

AI-Assisted Framework for Sustainable Human-Compatible Innovation and Transformation

Human, Soheil; Bernroider, Edward

Published in:
CENTERIS - International Conference on ENTERprise Information Systems

Accepted/In press: 01/11/2024

[Link to publication](#)

Citation for published version (APA):
Human, S., & Bernroider, E. (in press). AI-Assisted Framework for Sustainable Human-Compatible Innovation and Transformation. In *CENTERIS - International Conference on ENTERprise Information Systems*

AI-Assisted Framework for Sustainable Human-Compatible Innovation and Transformation

Soheil Human*, Edward Bernroider

Institute for Information Management and Control, Vienna University of Economics and Business, Vienna, Austria, Europe

Abstract

Introducing the “AI-Assisted Framework for Sustainable Human-Compatible Innovation and Transformation” (AI-FIT), this study investigates the use of Artificial Intelligence within innovation and transformation processes. It highlights AI’s role in addressing the shortcomings of traditional methods like Design Thinking and Human-Centered Design amidst the demands for resource-efficient, sustainable and human-compatible innovation. The study applies the Design Science Research (DSR) framework—from problem identification to designing a conceptual architecture—to explore how AI can improve innovation and transformation processes by enhancing outcomes, reducing costs, and facilitating the adoption of sustainable solutions aligned with human needs and capacities. Furthermore, it briefly reflects on potential drawbacks and concerns aiming to present a measured perspective on AI’s evolving role in future human-compatible innovation and transformation endeavours.

Keywords: Sustainable Human-Compatible Innovation; Sustainable Human-Compatible Digital Transformation; Artificial Intelligence

1. Introduction

The rapid expansion of innovation, particularly in digital solutions, compels organizations to invest heavily in digital transformation. However, challenges arise when these innovations and transformations do not meet human needs, values, capacities, diversity, and sustainability. Approaches like *Design Thinking* [1] and *Human-Centered Design* [2] aim to align innovations with human requirements, yet they often demand substantial time, financial

* Corresponding author. Tel.: +43-1-31336-5198.

E-mail address: soheil.human@wu.ac.at, edward.bernroider@wu.ac.at

resources, and effort, which can be prohibitive for resource-limited organizations. This research in progress seeks to address the question: *How can an AI-assisted integrative framework facilitate human compatibility and sustainability in digital transformation?* Here, we examine the challenges, objectives, and design of such a framework, focusing on the first three steps of the *Design Science Research (DSR)* methodology [3]: *problem identification, objectives definition, and the conceptual architecture* within the *design and development* step.

2. Applying the Initial Stages of the DSR Methodology to AI-Assisted Human-Compatible Innovation

Here, we briefly outline the first three stages of the DSR methodology in our study:

DSR Step 1 (Problem Identification and Motivation): Challenges in Human-Compatible Innovation and Transformation — Innovations must be sustainable and human-compatible, addressing technological, societal, and ecological dimensions. However, implementing this principle faces difficulties due to the diversity of [human] needs [4] and the conflicting priorities of business interests. Design Thinking [1] and Human-Centered Design [2] foster user-centered innovations, but traditional business models can resist these iterative approaches. Cultural resistance to these frameworks exacerbates challenges, particularly in integrating agile, experimental methods into corporate strategies. Additionally, the cost of these methodologies and the necessity of engaging diverse human input can discourage their adoption. These challenges limit the ability of organizations to foster human-compatible innovations. AI offers potential solutions by reducing the need for extensive human resources while enhancing the sustainability of these practices. This study examines how AI can mitigate these barriers and set the stage for sustainable and human-compatible innovations and transformations.

