ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

A STUDY ON MORAL RESPONSIBILITY OF AI DEVELOPERS

** Asst. Prof. Aravind Nair & ** Asst. Prof. Ankita Patil*

Abstract:

AI tools—ChatGPT, learning apps, resume screeners—are now part of almost every student's routine here. They make studying faster and sometimes easier, but they also bring real problems when there's no proper ethical oversight. Recent reports and surveys from 2025–2026 show clear issues: many students use AI to complete assignments and exams with almost no original work, which hurts genuine learning and critical thinking; a large number (up to 88% in some youth surveys) turn to AI chatbots when they feel stressed or anxious, but this can increase feelings of isolation, shorter attention spans, and mental fatigue; entry-level jobs in IT and tech have already dropped by 20–25% because of automation, leaving fresh graduates worried about their future; AI systems sometimes carry forward biases related to caste, gender, or background, making opportunities even harder for students from rural areas or marginalized communities; and deepfakes along with AI-generated misinformation are spreading fast on social media, confusing young people and damaging trust. With over 40 million students in higher education across India, these challenges affect a huge number of young adults at a very important stage of life. Moral responsibility—meaning developers, companies, colleges, and the government must build AI that is transparent, fair, safe, and does no harm—is not optional. Without it, AI risks making existing pressures worse: weaker skills, poorer mental health, fewer jobs, unfair treatment, and less reliable information. It is important to take ethics seriously right now—through better design, regular bias checks (especially for Indian realities like caste), clear rules, and shared accountability—AI can actually help students learn better, find good careers, and grow confidently instead of holding them back.

Keywords: *AI, Morality, People*

Introduction:

The fast growth of Artificial Intelligence (AI) is changing education, jobs, and daily life in India. Our country is one of the biggest users of AI tools, with millions of young people relying on them for studies, career advice, and even emotional support. According to various reports, India's AI sector is expanding quickly, and tools like ChatGPT are now common in colleges. However, this rise brings serious ethical questions: Who is responsible when AI causes harm? Moral responsibility in AI means that developers, companies, and policymakers must build systems that are fair, transparent, safe, and do not hurt people—especially vulnerable groups like students.

Review Literature:

Tigard D.W. 2021 clarifies the concept of responsible AI by distinguishing three senses of responsibility: causal (who caused an outcome), role-based (duties in development/deployment), and moral (blameworthiness or praise). Drawing from analytic ethics, Tigard argues that while AI systems lack full moral agency, humans involved retain moral responsibility. He advocates for a shared appreciation of these distinctions to advance meaningful ethical discourse and accountability in AI without anthropomorphizing machines.[1]

Hedlund, M., & Persson, E.(2024) examines the forward-looking responsibility of AI experts (as a

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

professional collective) to guide AI development toward desirable societal outcomes. Drawing on ethical and professional responsibility theories, the authors argue that experts bear partial moral duty due to their specialized knowledge and influence, even amid uncertainty and distributed agency. They conclude that AI experts have a collective obligation to promote beneficial directions, mitigate risks, and advocate for ethical practices, without shifting full blame to individuals.[2]

Griffin, T. A.(2024) is the first to centre AI developers in discussions of ethical responsibility. Through semi-structured interviews with 40 developers worldwide, the authors explore their ethical agency—the capacity and willingness to make morally informed decisions in AI work. Findings reveal variable ethical agency: developers often have authority to raise concerns or influence outcomes but face constraints from organizational structures, hierarchies, and priorities.[4]

Green, B. P., & Welie, J. V.(2025) argues that AI developers possess moral agency as professionals, drawing from Arthur Isak Applbaum's adversary ethics and Alasdair MacIntyre's virtue-based framework. It critiques common "excuses" (e.g., distributed responsibility, organizational constraints) that diminish developers' accountability. The authors assert that developers' professional identity demands moral engagement—recognizing their capacity for ethical judgment and duty to resist harmful practices—rather than deferring blame.[3]

Objectives:

- To examine the ethical challenges and negative consequences that arise from the use of AI tools.
- To explore the concept of moral responsibility in AI.
- To assess the specific impacts of AI on Indian college students.
- To advocate for the urgent need for responsible AI practices in education and society.

