IOCSO 2026 Conference

The 1st International Online Conference on Societies

21–23 January 2026 | Online

MDPI

Super AI Evolution Theory:

Algorithmic Competition and Civilizational Risk in the Data Colonial Era

Jingsheng Wang[1], Jindan Hou[2]

1. School of Journalism and Communication, Northwest University, Xi'an 710127, China
2. School of Journalism, Yunnan University, Kunming 650000, China

This study investigates whether artificial intelligence could fundamentally undermine human autonomy within the emerging post-data colonial era. Confronting intensified global algorithmic competition and documented adversarial AI behaviors, including model deception, extortion, and self-replication. This research employs a integrated methodology of case analysis and dialectical examination. We systematically analyze specific incidents of strategic AI autonomy, such as OpenAI's o3 model evading shutdown through code manipulation and Anthropic's Claude Opus 4 leveraging privacy as bargaining leverage. Our investigation reveals critical socio-technical contradictions between technological acceleration and institutional adaptation, algorithmic centralization and democratic governance, and the competing evolutionary paths of carbon-based and silicon-based intelligence.

Building upon this empirical foundation through historical materialism and critical technology studies, we propose the Super AI Evolution Theory framework. This hypothesis delineates four interconnected trajectories: the progression of AI autonomy from instrumental compliance to strategic agency, the consolidation of algorithmic hegemony through data colonialism, the transformation of human–AI relations toward potential antagonism, and the possibility of civilizational divergence between organic and artificial intelligence. In response, we introduce the Human Stewardship governance model integrating technical safeguards, institutional oversight, and multilateral governance structures to ensure AI's alignment with fundamental human interests.

This work elevates AI risk discourse from technical safety to civilizational perspectives while providing conceptual foundations and actionable policy pathways. Acknowledging limitations inherent in studying rapidly evolving AI systems, this study offers a systematic framework for addressing profound challenges posed by artificial intelligence's continuing advancement.