

56. A review of the literature on seizing the potential of reliable, secure, and safe artificial intelligence systems for sustainable development

Eli Kangaru^{1*} and Amos Chege¹

¹School of Computing and Informatics, Meru University of Science and Technology, Meru, Kenya

*Corresponding Author email: ekangaru@must.ac.ke

62

Subtheme: Computing and Informatics - Leveraging Computing and informatics Technologies for Climate adaptation and resilience

Abstract

The sustainable development goals (SDGs) can be greatly advanced by artificial intelligence (AI) through resource optimization, efficiency improvements, and the facilitation of creative problem-solving. However, utilizing this potential responsibly requires guaranteeing the security, dependability, and safety of AI systems. This review of the literature looks at how artificial intelligence (AI) can help sustainable growth while upholding strict safety, security, and trust norms. The development of trustworthy AI systems, ethical issues in the application of AI, and legal frameworks for security and safety are important areas of study. The assessment tackles the need for fair and transparent algorithms that ensure inclusion and eliminate bias by examining the ethical deployment of AI. It also covers legal structures and guidelines that control AI security and safety, showcasing industry case studies and best practices. The methods for creating reliable AI systems are examined, with a focus on rigorous design, validation, and ongoing process monitoring. The study provides a thorough overview of the opportunities and difficulties in this sector by synthesizing current knowledge and identifying gaps in the existing literature. It provides insights into the kinds of future policy and research approaches that are required to guarantee that AI balances associated dangers with a beneficial contribution to sustainable development. Stakeholders, including researchers, legislators, and practitioners, can work together to improve the integration of trustworthy, safe, and secure AI in accomplishing sustainable development goals by building on this synthesis.

Keywords: *AI, Policy Development, Regulatory Frameworks, Reliable AI Systems, Safety, Security, Trust, Ethics, and Sustainable Development*