

24 - 26 July 2026

Copenhagen , Denmark

Ethics of AI: The Trolley Problem... But for Real

Prof. Dr. Donna Roberts

Embry

Riddle Aeronautical University

Worldwide, the United States

Abstract

The trolley problem, a classic philosophical thought experiment, has long served as a tool for exploring ethical decision-making in life-or-death scenarios. With the advent of artificial intelligence systems capable of autonomous action, this once-hypothetical dilemma has become an urgent practical concern. This paper examines the ethical frameworks governing AI decision-making in high-stakes contexts, with particular focus on autonomous vehicles, medical triage algorithms, military drone systems, and predictive policing tools. Drawing on utilitarian, deontological, and virtue ethics traditions, the study analyzes how competing moral frameworks translate into algorithmic design and the governance structures that shape them. The paper further explores the critical question of moral agency. When an AI system makes a decision that results in harm, who bears responsibility—the developer, the deploying organization, or the regulatory body? Findings from the MIT Moral Machine experiment and documented case studies of AI decision failures are used to illustrate the gap between theoretical ethics and real-world implementation. The paper argues that effective ethical governance of AI requires not only technical solutions but a multidisciplinary approach incorporating social science, law, and philosophy. Recommendations for policy frameworks, including the EU AI Act and UNESCO AI Ethics guidelines, are evaluated for their adequacy in addressing these challenges. The paper concludes by calling for a global, inclusive dialogue to establish standards for ethically aligned AI systems.

Keywords: Accountability; Algorithmic Bias; Autonomous Systems; Trolley Problem; Moral Philosophy