Human Vs. Machine: Gen-AI Frameworks and Recommendations in HEOR

MSR41

Shilpi Swami, Tushar Srivastava

ConnectHEOR Ltd., London, UK. Email: shilpi.swami@connectheor.com

connectHEOR

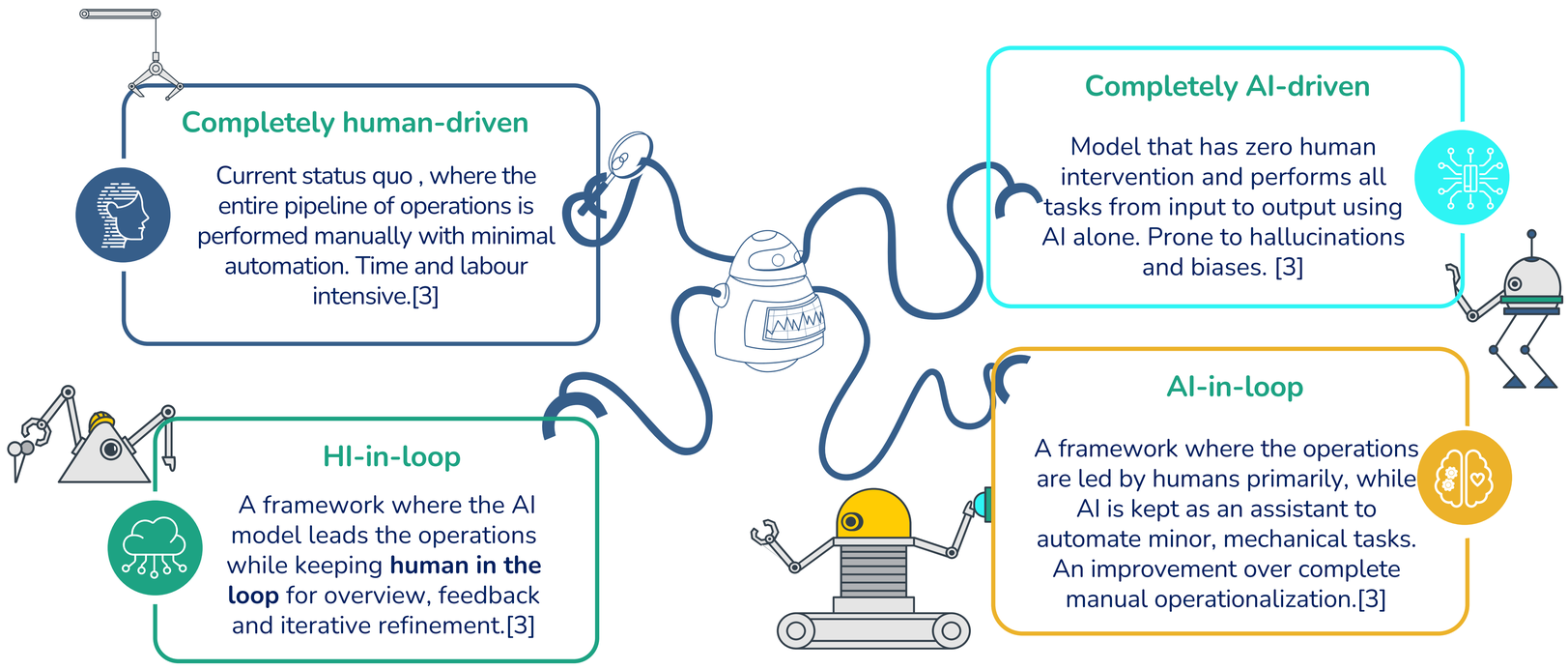
BACKGROUND

CONTEXT: Generative Artificial Intelligence (Gen-AI) and foundation models have begun to impact many areas of science and medicine. Applications in health economics and outcomes research (HEOR) are also emerging, with health technology assessment (HTA) agencies, such as the National Institute of Care and Excellence (NICE), in the process of developing guidelines for their use in submissions. As the field progresses, emerging techniques are being explored with the goal of improving the accuracy and usefulness of these models in health and medicine, and by extension, in HEOR.[1,2]

OBJECTIVES: To identify and evaluate distinct frameworks of potential Gen-AI integration in HEOR studies and recommend the most effective approach for maintaining ethical standards, accuracy, trust and efficiency.

METHODS

This study conducts a conceptual exploration of Gen-AI integration for HEOR based application. Four distinct frameworks of Gen-AI integration potentially useful for HEOR were identified and evaluated.



