Streszczenie w języku angielskim

Imię i nazwisko Doktoranta/ki	**Maciej Mróz**
Tytuł rozprawy doktorskiej	*Morality of Artificial Agents* *Moralność sztucznych agentów*

The rapid development and proliferation of artificial intelligence (AI) demand a diverse, multi-voiced debate about how to shape technological progress. AI's potential to reshape social and economic reality is arguably unprecedented for any technology in human history.

The dissertation aims to contribute to the debate on whether artificial agents can be considered moral agents. Although it focuses on artificial intelligence, it deliberately follows the argument that "artificial agency" offers better framing than "artificial intelligence" for describing the nature of AI systems. All parts of the thesis operate in the field of philosophy; however, they are supplemented by relevant discourse from AI theorists and researchers. An important feature of this dissertation is its grounding in the discipline of theology.

The study analyzes efforts to equip artificial agents with capacity for moral deliberation, focusing on concepts of machine ethics and Artificial Moral Agents (AMAs), and critically reviews the academic discourse on these subjects to date. The analysis also examines whether morality indeed can be computed. Additionally, it critically reviews attempts at building computational morality, exemplifying the problem of overly narrow and reductionistic interpretations of philosophical concepts by some researchers in robotics and artificial intelligence.

The study's key contribution introduces moral creativity as essential to genuine moral action. Drawing on Bergson, Tischner, and Wojtyła, it demonstrates that moral acts involve creative self-determination. The dissertation argues that this creative dimension cannot be computationally replicated.

The conclusion: artificial systems cannot achieve true moral agency but can serve as moral instruments augmenting human judgment. The theological perspective emphasizes irreducible dimensions of human nature that technology cannot replicate, while recognizing AI's role in the broader sociotechnical context.