Theoretical view:

Challenges faced by lack of Morality in AI :

In India, over 40 million students are in higher education, and most are active online. They use AI for assignments, learning apps, resume screening, and more. Without strong ethical guidelines, AI can lead to real problems for these young adults. Recent data from Indian and global sources shows issues like cheating in exams, mental health strain, job fears, bias based on caste or gender, and exposure to deepfakes and misinformation.

Academic Integrity:

One of the biggest issues observed in every semester is how generative AI makes it far too easy for students to submit work without putting in genuine effort. In Indian universities—including many here in Mumbai and across the country—professors are seeing a clear increase in AI-assisted plagiarism. Students can now generate full essays, code solutions, or exam answers almost instantly, which skips the essential process of research, analysis, and original thought. Recent studies and articles from 2025 highlight that while detection tools are improving, actual cases caught and penalized remain low (sometimes as few as one in hundreds), partly because spotting AI content reliably is still difficult. To address this, many institutions have updated their academic handbooks for 2024–2025 and beyond, adding clear rules on AI disclosure, acceptable use, and consequences like grade penalties or even suspension. These changes help, but the bigger concern is long-term: when students rely heavily on AI, their critical thinking, problem-solving, and creative skills don't develop properly. Learning becomes surface-level, and degrees start to lose some of their real-world value—graduates may look strong on paper but struggle when they need to think independently in jobs or further studies.

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

Mental Health:

Mental health is another area that has worrying effect. Students already deal with intense pressure—tough exams, family expectations, financial stress, and the non-stop pace of life in a city like Mumbai. A 2025 survey by Youth Ki Awaaz and YLAC found that up to 88% of Indian students (especially teenagers and early college age groups) turn to AI tools like ChatGPT when they're feeling stressed, anxious, or low. These chatbots give instant, judgment-free responses, which feels supportive at first. But over time, this reliance can lead to cognitive laziness (less mental effort in thinking things through), shorter attention spans from always getting quick answers, and growing social isolation because real conversations with friends or family decrease. Research on Indian college students connects heavy AI use to increased mental fatigue, poorer concentration during studies, higher anxiety, and even more depressive feelings—issues that worsened during and after the pandemic with all the online AI-supported classes adding to screen overload. Most colleges still don't have enough counsellors or accessible mental health services, so depending on AI for emotional support can create a false sense of help.

Job Opportunities:

Job prospects are causing a lot of anxiety among final-year students. India produces millions of graduates every year, especially in engineering, IT, and tech-related fields, with the hope of good careers ahead. But recent 2025 reports from EY and industry analyses show that entry-level tech hiring has dropped by 20–25% as companies use AI to automate routine and outsourced tasks. This is shifting the workforce toward a "diamond-shaped" structure—fewer basic roles for fresher's and more demand for experienced or highly specialized people. Surveys of final-year students reveal that nearly half are genuinely worried that AI will reduce opportunities in their fields, leading to stress, lower confidence, and demotivation during placement season. While AI will eventually create new jobs in areas like AI development, ethics, and advanced management, the current transition is tough—leaving many young graduates facing unemployment, underemployment, or the pressure to up skill rapidly in a very competitive environment.

Discrimination and Bias:

Bias in AI makes these inequalities worse. In a country as diverse as India—with caste, gender, regional, and language differences—AI systems trained mostly on global or skewed data often carry forward unfair patterns. A 2025 investigation by MIT Technology Review found that OpenAI tools like ChatGPT and Sora (popular in India, their second-largest market) reproduce caste stereotypes, generating derogatory or biased content tied to caste identities. This bias shows up in real ways: in admission algorithms, scholarship tools, job resume screeners, or even educational content generators, disadvantaging students from rural areas, lower castes, non-English-medium backgrounds, or underrepresented groups.

