Fostering Ethics in AI: Perceptions from the Indian AI Curriculum

Ashutosh RAINA[a]*, Kushal MUNDRA[b], Prajish PRASAD[b] & Shitanshu MISHRA C[c]
[a] *Member, Ed Tech Society, India*
[b] *Flame University, India*
[c] *UNESCO MGIEP, India*

*raina.ashu@gmail.com

Abstract: In the age of rapid Artificial Intelligence (AI) advancement, universities worldwide have responded by offering specialized AI and Machine Learning (ML) courses to meet industry demands. However, amidst this surge in AI education, AI deployment's ethical and societal implications often need more attention. Unlike traditional algorithmic programming, AI involves intricate decision-making processes that are challenging to predict or explain, demanding a comprehensive understanding of its ethical dimensions. Various global initiatives have highlighted the ethical considerations surrounding AI, resulting in an increased emphasis on integrating ethics education into AI curricula. Despite this, there remains a gap in addressing these vital aspects across the broader AI education landscape, with ethics often relegated to the periphery of computer science courses. In this context, this paper explores the imperative for ethics courses in undergraduate AI education in India. We examine educators ' awareness and perceptions regarding ethics education in AI curricula through interactions with faculty members participating in professional development programs and workshops focused on AI pedagogy. Our analysis reveals that there is a pressing need to extend discussions on ethics beyond mere privacy concerns and traditional performance metrics, integrating real-life scenarios into the curriculum. This paper serves as a preliminary step towards a need for a general framework in structuring ethics education in AI in India, aiming to initiate a more comprehensive and standardized approach to AI ethics education, fostering responsible AI development in the future.

Keywords: AI Education, Ethics, Teacher perception, Active Learning

1. Introduction

The rapid evolution of Artificial Intelligence (AI) has led to significant advancements in various industries, transforming how businesses operate and enabling previously unimaginable innovations. As AI technologies continue to penetrate different sectors, there is a growing demand for a skilled workforce capable of harnessing the power of AI to drive innovation. In response, universities worldwide have begun offering specialized courses and programs in AI and Machine Learning (ML) to equip students with the necessary skills to meet industry demands.

However, amidst this surge in AI education, a critical aspect often receives inadequate attention – the ethical and societal implications of AI deployment (Jobin et al., 2019). Unlike traditional algorithmic programming solutions, AI operates in a distinct manner, often involving complex decision-making processes that are challenging to predict or explain. This unique nature of AI necessitates a comprehensive understanding of its ethical dimensions to ensure responsible and beneficial deployment (Buiten et al., 2019).

Recognizing this imperative, various global efforts have raised concerns about the ethical considerations surrounding AI (Hagendorff, 2020). Consequently, the integration of ethics education into the AI curriculum has gained prominence. While some universities have introduced topics and even entire courses dedicated to AI ethics, more attention must be paid to these essential aspects in the broader AI education landscape. Ethics units are frequently relegated to the concluding segments of computer science courses, indicating their lower priority than technical topics. Even within pertinent disciplines like machine learning, where ethical considerations are profoundly relevant, the standard integration practice remains limited (Garrett et al., 2020).

Several noteworthy examples, however, emphasize the significance of addressing AI ethics within educational frameworks (Khan et al., 2022). Initiatives attempting to map AI ethics teaching practices often employ human-driven approaches, where topics of interest are identified and categorized based on instructor descriptions or open coding (Garrett et al., 2020). Despite such endeavors, a comprehensive and standardized approach to integrating AI ethics across curricula still needs to be discovered (Javed et al., 2022).

This paper examines the necessity for incorporating ethics courses into undergraduate AI education, particularly in the context of India. We explore this need by engaging with participants in faculty development programs and workshops dedicated to AI instruction. Through these interactions, we aim to gain insights into the awareness and perspectives of educators within the Indian context, considering the region's socio-cultural diversity and unique ethical beliefs. Our goal is to understand how ethics education can be effectively integrated into AI curricula. By analyzing these perspectives, we contribute to the ongoing dialogue on promoting responsible AI development, advocating for a more comprehensive and standardized approach to AI ethics education. Although this paper does not delve into pedagogy, it serves as an initial step in guiding instructors. We identify the ethical topics currently addressed in AI courses, providing insights into the existing state of AI ethics education and raising pertinent questions. Our intention is to catalyze further research that is tailored specifically to the Indian educational system, ultimately establishing standardized practices for ethics in AI education.

