Chapter 3
The Evolution of Ethical Standards and Guidelines in AI

Bodhibrata Nag
https://orcid.org/0000-0002-6621-359X
Indian Institute of Management Calcutta, Kolkata, India

ABSTRACT

AI ethics focuses on the systematic study of what is right and wrong, with the aim of providing principles to resolve ethical dilemmas. AI products, such as deepfakes, have raised concerns about their potential to disrupt traditional industries, devalue human work, and threaten intellectual property rights. AI ethics are intertwined with the need for an understanding of potential risks and benefits. We can categorize AI ethics into principles-based, processes-based, and ethical consciousness approaches. Key themes emerging from AI ethics include transparency, justice, fairness, non-maleficence, responsibility and accountability, privacy, beneficence, freedom and autonomy, trust in AI, dignity, sustainability, and solidarity. The development of AI ethics requires defining universally applicable guidelines and promoting global collaboration. Collaboration between industry, academia, and the public is critical for detecting and evaluating AI-generated content, addressing the social and economic impacts of AI displacements, and building public trust in AI technologies.

DOI: 10.4018/979-8-3693-9173-0.ch003

INTRODUCTION TO AI AND ETHICAL CONSIDERATIONS

Artificial Intelligence (AI) systems, particularly those based on machine learning (ML) and deep learning (DL), operate through complex algorithms that learn patterns from large datasets (Goodfellow, Bengio, & Courville, 2016) (LeCun, Bengio, & Hinton, 2015) (Schmidhuber, 2015).

Machine learning is the process of teaching computers to use data to make predictions or choices without being explicitly programmed to do so. Deep learning is a type of machine learning that uses neural networks with many levels (hence the name "deep") to model high-level concepts in data, like speech and image recognition. Neural networks (Hecht-Nielsen, 1992), which are based on the structure of the human brain, are made up of nodes (neurons) that are connected to each other and process input in layers. Each layer takes in info and changes it into a slightly less concrete form. As the model learns, the weights of these links are changed based on how wrong its predictions were. This is usually done using a method called backpropagation. This process keeps going over and over until the model meets a certain level of accuracy.

Artificial intelligence (AI) has swiftly progressed from a theoretical notion to a revolutionary technology that profoundly affects multiple facets of society. This chapter examines the historical progression of AI ethics, current obstacles, and forthcoming paths to offer a thorough comprehension of the ethical deliberations in AI.

Concluding Summary Points:

- AI's rapid development from theoretical concepts to practical applications.
- The necessity of ethical considerations to navigate the complexities of AI.
- Importance of understanding AI's impact on society and the need for robust ethical frameworks.

The foundational concepts of AI and its profound impact on various sectors lay the groundwork for understanding the evolution of ethical standards in this rapidly advancing field.

Historical Developments in AI and Ethics

Early Foundations (1940s-1950s): The foundation for AI was established in the 1940s with notable contributions by pioneers like Warren McCulloch and Walter Pitts, who invented neural networks. (McCulloch & Pitts, 1943). Alan Turing, a pioneer of modern computing, investigated the philosophical and ethical aspects of computing, including machine awareness and the societal influence of machines, in his seminal work "Computing Machinery and Intelligence" in 1950. He contemplated