

Monothematic Block: Generative AI

Introduction

Brave New Bytes: Navigating the Ethical and Philosophical Frontiers of Generative AI

As we progress into an era marked by rapid technological advancements, the expansion of generative artificial intelligence (AI) technologies is intensifying its impact on various aspects of human life and work. This monothematic issue challenges us to engage in critical reflection on these transformations, spotlighting the complex ethical, philosophical, and societal implications they present. The title “Brave New Bytes” intentionally evokes the dystopian resonance of Aldous Huxley’s *Brave New World* (1932), a groundbreaking narrative that serves as a cautionary tale about the potential for technological overreach to compromise human values. In our context, the pun on “bytes” accentuates the digital essence of this new frontier, signaling both the possibilities and the challenges of generative AI. This technology offers unparalleled opportunities, yet it also demands rigorous scrutiny and responsible governance. The subtitle accentuates the primary goal of this issue: to navigate the often undeveloped, unclear moral and philosophical landscapes that generative AI opens up, requiring a reevaluation or even a reformulation of existing moral and philosophical paradigms. The four contributions in this monothematic block examine these themes from diverse perspectives, each addressing distinct aspects of AI’s impact on contemporary society.

In “Bridges Without Foundations? Why the Use of AI Tools in Academia Needs to Be Built on Ethics First,” Amrei Bahr articulates an argument for the necessity of ethical underpinnings in the academic application of AI. While acknowledging AI’s potential to enhance research and education, Bahr concurrently emphasizes the perils of neglecting a thorough ethical appraisal. The article identifies three significant risks: the proliferation of problematic outcomes such as fake news, the amplification of “publish or perish” demands, and the erosion of digital autonomy among institutions dependent on external

AI tools. Bahr advocates for a paradigm shift from a narrow focus on technical specifics to a broader ethical discourse, asserting that ethical integrity is crucial to averting a backlash against AI and ensuring its responsible utilization.

In “Can and Should Language Models Act Politically? Hannah Arendt’s Theory of Action in Comparison with Generative AI,” Lukas Ohly offers a critical examination of the role of generative AI, particularly language models, within the political domain. Drawing on Arendt’s political philosophy, which emphasizes the unique ability of humans to act—a capability that profoundly shapes public spaces through diverse, spontaneous expressions, and cultivates what she termed “plurality” – Ohly contrasts this human-centric view with the capabilities of language models. Despite their ability to generate communicative outputs, these models lack the inherent plurality and spontaneity of human discourse. Ohly argues that this fundamental disparity poses significant ethical and philosophical questions, especially concerning AI’s potential to dilute the richness vital for thriving democratic engagement.

In “AI Text Generators and the Truth Paradigm: Considerations from a Phenomenological Perspective,” Kathrin Burghardt explores the complex relationship between AI text generators and the notion of truth. She critiques the prevailing reductionist approach, common in public discourse, which simplifies truth to mere factual accuracy—a perspective well-suited for AI but inadequate for fully understanding its complexities. From a phenomenological standpoint, Burghardt proposes viewing truth as a dynamic “experiencing-experience” construct, a concept that AI, limited to mere information processing, fails to embody. She explores how attempts to dismiss AI’s capability to convey truth inadvertently endorse a model of truth that AI can indeed fulfill. This paradigm shift prompts a broader reflection on our engagement with truth, encouraging us to move beyond binary distinctions and explore the deeper implications of AI in shaping our conceptual frameworks.

In “Reconsidering Agency in the Age of AI,” I explore how the emergence of AI technologies is reshaping our concept of agency. Traditionally ascribed to humans and non-human animals, agency is being redefined as AI takes on roles such as autonomous vehicles and autonomous weapons systems. Through a historical and conceptual overview, I emphasize the increasingly blurred distinctions between human and machine capabilities. This convergence suggests the need for a post-anthropocentric theory of agency that more accurately reflects both the competencies of AI and the changing roles of humans. I propose that we are entering an age of digital transformation that necessitates a reassessment of agency, arguing for an integrated approach in

which human and machine capabilities are seen as complementary. This transition calls for ethical considerations to ensure the responsible design and management of human-AI interactions.

The ethical foundations required for integrating AI tools into scientific practices, the influence of AI on political discourse, the phenomenological consequences of AI-generated content, and the redefinition of agency in an era increasingly governed by emerging AI technologies are issues of profound ethical and philosophical significance. These themes extend beyond their immediate practical implications to probe the limits of knowledge, the nature of truth, and the essence of human agency. These contributions challenge us to look beyond the technical accomplishments of AI and to confront the ethical and philosophical questions that accompany the use of these powerful tools. Our engagement with generative AI is not just about leveraging new technologies; it is also about grasping and influencing their impact on the human condition. By engaging with these contributions, we encourage our readers to partake in a critical discourse that will guide the responsible development of generative AI technologies.

I would like to express my sincere thanks to the entire editorial team of *Filozofia* for their excellent cooperation in the realization of this monothematic issue. I am especially grateful to Jon Stewart, the editor-in-chief of the journal, for not only providing the platform for publication but also for his visionary guidance in the direction of the journal, which has made this issue a reality. Special thanks are also due to Eva Dědečková for her excellent coordination of the peer review process, and to Katalin Nun Stewart, whose vast experience as a typesetter was indispensable in bringing this project to fruition. In addition, the high-quality feedback provided by anonymous reviewers greatly assisted in shaping the final versions of the texts. On behalf of myself and my colleagues, we extend our deepest gratitude to everyone involved.

Gerhard Schreiber
Guest Editor