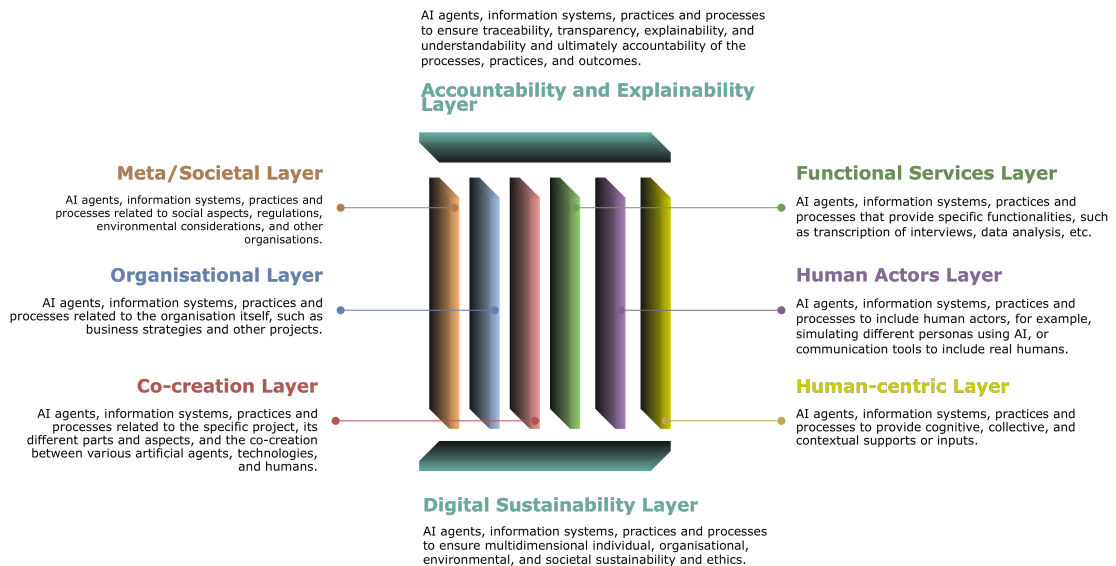


Figure 1. The Conceptual Architecture of AI-FIT, depicting a multilayered set of AI agents, information systems, practices, and processes, all designed to collaborate and ensure transparency and adaptability in fostering human-compatible innovation and transformation.

DSR Step 2 (Objective Definition): Enabling Sustainable Human-compatible Innovation and Transformations Through AI — AI, particularly generative AI, offers solutions to address challenges in human-centered design frameworks. For instance, AI can create diverse personas, consider and update knowledge bases, handle cognitive tasks, and perform repetitive processes, all while ensuring project alignment with regulatory and business strategies. These tasks, usually resource-intensive, can be managed efficiently by AI systems. Our

objective is to assess AI's potential in enhancing human-compatible innovation frameworks. AI-assisted frameworks could reduce costs, improve sustainability, and align innovations with human needs and organizational strategies. Although AI may seem to distance human involvement from the innovation process, *when designed and developed correctly*—and with humans kept in the loop—it can support the goal of human-compatible outcomes by addressing biases, promoting ethical practices, and ensuring transparency, accountability, and alignment with [human] needs and values.

DSR Step 3 (Conceptual Design): AI-Assisted Framework for Sustainable and Human-Compatible Innovation and Transformation (AI-FIT) — In the third DSR activity, we propose *AI-FIT*, with its conceptual architecture illustrated in Figure 1. AI-FIT integrates AI systems and methodologies to foster human-compatible innovation and transformation. It draws on theoretical foundations, among others, such as 4E Cognition Theory [5], which emphasizes the embodied and dynamic nature of cognition, and theories like Actor-Network Theory [6] and Co-production Theory [7], which examine the interrelations between human and technological actors. The AI-FIT framework's modular, multilayered architecture ensures adaptability, transparency, and accountability. Each AI agent and system operates at different layers, addressing societal, organizational, and human needs, while providing the tools and solutions necessary for innovation and transformation processes. This collaborative structure allows components to validate and audit each other, enhancing quality and reducing biases.

3. Discussion, Conclusion and Future Work

AI-FIT facilitates human-compatible innovation but faces significant barriers, including resistance to AI's role in creativity. While AI has the potential to make innovation processes more inclusive and democratic by managing and generating diverse datasets and inputs, challenges such as AI bias and the exclusion of marginalized groups remain. To ensure AI supplements rather than replaces human input, robust governance and ethical frameworks are essential. Technological advancements in AI, coupled with the growing recognition of sustainability, are key drivers for frameworks like AI-FIT. Support from academia, civil society, policymakers, and the private sector will be crucial for fostering innovation that aligns with societal values and ethical standards. Continuous research is necessary to ensure AI systems are fair, representative, and do not perpetuate existing inequalities. This study proposes AI-FIT as a framework to integrate AI into innovation and transformation processes, addressing the limitations of traditional methodologies. AI offers opportunities to enhance human compatibility and sustainability in innovation, but its adoption requires organizational shifts. Future research will advance the DSR stages, focusing on AI-FIT's further development, implementation and empirical evaluation in real-world contexts. Collaborations with international organizations will enable a rigorous assessment of its effectiveness in fostering human-compatible innovation. These efforts will generate valuable insights for both academic and practitioner communities, showcasing AI's transformative potential in driving inclusive and sustainable innovation and transformation.

References

- [1] Brown, T., 2008. Design Thinking, *Harvard Business Review* 86 (6), p. 84-92.
- [2] IDEO.org, 2015. The Field Guide to Human-Centered Design. IDEO.org.
- [3] Peffers, K., Tuunanen, T., Rothenberger, M. A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research, *Journal of Management Information Systems* 24 (3), p. 45-77.
- [4] Human, S., Watkins, R., 2023. Needs and Artificial Intelligence, *AI and Ethics* 3 (3), p. 811-826.
- [5] Newen, A., De Bruin, L., Gallagher, S., (Eds.), 2018. *The Oxford Handbook of 4E Cognition*, Oxford University Press, Oxford.
- [6] Law, J., 2008. Actor-Network Theory and Material Semiotics, *The New Blackwell Companion to Social Theory*, p. 141-158.
- [7] Jasanoff, S., 2004. The Idiom of Co-production, in *States of Knowledge*, Routledge, pp. 1-12.