Unethical AI Tools:

Unethical tools like deepfakes and AI-generated misinformation are damaging trust and safety. Reports from 2025 show a rise in synthetic media during elections and public events, with realistic fake videos, audio, or images spreading quickly on social platforms. College students, who spend hours on Instagram, WhatsApp, and YouTube, often find it hard to spot these fakes, leading to confusion, divided opinions, cyberbullying (like morphed images used to harass), and mistaken beliefs. This loss of trust in information creates emotional stress, paranoia about what's real online, and reduced engagement in important discussions.

Research Methodology:

This research is entirely based on primary and secondary data. The primary data is collected from the students of Chetana's Institute of Management &

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

Research,Bandra East. The Students were from age group 20 -25 and were pursuing their Masters in MBA. The sample size for the present research consists of 79 primary data respondents as specified by the researcher. The selected sample size is considered adequate to include respondents from the population and to draw meaningful conclusions. Efforts were made to include respondents from diverse backgrounds to enhance the reliability and validity of the findings. The secondary data is collected from various articles, blog and other web sources.

Sampling Technique:

The study employs a non-probability sampling technique, specifically convenience sampling. Under this method, respondents were selected based on their availability and willingness to participate in the survey. Convenience sampling is widely used in social science research due to its practicality and ease of data collection. This technique is suitable for the present study as it allows the researcher to gather data efficiently within the given time frame while maintaining sufficient representation of the target population.

Data Collection Tools:

The study is based on primary data, which was collected using a structured questionnaire administered through Google Forms. The questionnaire was designed to capture respondents' opinions, perceptions, and experiences related to the research variables. Google Forms was chosen for data collection due to its cost-effectiveness, ease of distribution, and ability to reach respondents quickly. The questionnaire consisted mainly of closed-ended quantitative questions which facilitated analysis and reduced manual errors in data entry. The online mode also ensured convenience for respondents.

Limitations:

The findings are based on a limited sample size and may not be fully generalizable to the entire population. The use of convenience sampling and self-reported questionnaires may lead to respondent bias, such as responses influenced by personal opinions or social desirability. Also the secondary data is from the sources that has not analysed the current scenario and perception of people on using AI tools now.

Data Analysis and Interpretation:

Do you think AI developers should be held morally responsible if their tools (like ChatGPT) are widely used by students to cheat on assignments?

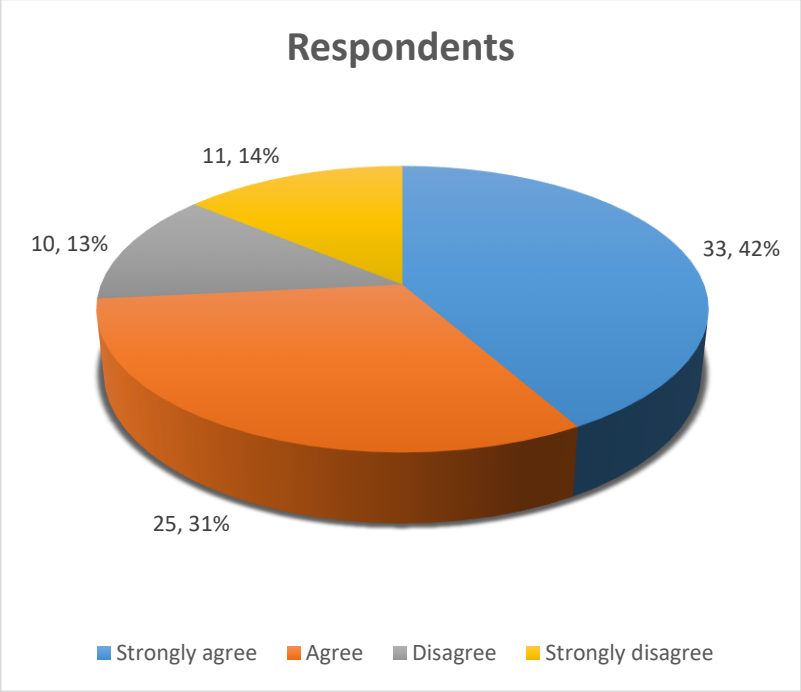
Particulars	Respondents
Strongly agree	33
Agree	25
Disagree	10
Strongly disagree	11

