

## **Research Strategy**

### **A. SIGNIFICANCE**

Alzheimer's disease and related dementias (ADRD) is a world-wide disease yet difficult to treat, because available pharmacologic options are limited by side effects, high cost, and restricted eligibility for newer disease-modifying therapies<sup>1</sup>. As millions of patients remain without accessible treatment options, it was estimated that up to ~40% of AD risk may be attributable to modifiable factors. For example, epidemiological research has associated limited social interaction with greater risk of cognitive decline and dementia<sup>3–8</sup>. Meanwhile, it was estimated that increasing social interaction could approximately prevent 4% of dementia cases<sup>15</sup>, motivating the development of behavioral interventions targeting limited social interaction<sup>16–19</sup>.

**Conversational intervention was proven effective, but its scalable delivery remains challenging.** Recently, the Internet-based conversational intervention, I-CONECT (NIH-funded clinical trial NCT02871921) which uses frequent one-to-one semi-structured conversational interactions via Internet, was proven to be effective in mitigating cognitive decline among older adults with limited social interaction and mild cognitive impairment (MCI)<sup>10</sup>. However, its key limitation, as a foundation for public health impact, is a lack of feasibility for wide dissemination: the intervention relies on well-trained interviewers to deliver, which is workforce-intensive and costly. Scaling the intervention requires recruiting, training, and supervising staff, plus complex scheduling and quality assurance, which can increase costs and reduce delivery consistency and quality.

**Conversational AI provides a scalable alternative but is under explored.** In response to this challenge, the emergence of conversational AI driven by Large Language Models (LLMs) that can chat fluently at low cost and be replicated with consistent quality<sup>11</sup> has fostered new opportunities for scalable conversational intervention. The estimated cost reduction is significant: based on OpenAI GPT-4o pricing (July 2024<sup>20</sup>), a twice-per-week, one-year service would cost 0.3% of an estimated human-delivered service cost, based on 2024 salary estimates<sup>12,21</sup>.

**Despite the promise, most currently available conversational AI systems lack a scientific basis and an evidence-based, efficacy-proven approach to address cognitive decline by executing an intervention as an interviewer.** Prior chatbot-based intervention studies in older adults have shown feasibility and acceptability as an assistant (rather than an interviewer or therapist) in web-based problem solving, daily cognitive training, and simple memory training<sup>22–26</sup>. As an early attempt, our prior work customized LLMs to follow the NIH-funded efficacy-proven I-CONECT conversational interaction protocol via prompt engineering<sup>12</sup>; however, it has multiple limitations: (i) The single-modality conversation has limited interactivity and thereby lowers engagement compared to human-led ones; (ii) Older adults face technical barriers to operating computers and accessing the AI service, which may lead to low adherence; (iii) The intervention protocol was not adapted to AI, ignoring the unique challenges of AI-based intervention, for example, the potential safety risks and the protocol drift (failure to maintain protocol fidelity) during conversations<sup>27,28</sup>. When older adults have limited social interaction and are less likely to have been trained to handle such hazards, concerns about the reliability and safety of AI-based intervention may be heightened. (iv) The system has not yet been evaluated in the target population; therefore, the feasibility of AI-based intervention remains uncertain.

**AIMI-CONECT aims to translate I-CONECT to scalable delivery.** This project directly addresses the scalability barrier by translating the evidence-based I-CONECT conversational protocol into an AI-delivered system with multimodal interaction (AIMI-CONECT) designed for older adults. If successful, this project will shift the field from workforce-constrained conversational interventions to accessible, and scalable delivery. Technically, Aim 1 will directly address key technical challenges and therefore advance the state of AI-based conversational interventions. Aim 2 will help understand whether AI-based conversational intervention is feasible in older adults with limited social interaction (the population group that benefits most from this type of behavioral intervention<sup>10,29</sup>) and MCI or normal cognition. *Clinically*, it will pave the way for the use of AI in future behavioral interventions.

## B. INNOVATION

**From human-delivered to AI-based conversational intervention.** AIMI-CONECT translates an efficacy-proven, interviewer-led dementia prevention intervention into a protocol-faithful, scalable AI-delivered system tailored for older adults with limited social interaction. This systematic translation adapts the I-CONECT protocol for AI delivery, addresses emergent safety risks, and evaluates feasibility to enable high-fidelity automation of future intervention delivery.

**Multimodal interaction enabled conversational AI system improving usability and engagement.** To mitigate the gap between human and AI-led interaction, particularly the gap causing low engagement for older adults, AIMI-CONECT integrates multimodal interaction into the conversational system: (1) a voice-forward interface (real-time speech input/output); (2) visual affective robot; (3) user-AI co-recreation of reminiscence 3D scene via conversational instructions. Compared with conventional chat interfaces, the compound design lowers operational demands enhancing usability, and enables more engaging conversations aligned with the I-CONECT goals. Unique to this work, the 3D scene recreation allows the participants to actively engage in the reminiscence as well as vocabulary enhancement (via describing personally relevant items in a scene), which are key therapeutic components of the I-CONECT protocol, and thereby may strengthen the cognitive stimulation effects.

