

LLM-driven Motivational Interviewing for Health Behaviour Change

Akshaye **Shenoi**, MSc; Tianze **Li**, UG, Singapore-ETH Centre; Ahmad **Jabir**, MSc, Singapore ETH-Centre; Elgar **Fleisch**, PhD, ETH Zurich; Tobias **Kowatsch**, PhD, University of Zurich, University of St.Gallen & ETH Zurich; Jacqueline **Mair**, PhD, Singapore-ETH Centre

Context. Conversational Agents (CAs), or chatbots, are increasingly used to deliver digital behaviour-change interventions. Many such CAs typically rely on strict rule-based dialogues crafted by domain experts to ensure that interactions with the user occur in a controlled and predictable manner. While this approach ensures therapeutic precision, it limits conversational flexibility with respect to an individual's unique needs and values. For instance, psychotherapeutic frameworks such as Motivational Interviewing (MI) emphasise reflective listening to elicit change talk and express empathy. Rule-based CAs are constrained by predefined scripts, making it infeasible to emulate such strategies. Advancements in Large Language Model (LLM) based CAs have shown considerable promise in digital psychotherapy given their ability to generate dynamic and contextually relevant responses. This capability facilitates more authentic and meaningful interactions that align with MI principles. Despite these advantages, the deployment of LLM-based CAs faces challenges. First, LLMs are not explicitly trained to apply psychotherapeutic frameworks such as MI systematically. While engineering prompts or fine-tuning models may offer a solution, their effectiveness has not yet been rigorously evaluated. Second, LLMs can occasionally generate irrelevant or even harmful content due to their "black box" nature. Given the sensitive settings in which these CAs operate, safety and ethical considerations are essential.

Methods. We present AIMI, an LLM-based CA that aims to deliver behaviour-change interventions using MI. AIMI follows an *agentic* design pattern that translates MI principles into a natural language capable CA. An agent is a specialized instance of a language model designed to perform a specific task, such as selecting the appropriate MI process (e.g., engaging or focusing) based on the conversational context. We encode distinct tasks into multiple agents so that each agent's function aligns with key components of the MI therapeutic approach. The agents then collaborate to produce systematic, cohesive, and contextually appropriate interactions. To improve safety, a guard agent further reviews each response to filter out harmful content.

Results. Preliminary evaluations suggest that AIMI can deliver an effective behavioural intervention to increase physical activity. We find that it generates contextually appropriate and empathetic responses with a systematic application of MI techniques such as reflection and affirmations. We also observe that our agentic design consistently outperforms simple prompt engineering approaches. Ongoing experiments will evaluate AIMI with respect to MI fidelity, linguistic quality, and users' motivation to change. Findings will be discussed alongside the challenges and opportunities for implementation and scaling in both clinical and non-clinical settings.