Chapter – 21

AI LITERACY, ETHICS AND DHARMA IN THE LIGHT OF *YAJNAVALKYA SMRITI*

Subhodeep Mukhopadhyay [1]

ISBN: 978-91-410-0142-8 | **DOI**: 10.25215/9141001427.21

Abstract

Artificial Intelligence (AI) is becoming increasingly prevalent across all domains, transforming how we live, work, and learn. As a result, there is a growing need for comprehensive AI awareness and education, particularly at the school level. AI literacy is a growing field with theoretical frameworks that emphasize ethical usage as a key component. Yet, the study of ethical and societal consequences of AI is approached predominantly from a western viewpoint, even within India. Often AI ethics is treated as a marketing tactic, and used for the sake of political correctness rather than as a fundamental aspect of AI systems. AI is frequently misused in matters concerning fairness, privacy, and manipulation, and is prone to biases. The focus of this paper is on the *dharmic* ethical and societal implications of AI. Different aspects of AI ethics are investigated from the perspectives of the nine *lakshanas* (characteristics) of universal *dharma* defined in *Yajnavalkya Smriti*. The paper uses a hermeneutics methodological approach to explore whether Vedic ideas of ethics can be conceptually integrated into school-level AI literacy programs to enhance learning outcomes for societal well-being.

Keywords: *AI Literacy, Dharma Shastra, Digital Ethics, AI Education, Yajnavalkya Smriti*

[1] PhD in Education (Student), Manipal GlobalNxt University

Introduction

Artificial Intelligence (AI) is rapidly transforming different aspects of human life, and influencing various fields, whether education, business, healthcare, transport, or other domains. At the same time, there are concerns about its ethical, political, economic, and societal impact (Stahl et al., 2023). Some of its impacts are undoubtedly beneficial, but widespread and lasting harm can also result from the technology. Concerns about bias, discrimination, racism, data security, and privacy in AI are growing, especially in Western nations where a large number of such issues have been brought to the public notice.

For example, in academic and employment settings, biases in AI algorithms frequently lead to racial minorities getting evaluated as less likely to succeed (UN Human Rights, 2024). The misuse of AI in data security and privacy is also on the rise with companies using sophisticated AI tools on private customer data to increase profits (Chin-Rothmann, 2024). Another issue is that of AI-driven manipulation of human behavior. Manipulative marketing strategies when combined with extensive data collection for AI algorithms, allow firms to significantly influence user choices and behaviors to boost profitability (Petropoulos, 2022).

Fostering comprehensive AI awareness and education is crucial, especially among younger generations. AI literacy, the understanding and responsible use of AI, is, therefore, gaining traction as a field of study. Various AI literacy frameworks have been developed by companies, trade and policy organizations, and academia. However, their perspectives, views and issues concerned, are largely borrowed from western ideas of ethics. In the Indian context, a balanced approach that integrates indigenous views on ethics such as those drawn from Vedic wisdom is largely missing.

This paper investigates the dharmic ethical implications of AI by examining traditional Indian views on digital ethics based on the nine *lakshanas* (characteristics) of universal dharma from the *Yajnavalkya Smriti*. By doing so, the research aims to explore whether these ancient ethical ideas can be incorporated into school-

level AI literacy programs, potentially enriching students' understanding of ethics and societal well-being.

Literature Review

AI Ethics

AI ethics provides a framework for building AI systems that benefit society by addressing its social, psychological, and political implications. However, there is little consensus on what defines "ethical AI" or the standards and regulations required to achieve it. Jobin et al. (2019) propose five fundamental dimensions of AI ethics: transparency, justice and fairness, non-maleficence, responsibility, and privacy. Kazim and Koshiyama (2021) expand AI ethics to include psychological, social, and political aspects. The psychological dimension involves mental autonomy, protection from manipulation, and transparency in AI interactions. The social dimension covers justice, fairness, and environmental concerns, while the political dimension addresses AI's impact on democracy and economic structures, emphasizing the protection of civic rights. The issues are manifold, and the frameworks are numerous. However, AI ethics guidelines are often ineffective in solving such issues, and deviations frequently go unpunished. In fact, institutions often adopt ethical AI guidelines simply as a marketing tool (Hagendorff, 2020).

Dharma and *Yajnavalkya Smriti*

Dharma is a foundational concept in Indian traditions, derived from the root *dhri*, meaning "that which upholds." It is defined As*tha* principle which sustains and nurtures a civilized society. For millennia, it has served as the guiding light of Indian civilization, informing people of their social and ethical duties, roles, and responsibilities. *Dharma* fosters a balanced, sustainable civilization by promoting harmony between individuals, society, and the environment

“Dharma sustains the society, Dharma maintains the social order, Dharma ensures the well-being and progress of humanity.” – Mahabharata Karna Parva 69.581

The *Yajnavalkya Smriti* is an important text within the *Dharma-shastra* tradition, which encompasses the codes of Hindu rites and customs with a strong emphasis on social ethics. It provides an ethical framework to guide interpersonal and social relationships. *Yajnavalkya Smriti* is often regarded as the "best composed" and "most homogeneous" of the *Dharma-shastra* texts, noted for its refined vocabulary and sophisticated style (Lingat, 1973, p. 98). It is divided into three *Adhyayas* (books): *achara* dealing with moral code, *vyavahara* dealing with civil code, and *prayaschitta* dealing with penal code (Vidyarnava, 1918).

