


# Chapter 3

## Ethical Theories in AI: Deontology, Consequentialism, Virtue Ethics, and Their Role in Fairness Accountability Transparency and Privacy

**Ilham Firmansyah**

 <https://orcid.org/0009-0007-6104-9966>

*Universitas Esa Unggul, Indonesia*

### **ABSTRACT**

*This chapter examines the application of classical ethical theories deontology, consequentialism, and virtue ethics to the governance of artificial intelligence (AI). It argues that the principles of fairness, accountability, transparency, and privacy cannot be adequately addressed without grounding in established moral philosophy. Through detailed analysis, the chapter demonstrates how deontological duties, consequentialist outcomes, and virtue-based character reasoning provide complementary insights into the ethical challenges of AI across healthcare, finance, law enforcement, and social media. A hybrid ethical framework is proposed, integrating duties, consequences, and virtues into a coherent approach to AI governance. By synthesizing theoretical foundations, case studies, and practical recommendations, this chapter contributes to interdisciplinary debates on responsible AI and offers pathways for aligning innovation with human rights, social trust, and global regulatory standards.*

DOI: 10.4018/979-8-3373-7011-8.ch003

Copyright © 2026, IGI Global Scientific Publishing. Copying or distributing in print or electronic forms without written permission of IGI Global Scientific Publishing is prohibited. Use of this chapter to train generative artificial intelligence (AI) technologies is expressly prohibited. The publisher reserves all rights to license its use for generative AI training and machine learning model development.

# 1. INTRODUCTION

Artificial intelligence (AI) has rapidly transformed from a domain of experimental research into a pervasive technology that informs, mediates, and shapes everyday human interactions. From algorithmic decision-making in financial systems to diagnostic support in healthcare and surveillance applications in public spaces, AI systems now exercise an unprecedented level of influence over human affairs. While these technologies promise increased efficiency, predictive power, and innovation, they simultaneously raise profound concerns regarding fairness, accountability, transparency, and privacy. These concerns are not merely technical in nature but are fundamentally ethical and philosophical, demanding systematic frameworks for analysis and governance.

The ethical evaluation of AI cannot rely solely on pragmatic or ad hoc principles (Nunes et al., 2025). Instead, it requires robust grounding in well-established traditions of moral philosophy that have historically guided human conduct and institutional practices. Among the most enduring and relevant frameworks are deontological ethics, consequentialism, and virtue ethics. Each of these traditions offers a distinct perspective on moral reasoning: deontology emphasizes duties and rights, consequentialism evaluates outcomes and social utility, and virtue ethics highlights the cultivation of moral character and virtues. By applying these philosophical lenses to AI governance, we can obtain deeper insights into the obligations of developers, the consequences of algorithmic deployment, and the cultivation of responsible institutional practices.

AI's expansion into sensitive domains has revealed a series of tensions that illustrate why ethical theory remains indispensable. Consider algorithmic decision-making in hiring processes. On the one hand, AI can streamline recruitment by processing thousands of applications efficiently. On the other, if the data used to train such systems reflects historical bias, the AI may reproduce discriminatory practices under the guise of neutrality. A consequentialist lens would evaluate whether the outcomes reduce or amplify social inequality, while a deontological perspective would interrogate whether such practices violate the duty to treat candidates with fairness and dignity. Virtue ethics, in contrast, would ask whether the institutions deploying such systems are cultivating justice, integrity, and responsibility in their hiring practices. Together, these theories illuminate the ethical stakes beyond technical optimization (Deng et al., 2025).

The triad of fairness, accountability, and transparency (FAT) has emerged as a central framework in AI ethics debates. Fairness speaks to distributive justice and the equitable treatment of individuals and groups. Accountability concerns the assignment of responsibility in cases of harm or misconduct, particularly in systems that operate with high autonomy and opacity. Transparency addresses the

34 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/ethical-theories-in-ai/396080](http://www.igi-global.com/chapter/ethical-theories-in-ai/396080)

## Related Content

---

### Data Privacy and Security Safeguarding Customer Information in ChatGPT Systems

Wasswa Shafik (2024). *Revolutionizing the Service Industry With OpenAI Models* (pp. 52-86).

[www.irma-international.org/chapter/data-privacy-and-security-safeguarding-customer-information-in-chatgpt-systems/345285](http://www.irma-international.org/chapter/data-privacy-and-security-safeguarding-customer-information-in-chatgpt-systems/345285)

### AI-Driven Decision Support System for Intuitionistic Fuzzy Assignment Problems

P. Senthil Kumar (2024). *Using Traditional Design Methods to Enhance AI-Driven Decision Making* (pp. 352-398).

[www.irma-international.org/chapter/ai-driven-decision-support-system-for-intuitionistic-fuzzy-assignment-problems/336707](http://www.irma-international.org/chapter/ai-driven-decision-support-system-for-intuitionistic-fuzzy-assignment-problems/336707)

### Software Effort Estimation: Harmonizing Algorithms and Domain Knowledge in an Integrated Data Mining Approach

Jeremiah D. Deng, Martin Purvis and Maryam Purvis (2011). *International Journal of Intelligent Information Technologies* (pp. 41-53).

[www.irma-international.org/article/software-effort-estimation/58055](http://www.irma-international.org/article/software-effort-estimation/58055)

### Real-Time Vehicle Traffic Prediction in Apache Spark Using Ensemble Learning for Deep Neural Networks

Anveshrihaa Sundareswaran and Lavanya K. (2020). *International Journal of Intelligent Information Technologies* (pp. 19-36).

[www.irma-international.org/article/real-time-vehicle-traffic-prediction-in-apache-spark-using-ensemble-learning-for-deep-neural-networks/262977](http://www.irma-international.org/article/real-time-vehicle-traffic-prediction-in-apache-spark-using-ensemble-learning-for-deep-neural-networks/262977)

### Deep Learning Perspectives for Prediction of Diabetic Foot Ulcers

Aman Sharma, Archit Kaushal, Kartik Dogra and Rajni Mohana (2024). *Metaverse Applications for Intelligent Healthcare* (pp. 203-228).

[www.irma-international.org/chapter/deep-learning-perspectives-for-prediction-of-diabetic-foot-ulcers/334350](http://www.irma-international.org/chapter/deep-learning-perspectives-for-prediction-of-diabetic-foot-ulcers/334350)