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

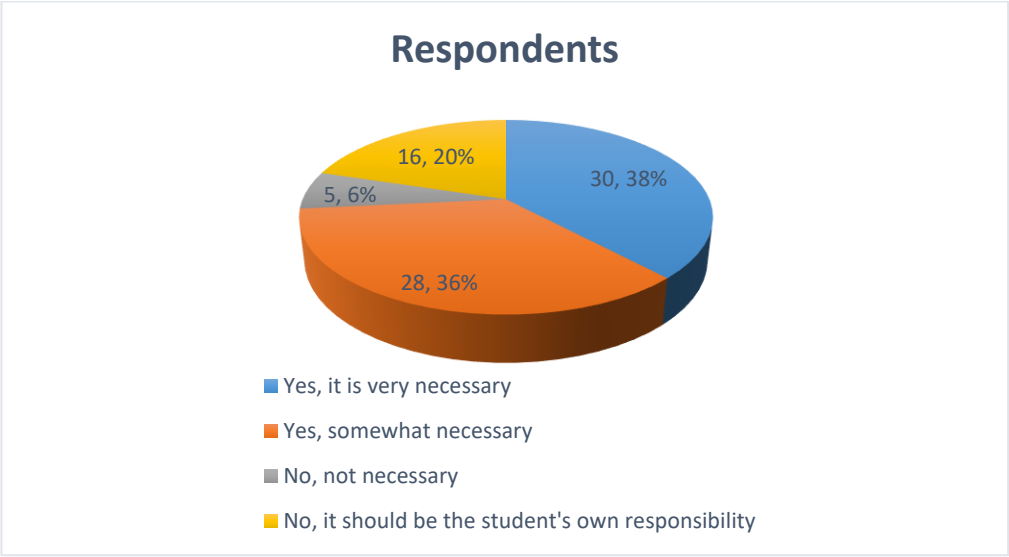
OPEN ACCESS Original Research Article



The above chart shows that majority of students 33 (42%) believe AI developers should be held morally responsible if their tools (like ChatGPT) are widely used by students to cheat on assignments.

In your opinion, should AI companies be required to build in features that prevent students from over-relying on AI and losing critical thinking skills?

Particulars	Respondents
Yes, it is very necessary	30
Yes, somewhat necessary	28
No, not necessary	5
No, it should be the student's own responsibility	16



ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

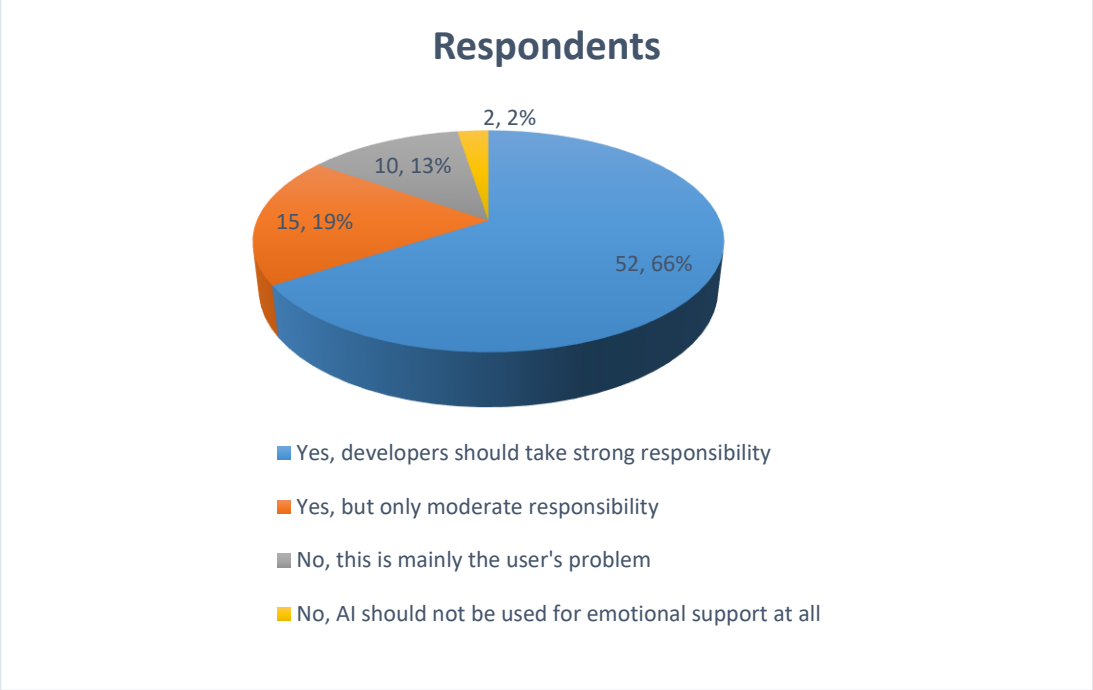
Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