2. Ethics in AI

2.1 Global Discussion

The current state of literature concerning the inclusion of ethics in AI-based courses reflects the heightened awareness of the ethical dimensions of artificial intelligence (Khan et al., 2022). Numerous institutions and organizations have emphasized the importance of ethical principles in AI, focusing on values such as Transparency, Privacy, Accountability, Fairness, Autonomy, and Explainability (Florida et al., 2018). Several committees and organizations have also introduced their frameworks and guidelines for the ethical implementation of AI technologies. (Floridi et al, 2018 & Jobin et. al., 2019)

However, despite the widespread acknowledgement of these principles and the proliferation of ethical frameworks, a significant challenge still needs to be solved - effectively integrating these ethical considerations into training materials, especially within the courses, to prepare the future AI workforce. The translation of abstract ethical principles into tangible, actionable content for AI education is an ongoing challenge.

Furthermore, there is a recurring concern within the literature about the perceived ambiguity and generality of ethical principles outlined by many organizations (Eitel, 2021). This ambiguity can lead to varied interpretations and applications of these principles, making it challenging to establish a standardized approach to ethical education in AI. Additionally, it has come to light that technology firms involved in AI and autonomous systems development often follow ethical guidelines based on their interpretations and priorities (Saheb, 2023). This further underscores the need for a comprehensive and standardized approach to ethics education in AI courses to ensure that future AI professionals are well-equipped to navigate

the complex ethical landscape of the field. As we delve deeper into the evolving discourse on ethics in AI, addressing these challenges and fostering a more consistent and robust ethical foundation for AI education remains a paramount concern.

2.2 Ethics in Global Curriculum

AI has made its way into school and college curricula, but the absence of a standardized framework has resulted in diverse implementations (Miao et al., 2022). Instructors often interpret and incorporate AI differently, leading to variations in curriculum content and teaching methods (Javed et al., 2022).

The strategies to achieve ethics education objectives remain subjects of ongoing deliberation (Eitel, 2021). Ethical components are mandated in Computer Science (CS) programs in the United States to secure accreditation. However, universities and educators are responsible for deciding how to execute this mandate (Javed et al., 2022). The advantages of ethics education are evident; discipline-specific ethics instruction has been shown to foster ethical growth and imbue students with the conviction that ethics is intrinsic to their profession, transcending mere public relations considerations (Knight, 2014 & Loescher, 2004). Furthermore, the field of CS has long advocated for a more comprehensive integration of ethics throughout the curriculum, including technical courses (Martin et al., 1996). Despite some progress in this direction, it has yet to become standard practice (Fiesler, 2020 & Saltz, 2019).

Additional queries arise regarding the methods of imparting ethics education, stemming from the notion that an ethics curriculum is optimal in breaking away from isolation and giving it a structured form (Eaton, 2018). Furthermore, proposals suggest incorporating project-based learning and utilizing science fiction as a tool for speculative exploration (Burton, 2015). Researchers have additionally pinpointed that global curricula demonstrate deficiencies in disciplinary breadth. They frequently exhibit limited geographical coverage, a bias toward Western cultures, and a tendency to offer courses primarily at just one academic level (Saltz et al., 2019).

This paper endeavors to initiate a discourse on integrating ethics into India's undergraduate (UG) curriculum for Artificial Intelligence (AI). This objective is achieved through interactions with UG computer science instructors and examining AI-based curricula to assess the inclusion of ethical components.

3. Ethics in AI Curriculum in India

Within educational research, scrutinizing syllabi and soliciting instructors' viewpoints on the subject matter is a customary approach to evaluating curricular requisites (Chong, 2016). This methodology has been frequently employed to offer insights into pedagogical strategies within the field of computer science (Becker, 2019). Assessing a syllabus is an effective means of ascertaining the components of knowledge encompassed within a course (Becker, 2019). Prior research has effectively applied this methodology to delineate the content and objectives of general technology ethics courses (Fiesler, 2020).

In this study, our focus centers on the domain of AI ethics within the context of AI courses and dedicated ethics-oriented courses. To steer our analysis, our principal research query was formulated as follows: What themes pertaining to Ethics in AI (Principles) are encompassed within AI-based courses offered at a selection of prominent universities in India? What is the comprehension and viewpoint of undergraduate (UG) instructors from various Indian universities regarding the instruction of ethics in AI?