Gen-AI Applications in HEOR

APPROACH	DEFINITION	APPLICATION IN HEOR
Prompt Engineering	Specialized prompting techniques that modify the in-session reasoning algorithm of the LLM to guide towards a structured response.[2]	Guide LLMs to simulate human-like thinking and reasoning process to tackle de-novo problems in healthcare.
Retrieval Augmented Generation (RAG)	RAG is a sophisticated method that combines the pre-trained knowledge of LLMs with external, domain-specific knowledge. Conceptually a RAG system embeds data on multidimensional vectors and stores in a vector database to be retrieved and augment the LLM's output.[2]	Helps in transfer learning by training the LLM on health economic specific data including non-public information that can be queried from a database to enhance and refine the LLM's answers.
Model Fine-Tuning	Fine-tuning is a specialized technique where a pre-trained LLM undergoes additional training using targeted datasets or hyperparameter tuning to refine its capabilities for specific tasks. Fine tuning might also take the shape of instruction tuning, using high-quality instruction-response pairs.[2]	LLMs fine-trained on clinical data with tuned hyperparameters that are learned from the data itself improve predictions of clinical outcomes and outperform general models, enhancing accuracy in healthcare research and decision-making.
Domain-Specific Foundation Models	Specifically programmed LLMs that have been "pre trained" on HEOR related data only as opposed to general use LLMs.[2]	Their parameters and algorithms are specifically set to handle HEOR based problems. They can act as chatbot assistants in solving complex HEOR problems.

Strengths and limitations

Strengths of generative AI:

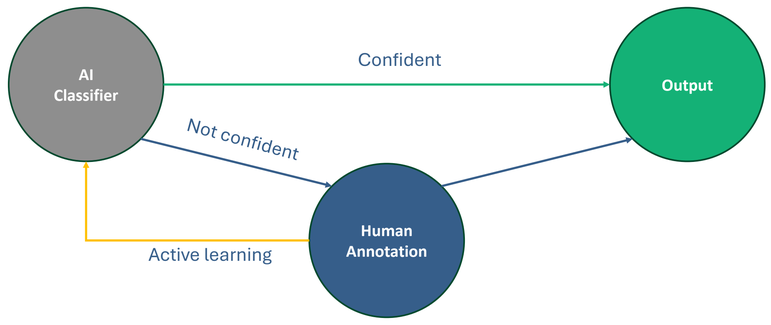
- **Advanced capabilities:** Can create new content, synthesize data, and provide innovative solutions to complex problems.
- **Continuous improvement:** Models are continually improving, offering increasing accuracy and efficiency in HEOR tasks.

Limitations of generative AI in HEOR:

- **Bias, equity, and ethical concerns**: AI models may reinforce biases present in training data, leading to unequal healthcare outcomes.
- **Adoption, integration, and regulatory challenges**: Integrating AI into HEOR workflows faces technical and organizational hurdles, including resistance, lack of expertise, high costs, and the need to comply with regulations.

Conclusions

- Emphasizing Gen-AI as a tool to assist, not replace, human capabilities, alongside continuous capacity building of AI is the way forward.



- AI-in-loop framework is preferred when expert judgment and a deeper understanding of complex healthcare issues are needed such as analyzing patient behavior, healthcare policies, or regional differences in health outcomes. In these cases, human expertise helps make sure the results are meaningful and aligned with real-world situations while AI assists with minor, mechanical tasks to save time.
- HI-in-the-loop is more suitable when handling large amounts of data or identifying patterns quickly is the main goal. For example, AI can help analyze big datasets to find trends or predict outcomes, while humans can step in to validate the findings and make final decisions.

References.

1. Memarian B, Computers in Human Behavior: Artificial Humans, 2024;2(1).
2. Fleurence RL. Generative AI in Health Economics and Outcomes Research: A Taxonomy of Key Definitions and Emerging Applications - an ISPOR Working Group Report, 2024.
3. Wu, X., IEEE Transactions on Human-Machine Systems, 52(2), 222-233.

Acknowledgments

The authors acknowledge Shashwat Gaur and Hanan Irfan (ConnectHEOR, Delhi, India) for developing and designing the content of the poster.

Financial Disclosure

The authors are employees of ConnectHEOR Limited and no external funding was received to conduct this research. The authors have no conflict of interest to declare. **Manuscript in writing.**

Poster presented at the ISPOR EU 2024, November 17-20, Barcelona, Spain