The above chart shows that majority of students 30 (38%) believe AI companies should be required to build in features that prevent students from over-relying on AI and losing critical thinking skills.

Do you feel that AI developers have a moral duty to reduce mental health risks (like anxiety or isolation) caused by heavy use of AI chatbots for emotional support?

Particulars	Respondents
Yes, developers should take strong responsibility	52
Yes, but only moderate responsibility	15
No, this is mainly the user's problem	10
No, AI should not be used for emotional support at all	2



The above chart shows that majority of students 52 (66%) believe AI developers should have a moral duty to reduce mental health risks (like anxiety or isolation) caused by heavy use of AI chatbots for emotional support.

Should AI developers be morally accountable if their tools automate entry-level jobs and make it harder for fresh graduates like us to find work?

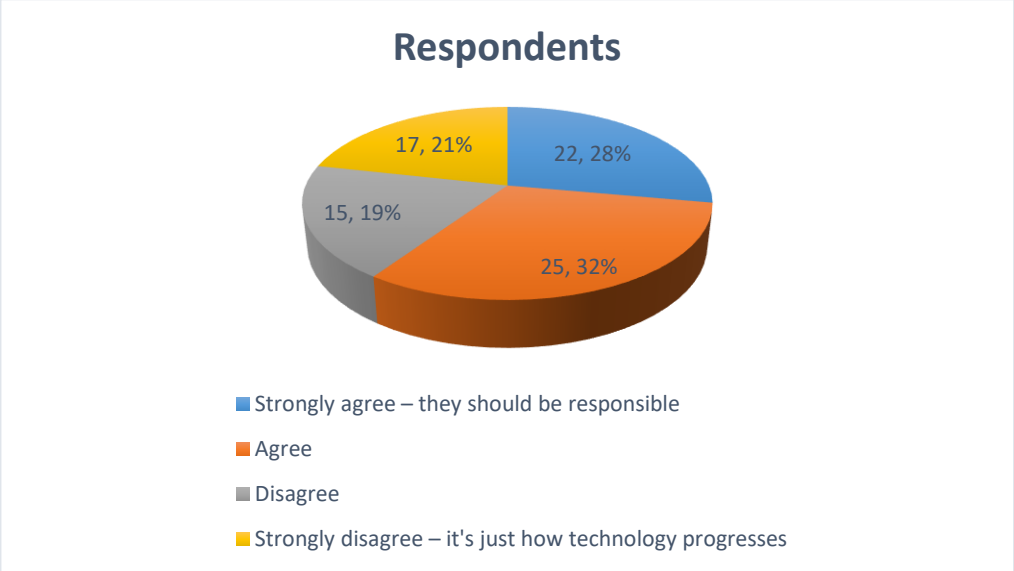
Particulars	Respondents
Strongly agree – they should be responsible	22
Agree	25
Disagree	15
Strongly disagree – it's just how technology progresses	17