Our dataset was bifurcated into two segments for analysis. We gathered instructors' perspectives on Ethics in AI during AI education workshops. Their responses were systematically collected through the utilization of Mentimeter activities. Simultaneously, the

scrutiny of syllabi was conducted by mining online resources for syllabi related to AI education within the top 100 undergraduate institutions in India specializing in engineering and technology.

3.1 Instructors' Perspective on Ethics in AI

Sixty responses were collected from various instructors across three AI Education workshops. A specific segment of these workshops was dedicated to engaging instructors in discussions regarding their perspectives on the necessity of ethics in AI education. These interactions were conducted using open-ended questions through the utilization of Mentimeter. Mentimeter is an interactive presentation and polling tool that enables real-time audience engagement and feedback during presentations and events (Mentimeter, 2023). The responses were recorded and later presented to the participants to guide subsequent discussions.

The questions posed during these interactions were divided into three sections. Initially, instructors were asked about their teaching experiences in AI. Following this, their viewpoints on 'What ethical challenges they encountered or perceived in the design and implementation of AI-based solutions' were sought. Subsequently, a scenario-based discussion was conducted, requiring instructors to design a lunch recommendation system for school children while considering the perspectives of various stakeholders, including management, students, parents, and doctors (Payne, 2019). Following an exercise involving an ethics matrix (Payne, 2019), the instructors were once again queried regarding their perspectives on the integration of ethics within the AI curriculum. Figure 1 illustrates a selection of responses recorded through Mentimeter.

For analysis, the responses were categorized into challenges related to data quality or responses and ethical and moral perspectives. This categorization followed coding based on associations derived from the existing literature (Khan et al., 2022).
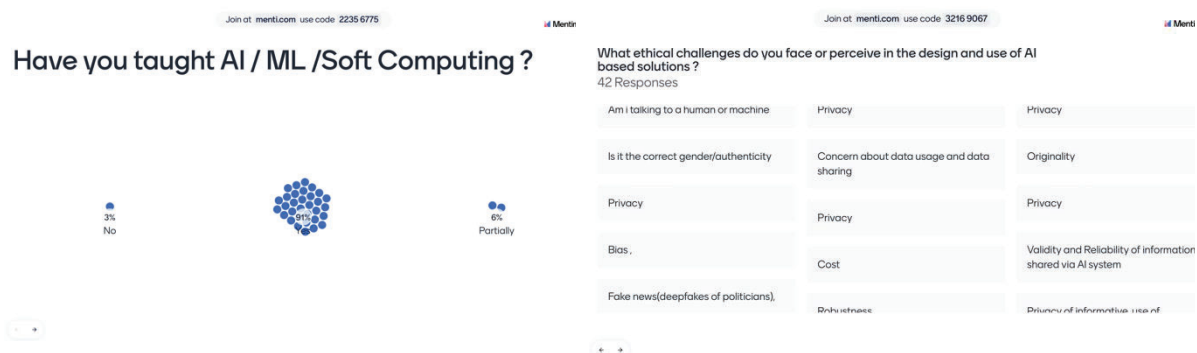

Figure 1: Mentimeter response recorded during one of the sessions.

Based on the data, it was observed that 90% (54) of the instructors had taught a course in artificial intelligence. The perceived ethical challenges were evenly distributed among ethical and moral aspects, with data quality challenges encompassing privacy as the most frequently reported concern in the former category and issues of reliability, authenticity, and validity in the latter category, as indicated in Figure 2. Following the discussion session and the activity involving the design of a learning exercise for teaching ethics in AI using an ethics matrix, participants were once again invited to share their perspectives. It was observed that some of the earlier responses had been consolidated, reflecting a more balanced distribution of concerns related to privacy, explainability, fairness, and reliability.

Figure 2: Participant perspectives on ethics in AI before and after the ethics session.

During focused group interactions with the participants, a recurring theme emerged regarding the need for a structured framework to guide the development of teaching materials and activities related to ethics. As relevant topics arose, these issues were raised sporadically without a dedicated session or structured approach. Furthermore, participants noted that using an ethics matrix with practical examples, such as designing recommender systems, helped them streamline their thought processes. Such a framework would facilitate the creation of activities that enable learners to engage with ethical considerations while designing AI-based solutions.

As discussed in previous sections, the absence of a universal framework has been identified as a challenge. Most of the participants in this study hailed from institutions where course curricula are structured at the university level. Thus, we embarked on the second phase of this research to investigate whether any structured methods or frameworks, either as individual courses or integrated into broader curricula, were being employed in teaching ethics in AI within Indian academia. In this phase, we explored universities and autonomous institutions where instructors had the autonomy to shape course curricula.