Yajnavalkya, mentions nine qualities or virtues to be adopted by every individual member of the society: non-harm, truthfulness, not stealing, purity, control of senses, charity, mercy, self-restraint, and forgiveness.

“ahiṃsā satyamasteyaṃ śaucamindriyanigraḥ/ dānaṃ damo dayā kṣāntiḥ sarveṣāṃ dharmasādhanaṃ.” – Yajnavalkya Smriti 5.122

Methodology

This study adopts a hermeneutic approach to analyze and understand Hindu texts such as the *Yajnavalkya Smriti*. This approach provides a structured framework for exploring the deeper meanings within these texts by considering their historical, cultural, and philosophical contexts (Gadamer, 2013; Schmidt, 2016). By applying hermeneutics to the *Yajnavalkya Smriti*, scholars can interpret concepts like *dharma* (moral duty and responsibility) and the *lakshaṇas* (qualities) in a way that is relevant to contemporary issues, including AI ethics. The study will be focusing on three key ethical issues in AI:

- justice and fairness which pertains to addressing biases and ensuring equitable decision-making
- privacy and data security which deals with protecting individual privacy and securing data from misuse
- manipulation and autonomy in order to prevent undue influence and enabling informed decision-making

These ethical issues will be analyzed alongside the nine universal *lakshanas* (qualities) of *dharma* according to Yajnavalkya: non-harm, truthfulness, non-stealing, purity, control of senses, charity, mercy, self-restraint, and forgiveness. Each issue will be analyzed to assess their relevance and adaptability for school-level AI education. Insights drawn from this mapping will guide the potential incorporation of dharmic values into AI literacy programs, aiming to foster a holistic and culturally resonant ethical framework in AI education

Results and Discussion

The application of each of the nine *lakshanas* of *sadharana dharma*, (the none attributes of Hindu universal code of conduct)—non-harm, truthfulness, non-stealing, purity, control of senses, charity, mercy, self-restraint, and forgiveness—is outlined below. Each *lakshana* can serve as a foundational principle aligned with key ethical domains, offering a culturally rooted framework to evaluate and enhance AI practices.

***Ahimsa* (non-harm) –** The *Yajnavalkya Smriti* emphasizes the importance of *ahimsa*, advising against harm or injury to any living being. It is a foundational principle in the Indian knowledge system, present across the Vedas and central to Jain philosophy. In the Mahabharata, we find the verse "*ahiṃsā paro dharma-stathāhiṃsā paro damaḥ ahiṃsā paramaṃ dānamahiṃsā paramaṃ tapaḥ*."

"Ahiṃsā is the greatest Dharma. Ahiṃsā is the supreme restraint. Ahiṃsā is the highest charity. Ahiṃsā is the greatest penance." Mahābhārata 13.117.37

In the context of AI systems, the principle of *ahimsa* advocates for safeguarding personal data to prevent misuse and minimize harm to users. By integrating *ahimsa,* AI can prioritize ethical data handling and user protection.

***Satya* (Truthfulness)** – Truthfulness is the cornerstone of all schools of Hindu philosophy, and the mantra "*Satyameva jayate*" (Truth alone triumphs) from *Mundaka Upanishad* is the national motto of India. Truthfulness in AI interactions can ensure transparency, making users aware when engaging with an AI system

rather than a human. This reduces the risk of deception and empowers users to make informed decisions based on clear, accurate information.

***Asteya* (Non-stealing)** – *Asteya* means to refrain from theft. It isalso one of the five *yamas* in the *Patanjali Yoga Sutra*, advocating that one should not take anything unearned. This principle applies to material possessions, as well as respecting others' time, energy, and efforts (Prasad, 1998). Non-stealing in the AI context promotes fairness by discouraging exploitative practices such as unauthorized data collection or covert tracking.

***Shaucha* (Purity)** – The shastras say that purity and cleanliness ought to be practiced at three levels: diet and personal hygiene, purity of mind, and clean environment. In the context of AI, *shaucha* can be used to emphasize the purity of intent and ethical data practices. This includes maintaining the confidentiality of user data and protecting it from external threats, aligning with the moral responsibility of safeguarding user information.

***Indriya Nigraha* (Control of Senses)** – Controlling the senses and restricting sensual pleasure are key aspects of the practice of *Indriya Nigraha.* When employed in AI systems, it will facilitate respect towards user autonomy. This principle discourages exploiting user vulnerabilities and promotes AI designs that support lawful and ethical interactions, fostering user trust and well-being.