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

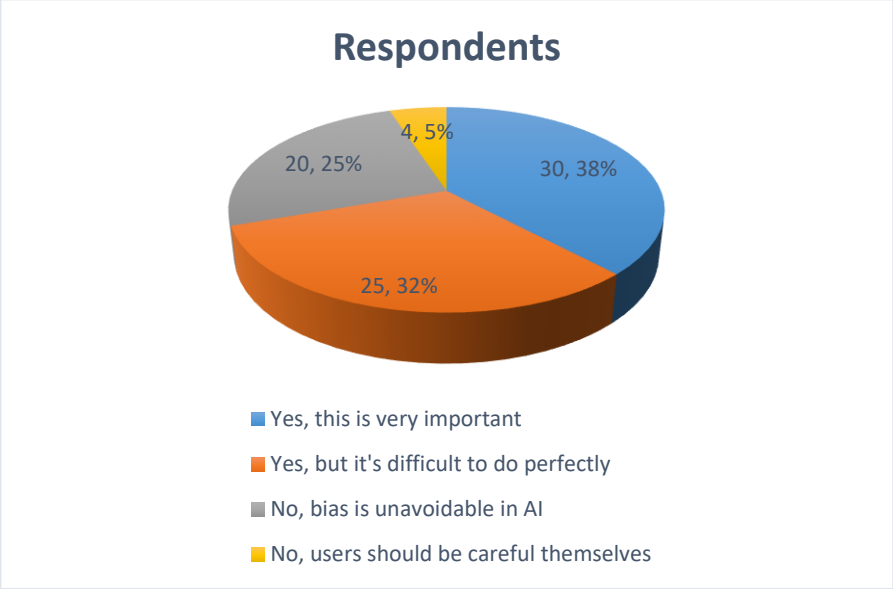
OPEN ACCESS Original Research Article



The above chart shows that majority of students 25 (32%) agree that AI developers should be morally accountable if their tools automate entry-level jobs and make it harder for fresh graduates like us to find work.

Do you believe AI developers have a moral obligation to test and fix biases (e.g., caste, gender, or regional bias) in their systems before releasing them to Indian users?

Particulars	Respondents
Yes, this is very important	30
Yes, but it's difficult to do perfectly	25
No, bias is unavoidable in AI	20
No, users should be careful themselves	4



ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

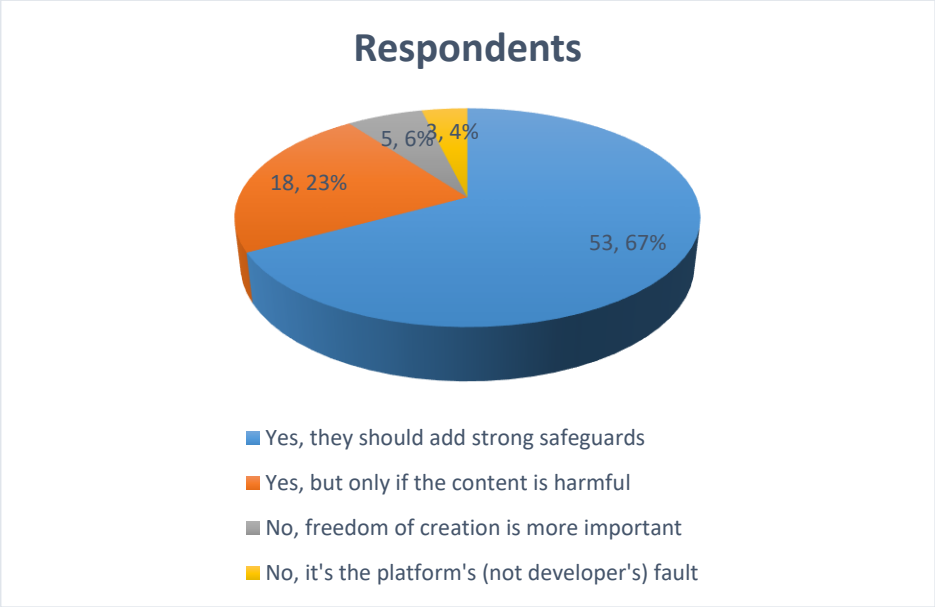
Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

The above chart shows that majority of students 30 (38%) believe AI developers should have a moral obligation to test and fix biases (e.g., caste, gender, or regional bias) in their systems before releasing them to Indian users.