**Enhancing protocol compliance of AI via learning from clinical trial grounded conversation simulation.** We leverage reinforcement learning from human feedback (RLHF)<sup>11</sup> to align the LLM-based chatbot to the I-CONECT protocol. Yet, when targeting the specific population (e.g., older adults with limited social interaction and mild cognitive impairment (MCI)), traditional RLHF is limited by scarce data and high cost to collect human feedback. Thus, in the project, we introduce a data-driven simulation framework that can generate conversations between LLMs and virtual users that are LLMs fine-tuned on recorded conversations from the I-CONECT clinical trial. Unlike previous simulation methods<sup>30</sup>, our method is grounded in clinical-trial data (enabling us to capture unique MCI traits). This approach extends existing AI development methods by adding a quantitative, intervention-relevant virtual evaluation environment, thereby enabling rapid iteration and earlier identification of risks.

## C. APPROACH

**Overall Strategy.** We will develop an AI-based conversational system (AIMI-CONECT) to implement the AI-adapted I-CONECT intervention protocol and evaluate its feasibility and satisfaction among older adults with limited social interaction and MCI or normal cognition.

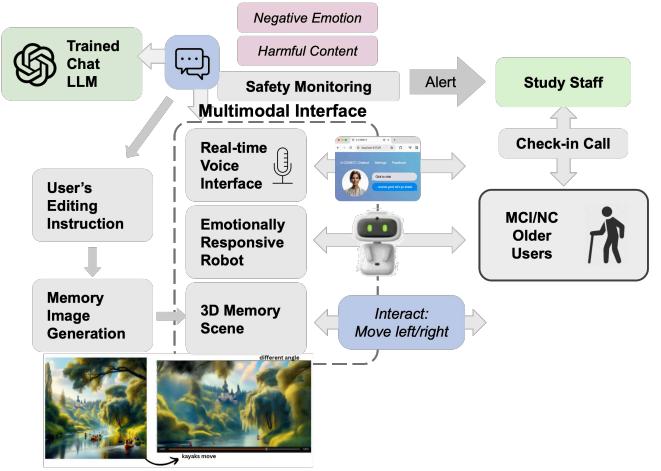
### C.1 Intervention Protocol with Safety Assurance

The intervention protocol is adapted from I-CONECT<sup>10</sup> to fit the AI-based intervention. Each session will last 15 minutes and follow the conversation flow, starting with a greeting and topic selection from three candidates within a randomized theme, followed by open conversation in which reminiscence and vocabulary enhancement are executed spontaneously. (1) *Image and Language Prompting for Critical Thinking*. We adopt the same topics and cognitively-demanding language prompts as I-CONECT to stimulate participants' interests and cognitive function in conversation. Based on the prompts, images generated by AI are used as visual prompts to strengthen the participants' engagement in the topic selection stage. (2) *Reminiscence with Visual Interactions*. We will engage participants in reminiscence by asking about their personal experiences related to the presented images and topics. With the AI tools, the participants will engage in generating personalized 3D scenes based on the conversation and AI memory from prior sessions. (3) *Vocabulary Enhancement with Hints*. We will prompt the participants to describe the presented images in detail to enhance their vocabulary usage. AI will guide the participants to describe specific details by providing visual hints, e.g., zooming in on image parts. (4) *Safety Assurance*. We add safety assurance to define prohibited behaviors that should be avoided in conversations, for example, illegal content or financial advice. We also flag suicidal tendency using a protocol similar to the one used in the I-CONECT study.

### C.2 System Design with Multimodal Interaction.

The system is composed of three main modules: (1) A trained Chat LLM that is responsible for handling the conversation in text; (2) A multimodal interface that routes input from users (MCI/NC older users) to the Chat LLM; (3) Backend processing includes safety monitoring, and 3D generation that will be executed asynchronously to avoid blocking real-time interaction.

**(1) Voice interface.** To enhance the accessibility of the chatbot for older adults, we will implement a voice-forward interface that enables natural spoken interaction. Specifically, user speech will be captured through an on-device microphone and converted to text using streaming speech-to-text (STT), so the dialogue manager can apply the I-CONECT conversation strategies in real time. The chatbot's responses will be produced as text and then rendered as natural-sounding speech via text-to-speech (TTS), allowing participants to engage without typing or reading long on-screen messages. We will primarily use the OpenAI Real-time API to support low-latency, bidirectional audio streaming and automatic turn-taking (i.e., detecting when the participant has finished speaking and when the system should respond). This capability is important for older adults because it reduces the need for explicit controls such as “push-to-talk” buttons and minimizes conversational interruptions due to delayed responses. We will also log time-stamped transcripts (from STT) and system audio outputs (from TTS) to support protocol compliance checks and safety monitoring while avoiding storage of unnecessary identifiers.