***Dana* (Charity)** – *Dana* can promote AI practices that serve the greater good, encouraging accessibility, inclusivity, and fairness. By integrating the concept of charity or social profits instead of only financial profits, AI can be developed to prioritize societal benefit, ensuring equitable access and reducing disparities.

***Daya* (Compassion)** – Recognizing and actively alleviating suffering, known as *daya* is crucial for oneself and others. In AI design, the ideal of *daya* can be used to deal with manipulative tactics which cause suffering, thus respecting user autonomy. Compassionate AI systems treat users with care, fostering trust and ethical interaction.

***Dama* (Self-restraint)**– In Hinduism, *dama* is a concept that refers to self-control and self-restraint. *Dama* can support ethical AI by advocating for restraint in data collection and processing. This aligns AI practices with privacy protection, ensuring that only necessary data is used and user consent is respected.

Dama along with *dana* and *daya* are also mentioned in Brihadaranyaka Upanishad (5.2.3), which states that three characteristics of a good, developed person are self-restraint, compassion and love for all sentient life, and charity (Swami Madhavananda, 1950).

***Kshanti* (Forgiveness)** – *Kshanti* meaning patience, forbearance, or forgiveness, is an important concept in Hinduism. It is about self-regulation of emotions and behavior developed through mental training. In Buddhism too, *kshanti*or *khanti* (in Pali) is one of the ten perfections (parami) and is considered a noble quality.

Kshanti in AI can encourage continuous improvement by addressing and correcting errors, particularly those related to biases. This approach ensures that AI systems are accountable and responsive to ethical challenges

The results are summarized in Table 1 which illustrate how the nine *lakshanas* of *dharma* according to Yajnavalkya can be applied to AI Ethics.

Table 1

Mapping Dharma-Lakshana and AI/Digital Ethics

Lakshana	Justice and Fairness	Privacy and Data Security	Manipulation and Autonomy
Ahimsa		✓	
Satya			✓
Asteya	✓		
Shaucha		✓	
Indriya Nigraha			✓
Dana	✓		

Lakshana	Justice and Fairness	Privacy and Data Security	Manipulation and Autonomy
Daya			✓
Dama		✓	
Kshanti	✓		

These results highlight the potential of integrating *dharmic* principles to enhance AI literacy by embedding values such as compassion and non-violence. This approach can help students understand the societal impact of AI more deeply.

Current AI literacy frameworks often lack a foundation in personal morality or cultural ethics and are typically reactive, addressing issues only after they emerge. Incorporating *dharmic* values into school curricula could foster a safer, more responsible, and inclusive approach to AI ethics, allowing students to consider AI's ethical implications through culturally rooted values. The *dharmic* approach improves upon existing frameworks by prompting learners to think about the moral dimensions of AI use beyond mere legal compliance.

Conclusion

As AI continues to reshape modern society, using this technology ethically and responsibly is essential. The study makes a strong case to integrate the moral principles of Vedic knowledge with ethical concerns in AI literacy frameworks. By referencing the nine *lakshanas* of human duties from the Yajnavalkya Smriti, the paper illustrated how themes of AI fairness, privacy, and autonomy could be addressed. *Dharma-shastras* have the potential to provide practical guidance to address ethical challenges in AI and the digital world. This integration of *dharmic* principles into AI ethics offers a culturally grounded approach that can enrich AI literacy, fostering a more ethical, inclusive, and socially responsible use of technology

References

Chin-Rothmann, C. (2024, July 18). Protecting Data Privacy as a Baseline for Responsible AI. Center for Strategic & International Studies. https://www.csis.org/analysis/protecting-data-privacy-baseline-responsible-ai

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Gadamer, H. G. (2013). *Truth and method*. A&C Black.

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

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

Lingat, R. (1973). *The Classical Law of India*. University of California Press.

Petropoulos, G. (2022). The dark side of artificial intelligence: manipulation of human behaviour. *Bruegel*.

Prasad, R. (1998). *Patanjali's Yoga Sutras with the Commentary of Vyasa and the Gloss of Vachaspati Misra*. (Original work published 1922). Munshiram Manoharlal Publishers.

Schmidt, L. K. (2016). *Understanding hermeneutics*. Routledge.

Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., Lindqvist, B., ... & Wright, D. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review*, *56*(11), 12799-12831.

Swami Madhavananda. (1950). *Bṛhadāraṇyaka Upaniṣad*. Advaita Ashrama.

UN Human Rights (2024, July 30). Racism and AI: “Bias from the past leads to bias in the future.” https://www.ohchr.org/en/stories/2024/07/racism-and-ai-bias-past-leads-bias-future

Vidyarnava, S. C. (Trans.). (1918). *Yajnavalkya Smriti: With the commentary of Vijnaneshvara called The Mitaksara and notes from the gloss of Balambhatta* (Book 1). The Panini Office.