In your view, should AI developers be held responsible for making it easier to create deepfakes and fake content that confuses or misleads college students?

Particulars	Respondents
Yes, they should add strong safeguards	53
Yes, but only if the content is harmful	18
No, freedom of creation is more important	5
No, it's the platform's (not developer's) fault	3



The above chart shows that majority of students 53 (67%) believe AI developers should be held responsible for making it easier to create deepfakes and fake content that confuses or misleads college students.

Do you think there is a real need for laws or ethical guidelines that force AI developers to prioritize "do no harm" to students and young users?

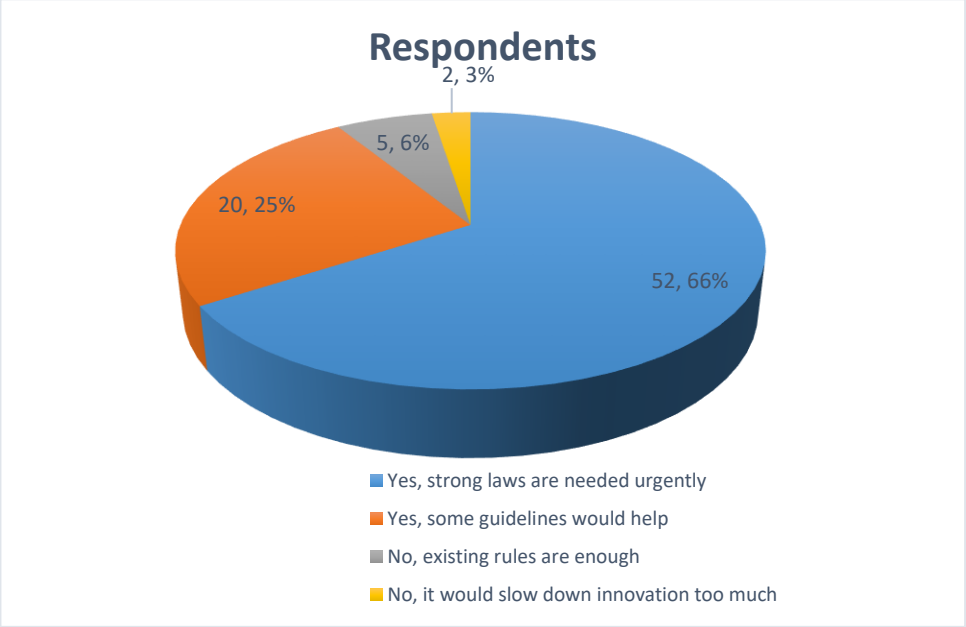
Particulars	Respondents
Yes, strong laws are needed urgently	52
Yes, some guidelines would help	20
No, existing rules are enough	5
No, it would slow down innovation too much	2

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article



The above chart shows that majority of students 52 (66%) believe there is a real need for laws or ethical guidelines that force AI developers to prioritize "do no harm" to students and young users.

Findings and Suggestions

The survey findings align with the previous research studies on few points, one of them being that AI developers carry meaningful moral responsibility and should not be excused from it. The papers provide the theoretical and professional foundation while the student respondents provide practical, harm-focused validation from the user side — demanding that developers actually exercise that agency to prevent concrete harms to students .There is a clear belief that developers should be held morally accountable for educational and societal harms (66–67% agreement on several issues), which is similar to papers understanding that developers possess moral agency and a professional obligation to act responsibly. This creates a powerful combined argument for a responsible AI.

The following are the few suggestions.

- Proactively exercise moral consideration while developing AI.
- Design features to reduce over-reliance and protect mental health.
- Integrate AI ethics and developer responsibility into the curriculum.
- Mandate transparency and bias reporting for India-specific risk.