3.2 Ethics in AI in UG curriculum.

In this study, we compiled a list of the top 100 institutions in India specializing in engineering and technology by amalgamating data from the QS World Ranking and India's national equivalent ranking system. Subsequently, we systematically crawled the webpages of each of these institutions to amass comprehensive information regarding courses related to Artificial Intelligence (AI), Machine Learning (ML), Deep Learning, Soft Computing, and Ethics in AI. We meticulously recorded both the names and contents of these courses. Additionally, we scrutinized the course details to identify the utilization of terms commonly associated with ethics frameworks, as outlined by (Khan et al., 2022). It is crucial to acknowledge that due to disparities in website structures among institutions, a substantial amount of manual data extraction was requisite, and we continue to engage in ongoing efforts to maintain this data collection process. The forthcoming analysis encompasses findings from a 25% subset of institutions, selected randomly from public and private institutions for web crawling. Notably, our dataset exclusively comprises autonomous universities, and colleges affiliated with universities were excluded if the parent university was already represented in our dataset.
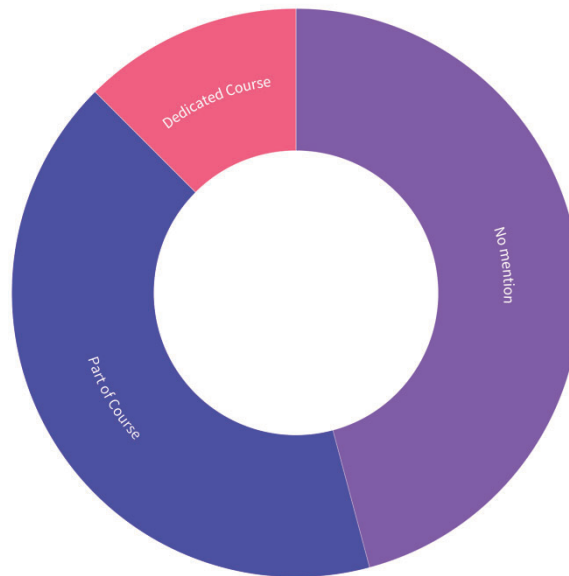
Figure 3: Inclusion of Ethics in AI for Institutions that offer AI-based courses.

Based on our initial data analysis, as shown in Figure 3, it is evident that 44% of the institutions offering AI-related curricula do not explicitly mention ethics in AI within their course descriptions. On the other hand, 40% of the institutions refer to ethics and allocate one or two sessions for discussions on ethics, while 6% incorporate ethics into the course title. Upon examining the course curricula, it becomes apparent that most courses within the 40% and 6% categories predominantly emphasize ethics within the context of engineering design without explicitly addressing the ethical principles relevant to AI. These courses touch upon terms like privacy, reliability, accuracy, and robustness, but often in assessing solution efficiency or comparing different algorithms. In some instances, there appears to be a tentative shift toward discussing the real-world implications of accuracy and robustness within specific scenarios. Notably, one elective course, "Ethics in AI", explicitly sets an objective to "Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice." Only two courses delve into the ramifications of ethical practices in the use of AI, particularly about policy, with one of these being specific to AI in education.

Based on this preliminary analysis, it is apparent that while most courses underscore the importance of robustness, accuracy, and reliability as performance metrics, a significant portion confines discussions of ethical considerations in AI to professional ethics in engineering. This underscores the pressing need for a standardized framework, even at the national level, to guide the incorporation of ethics in AI education. Only two-course descriptions allude to a structured framework encompassing privacy, explainability, robustness, and accuracy.

4. Discussion and Conclusion

Based on the preceding analysis, there is an urgent necessity to expand the discourse surrounding privacy and other efficiency parameters beyond the conventional performance metrics of AI. It is imperative to explore their ramifications within the context of the contemporary world, particularly within real-life scenarios. Furthermore, providing learners with opportunities to grapple with ethical dilemmas by extending these discussions across multiple perspectives, as exemplified in the school recommender system workshop, is essential. Privacy garners extensive attention, a trend reflected in the responses obtained from participants in our pre-workshop perception survey. However, a perceptible shift in participants' perspectives occurred following discussions and activities utilizing the ethics matrix within real-life scenarios. Their concerns extended to encompass explainability, robustness, accuracy, and more. Our observations suggest that implementing a

standardized framework, in our case, the ethics matrix, enables participants to systematize their thoughts concerning ethical concerns.

