AI Ethics:
A Bibliometric Analysis, Critical Issues, and Key Gaps

Di Kevin Gao, California State University, East Bay, USA*

iD https://orcid.org/0009-0008-7391-5208

Andrew Haverly, Mississippi State University, USA

Sudip Mittal, Mississippi State University, USA

iD https://orcid.org/0000-0001-9151-8347

Jiming Wu, California State University, East Bay, USA

Jingdao Chen, Mississippi State University, USA

iD https://orcid.org/0000-0002-5133-9552

ABSTRACT

Artificial intelligence (AI) ethics has emerged as a burgeoning yet pivotal area of scholarly research. This study conducts a comprehensive bibliometric analysis of the AI ethics literature over the past two decades. The analysis reveals a discernible tripartite progression, characterized by an incubation phase, followed by a subsequent phase focused on imbuing AI with human-like attributes, culminating in a third phase emphasizing the development of human-centric AI systems. After that, they present seven key AI ethics issues, encompassing the Collingridge dilemma, the AI status debate, challenges associated with AI transparency and explainability, privacy protection complications, considerations of justice and fairness, concerns about algocracy and human enfeeblement, and the issue of superintelligence. Finally, they identify two notable research gaps in AI ethics regarding the large ethics model (LEM) and AI identification and extend an invitation for further scholarly research.

KEYWORDS

AI Ethics, AI Identification, Artificial Intelligence Ethics, Bibliometric Analysis, Human-Like Machine, Large Ethics Model (LEM), Literature Review, Machine-Like Human

INTRODUCTION

In November 2022, OpenAI took the world by storm with the debut of ChatGPT. The subsequent release of Bard by Google in February 2023 opened the floodgate of the once carefully guarded AI underworld. It also laid bare the upcoming breakneck competition that will reshuffle the winners and losers at the pivotal moment of Industry 4.0 (McKinsey, 2022), where a trivia chat-bot mistake cost

*Corresponding Author

Google $100 billion in market capitalization (Wittenstein, 2023) and a tiny graphics chip propelled NVIDIA into the trillion-dollar club (Fitch, 2023). Tech giants such as Apple, Amazon, and Facebook swiftly joined the race (Gurman, 2023; Dotan, 2023; Hao, 2023). The burgeoning excitement for anything AI also fueled a powerful surge in AI startups, with Anthropic AI and Inflection AI becoming newly minted unicorns despite being in the business for less than two years (Hu & Shekhawat, 2023; Konrad, 2023). New AI startups flourished.

The study of AI ethics is the study of the ethical and responsible development and deployment of artificial intelligence technology. Its significance is underscored by the rapid advancements in AI technology and the potential disruptions it may bring to our society. However, crucial questions need to be answered: How has AI ethics evolved, and what are the critical issues and debates? Additionally, what are the key gaps that necessitate further scholarly research? Our research is built on prolific AI ethics literature published between 2004 and 2023, a span of 20 years. Utilizing keyword patterns, we systematically analyze the development phases and trends in AI ethics. Drawing from comprehensive literature reviews, we present seven key issues that continue to be subjects of research and debate. Finally, we extend an invitation for additional scholarly research on the large ethics model (LEM) and AI identification.

This article provides a distinctive contribution to AI ethics across four areas:

- The delineation of the origins of modern AI ethics
- The contrast of human-like machines versus human-centric machines, highlighting two pivotal phases in AI development
- The initiation of the LEM discussion

AI ethics research may leverage the approaches used by the large language model (LLM) and get away from the bounds of conventional approaches of theories, principles, and frameworks.

- The initiation of discussions on AI identification

AI ethicists can remove mysteries and nebulosity on AI by pioneering approaches to identify and rate AI instances.

DEFINITIONS

Artificial intelligence (AI) was coined in 1955 by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon during the preparation for the Dartmouth Workshop (Dartmouth, 1956). John McCarthy defined AI as:

The science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. (2007, p. 2)

The term *ethics* originates from the Greek word "ethos" meaning "character." In the field of philosophy, ethics is the field that investigates individual behavior in society, providing rational justification for moral judgments, discerning what is morally right or wrong, and distinguishing what is morally just or unjust (Cornell, 2023). Artificial intelligence ethics is the study of rational justifications, what is morally right or wrong and just or unjust, for the responsible development and deployment of artificial intelligence technology.

17 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/article/ai-ethics/338367

Related Content

Data Profiling and Data Quality Metric Measurement as a Proactive Input into the Operation of Business Intelligence Systems
Scott Delaney (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications (pp. 2171-2188).*
www.irma-international.org/chapter/data-profiling-and-data-quality-metric-measurement-as-a-proactive-input-into-the-operation-of-business-intelligence-systems/142722

Modeling Energy Portfolio Scoring: A Simulation Framework
Rafael Diaz, Joshua G. Behr, Rafael Landaeta, Francesco Longoand Letizia Nicoletti (2015). *International Journal of Business Analytics (pp. 1-22).*
www.irma-international.org/article/modeling-energy-portfolio-scoring/132799

Classification Trees as Proxies
Anthony Scime, Nilay Saiya, Gregg R. Murrayand Steven J. Jurek (2015). *International Journal of Business Analytics (pp. 31-44).*
www.irma-international.org/article/classification-trees-as-proxies/126244

Supply Chain Analytics: Challenges and Opportunities
Xiuli He, Satyajit Saravaneand Qiannong Gu (2014). *Encyclopedia of Business Analytics and Optimization (pp. 2364-2375).*
www.irma-international.org/chapter/supply-chain-analytics/107420

Data Protection and BI: A Quality Perspective
Daragh O. Brien (2014). *Information Quality and Governance for Business Intelligence (pp. 389-413).*
www.irma-international.org/chapter/data-protection-and-bi/96162