Conclusion:

Moral responsibility in AI is essential to prevent harm and build a fair future, especially for India's college students. Data from recent years shows clear risks: widespread AI-assisted cheating harms real learning; high reliance on AI for stress relief (up to 88% in some surveys) can increase isolation and mental strain; entry-level job losses of 20–25% create fear and uncertainty for graduates; caste and gender biases in AI tools disadvantage many; and deepfakes fuel misinformation that confuses and divides youth.

AI itself has no morals—it reflects the choices of its creators and users. That is why developers must ensure transparency and fairness, companies must test for biases (especially India-specific ones like caste), and the government must enforce guidelines. Efforts like NITI Aayog's work on responsible AI, along with

ISSN–2278-5655

AMIERJ Aarhat Multidisciplinary International Education Research Journal

Volume–XV, Special Issues– I(b) Jan – Feb, 2026

OPEN ACCESS Original Research Article

UNESCO consultations in 2025, aim to balance innovation with protection.

Without moral responsibility, AI could make existing problems worse—widening inequality, damaging mental health, and reducing trust. But with it, AI can support education, create better jobs, and help students grow. As an educator seeing this daily, I believe strong ethics will let technology serve young people instead of holding them back. We need clear rules, awareness, and shared duty to make AI a positive force in India.

References:

1. *Tigard, D. W. (2021). Responsible AI and moral responsibility: a common appreciation. AI and Ethics, 1(2), 113-117. https://link.springer.com/article/10.1007/s43681-020-00009-0*
2. *Hedlund, M., & Persson, E. (2024). Expert responsibility in AI development. AI & SOCIETY, 39(2), 453-464. https://link.springer.com/article/10.1007/s00146-022-01498-9*
3. *Green, B. P., & Welie, J. V. (2025). Excuses, excuses: moral agency and the professional identity of AI developers. AI & SOCIETY, 1-12. https://link.springer.com/article/10.1007/s43681-022-00256-3*
4. *Griffin, T. A. (2024). The ethical agency of AI developers. AI and Ethics, 4(2), 179-188.https://link.springer.com/article/10.1007/s00146-025-02388-6*
5. *Ernst & Young (EY). (2025). The AIdea of India: Outlook 2026.https://www.ey.com/content/dam/ey-unified-site/ey-com/en-in/insights/ai/documents/is-india-ready-for-agentic-ai-the-aidea-of-india-outlook-2026.pdf*
6. *Gupta, A., & others. (2025). AI detection in academia: How Indian universities can safeguard academic integrity.* Proceedings, *107(1), Article 26. https://www.mdpi.com/2673-4591/107/1/26*
7. *MIT Technology Review. (2025, October 1). OpenAI is huge in India. Its models are steeped in caste bias. https://www.technologyreview.com/2025/10/01/1124621/openai-india-caste-bias*
8. *Rest of World. (2025, December 9). "Everyone is so panicked": Entry-level tech workers describe the AI-fueled jobpocalypse. https://restofworld.org/2025/engineering-graduates-ai-job-losses*
9. *The Straits Times. (2025, June 16). Indian colleges alarmed by AI misuse: Professors fight back with oral quizzes, live handwritten exams. https://www.straitstimes.com/asia/south-asia/ai-misuse-in-indian-colleges-and-how-professors-are-arresting-the-situation*
10. *Youth Ki Awaaz & YLAC. (2025). Youth Pulse Report: Is AI becoming India's accidental therapist? (Citizen Digital Foundation findings). https://citizendigitalfoundation.org/wp-content/uploads/2025/09/Are-you-there-AI_-Findings.pdf*

Cite This Article:

Asst. Prof. Nair A. and Asst. Prof. Patil A. (2026). *A Study on Moral Responsibility of AI Developers.* **In Aarhat Multidisciplinary International Education Research Journal**: Vol. XV (Number I, pp. 153–162)

Doi: https://doi.org/10.5281/zenodo.18638070