provenance systems can enhance user trust and confidence in digital information environments. Over time, increased transparency may contribute to more reliable information sharing, greater user engagement, and sustained public trust in AI-enabled media systems.

7. Limitation

This study is limited by its small sample size (n = 40) and its geographic concentration within urban regions of India. As a result, the findings may not be generalizable to broader or cross-national populations. Future research should incorporate larger, more diverse samples, cross-country comparisons, longitudinal designs, and inferential statistical analyses to validate observed trends, providing stronger evidence for the patterns identified here. Moving ahead, studies should blend technological, legal, and behavioral insights to build AI governance tools that remain people-centered-where innovation meets responsibility.

8. Conclusion

This research examined India's "AI for All" alongside the U.S. NIST AI RMF 1.0, using combined methods to assess public knowledge and confidence in generative-AI ethics. Results reveal a disconnect between usage and understanding: while three-quarters engage with AI every day, just 30 percent grasp Responsible AI concepts. Confidence is weak-nearly two-thirds are unsure or skeptical, but those familiar with openness measures tend to feel more assured. Support for tagging content is high (65%), while backing for global collaboration stands at 45%, pointing to a need for clearer, stronger oversight. To close this gap, two paths must align-one enforcing origin rules like C2PA; another expanding public understanding of AI ethics. In time, ethical AI will not rely on laws alone, but will also grow from everyday digital habits shaped by shared values.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable

Use of artificial intelligence (AI)-assisted technology for manuscript preparation

The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using AI.

References

[1] A. Jobin, M. Ienca, E. Vayena, The global landscape of AI ethics guidelines, *Nature Machine Intelligence*, 2019, **1**, 389–399, doi: 10.1038/s42256-019-0088-2.

[2] PwC, Global AI Adoption and Risk Report, PwC, 2025.

[3] C. Cath, Governing artificial intelligence: Ethical, legal, and technical challenges, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2018, **376**, 20180080, doi: 10.1098/rsta.2018.0080.

[4] M. K. Pathan, A. Shah, Ethical governance of generative AI: A systematic review, *Journal of Ethics in Emerging Technologies*, 2024, 4(1).

[5] P. Siddhapura, V. Patel, Developing ethical AI frameworks: a comparative analysis of global standards and practices, *TIJER – International Research Journal*, 2024, 11, a989–a996.

[6] NITI Aayog, National Strategy for Artificial Intelligence – AI for All, Government of India, 2018.

[7] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0), U.S. Department of Commerce, 2023.

[8] S. Lucas, R. M. Heinitz, S. J. Becker, J. E. Charton, Developing a framework for addressing ethical challenges in generative AI, *Journal of Information Technology Case and Application Research*, 2025, 1–15, doi: 10.1080/15228053.2025.2558443.

[9] Coalition for Content Provenance and Authenticity (C2PA), C2PA Technical Specification Version 2.0, *C2PA*, 2024.

[10] European Union, Artificial Intelligence Act (AI Act), Regulation (EU) 2024/1689, *European Union*, 2024.

[11] OECD, Recommendation on Artificial Intelligence (AI Principles), *OECD Publishing*, 2019.

[12] UNESCO, Recommendation on the Ethics of Artificial

Intelligence, *UNESCO*, 2021.
[13] NITI Aayog, NITI Frontier Tech Hub: AI for Viksit Bharat, Government of India, 2024.
[14] P. Radanliev, AI Ethics: Integrating Transparency, Fairness, and Privacy in AI Development. *Applied Artificial Intelligence*, 2025, 39, doi: 10.1080/08839514.2025.2463722.
[15] P. Kumari, Legal frameworks for AI regulation: a comparative study, *Advances in Consumer Research*, 2025, **2**, 216-224.