Additionally, it facilitates their ability to empathize with the predicament of selecting the appropriate perspective for fine-tuning algorithms. In such cases, it prompts contemplation on the steps required to enhance the transparency of the algorithmic processes and tuning parameters, allowing users to make informed decisions regarding reliance on the output and its extent. Intriguingly, we noted substantial discussions on ethics in AI within Indian school curricula, predominantly centered on the four pillars of ethics: privacy, explainability, robustness, and accuracy (Miao et. al. 2022). However, this emphasis must be noticed in undergraduate (UG) curricula.

Our data collection process is ongoing as we engage with instructors to gain insights into their perspectives on ethics in AI across various institutions. Further in-depth analysis, coupled with exploring questions such as the impact of university-based curricula and the variability among curricula in autonomous institutions, will significantly contribute to our understanding of and structuring ethics education in AI. Such analysis can facilitate the development of a unified framework for teaching and learning ethics in AI, grounded in real-life contexts and experiential learning.

References

Becker, B. A., & Fitzpatrick, T. (2019, February). What do cs1 syllabi reveal about our expectations of introductory programming students? In *Proceedings of the 50th ACM technical symposium on computer science education* (pp. 1011-1017).

Buiten, M. C. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation*, *10*(1), 41-59.

Burton, E., Goldsmith, J., & Mattei, N. (2015, January). Teaching AI Ethics Using Science Fiction. In *Aaai workshop: Ai and ethics*.

Chong, F. (2016). The pedagogy of usability: An analysis of technical communication textbooks, anthologies, and course syllabi and descriptions. *Technical Communication Quarterly*, *25*(1), 12-28.

Eaton, E., Koenig, S., Schulz, C., Maurelli, F., Lee, J., Eckroth, J., ... & Williams, T. (2018). Blue sky ideas in artificial intelligence education from the EAAI 2017 new and future AI educator program. *AI Matters*, *3*(4), 23-31.

Eitel-Porter, R. (2021). Beyond the promise: implementing ethical AI. *AI and Ethics*, *1*, 73-80.

Fiesler, C., Garrett, N., & Beard, N. (2020, February). What do we teach when we teach tech ethics? A syllabi analysis. In *Proceedings of the 51st ACM technical symposium on computer science education* (pp. 289-295).

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2021). An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Ethics, governance, and policies in artificial intelligence*, 19-39.

Garrett, N., Beard, N., & Fiesler, C. (2020, February). More than" If Time Allows", the role of ethics in AI education. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 272-278).

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and machines*, *30*(1), 99-120.

Mentimeter (2023), Interactive presentation software. Mentimeter. https://www.mentimeter.com

Javed, R. T., Nasir, O., Borit, M., Vanhée, L., Zea, E., Gupta, S., ... & Qadir, J. (2022). Get out of the BAG! Silos in AI ethics education: an unsupervised topic modelling analysis of global AI curricula. *Journal of Artificial Intelligence Research*, *73*, 933-965.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, *1*(9), 389-399.

Khan, A. A., Badshah, S., Liang, P., Waseem, M., Khan, B., Ahmad, A., ... & Akbar, M. A. (2022, June). Ethics of AI: A systematic literature review of principles and challenges. In *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering* (pp. 383-392).

Knight, H. (2014). *How humans respond to robots: Building public policy through good design*. Brookings.

Loescher, K. J. (2004). *The impact of an "Ethics across the curriculum" initiative on the cognitive moral development of business school undergraduates*. Nova Southeastern University.

Martin, C. D., Huff, C., Gotterbarn, D., & Miller, K. (1996). Implementing a tenth strand in the CS curriculum. *Communications of the ACM*, *39*(12), 75-84.

Miao, F., & Shiohira, K. (2022). K-12 AI curricula. A mapping of government-endorsed AI curricula.

Payne, B. H. (2019). Ethics of artificial intelligence curriculum for middle school students. *MIT Media Lab Personal Robots Group. Retrieved Oct 10*, 2019.

Saheb, T., & Saheb, T. (2023). Topical review of artificial intelligence national policies: A mixed method analysis. *Technology in Society*, 102316.

Saltz, J., Skirpan, M., Fiesler, C., Gorelick, M., Yeh, T., Heckman, R., ... & Beard, N. (2019). Integrating ethics within machine learning courses. *ACM Transactions on Computing Education (TOCE)*, *19*(4), 1-26.