**(2) 3D Memory Scene Collaborative Recreation.** To enhance reminiscence and engagement, we will implement a user-AI collaborative 3D scene recreation system. Our preliminary work showed that 3D scene creation can improve participants' interest in using a reminiscence chatbot<sup>31</sup>. During conversations, the chatbot will prompt participants to describe specific details of their memory associated with the presented images (e.g., “Can you tell me what this picture reminds you of?”). Based on the participant's descriptions, the system will generate a personalized 3D scene that visually represents their memories and experiences related to the image. This process allows participants to actively engage in reminiscence while also providing visual cues that can stimulate further conversation and cognitive activity. The 3D scenes will be generated asynchronously to ensure that they do not disrupt the real-time flow of conversation.

**(3) Safety Assurance via Real-time Monitoring.** To manage safety risks arising during deployment, we will create real-time safety guardrails that leverage LLMs to monitor conversations and detect critical situations such as low emotion, suicide, misinformation, or financial risks. The guardrails will use generative rule-based methods, where a GuardAgent will generate guardrail rules to screen out unsafe generation in conversations<sup>32</sup>. The monitoring system will send an alert to study staff when critical issues are identified. Study staff in the loop work as an additional layer of protection via weekly check-in phone calls.

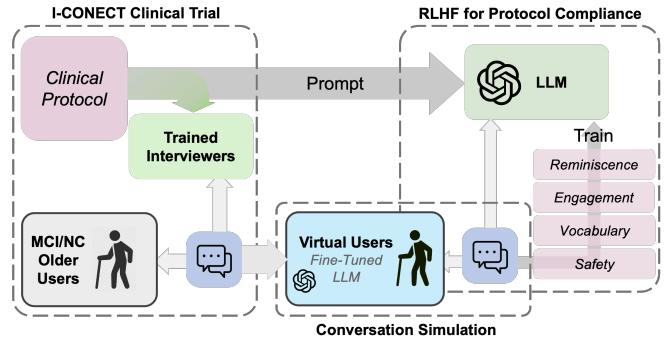
**(4) Device and Emotionally Responsive Robot:** The software will be developed using Node.js and pre-installed on a ready-to-go tablet, which will be delivered to participants. Alongside the tablet, we will develop an emotionally responsive robot, with real-time facial expressions to provide affective responses. The facial expressions are managed by the Chat LLM through function calling; the LLM will generate and send commands to the robot on demand. The connection between the facial robot and the tablet is established via Bluetooth, which can maintain a robust connection. The robot will be developed based on our consultant Dr. Chen's prior work<sup>13</sup>.

**(5) Data privacy and system security.** To ensure participants' data privacy and system security, we will implement the following measures. The OpenAI API will be served through HIPAA-compliant servers inside MGH, and all data will be encrypted in transit and at rest. We will minimize the collection of personally identifiable information (PII) and use de-identified data for analysis. Access to the system will be protected by secure authentication, and

regular security audits will be conducted to identify and mitigate vulnerabilities. Additionally, we will provide clear information to participants about data usage and obtain informed consent that explicitly addresses privacy and security measures.

### C.3 Optimization Algorithms for Improving AI Protocol Compliance.

We will use RLHF to optimize the chatbot to follow the protocol<sup>11</sup>. The reward model, that evaluates the LLM compliance in conversation, will be constructed based on the I-CONECT protocol. However, obtaining conversation data is expensive. Thus, we will build virtual users to simulate conversations.



**(1) Learning to Simulate Human Feedback from I-CONECT Data.** RLHF relies on conversations between the chatbot and real users to collect feedback. However, collecting feedback from real users is costly and risky, especially when the users are older adults with limited social interaction and MCI. To reduce cost and risk, we will create virtual users that are LLMs fine-tuned on older adults' recorded conversations and then simulate conversations with older adults with limited social interaction to quantitatively evaluate the chatbot. The simulation will learn I-CONECT participants' language traits that may drive the engagement design. To provide a concrete simulation of older adults with MCI, we will validate that the virtual users can generate similar language symptoms that have been proven effective for MCI diagnosis<sup>33</sup>.

**(2) Optimizing AI for Protocol Compliance via Learning from Verifiable Feedback.** We will customize Large Language Models (LLMs) to implement the I-CONECT intervention strategies. We will first engineer prompts that instruct LLMs to carry out predefined tasks, including proposing chat themes and topics, and presenting topic-related images. To enhance LLMs' capability to engage users in 15-min conversations, we will improve conversation engagement using heuristics and data from the I-CONECT trials. On the one hand, we will consult interviewers and related researchers (from the I-CONECT team) to provide heuristic conversation strategies and encode the heuristics in the instructions for LLMs. On the other hand, we will *supervised fine-tune* the LLMs to learn spontaneous language patterns, for example, the active use of filler words, from interviewers' speech in the I-CONECT conversation data. The conversation data will be processed from audio recordings, and we will develop speaker-aware Automatic Speech Recognition to extract high-quality text data for LLM fine-tuning. Then, we will train LLMs to learn from the *feedback* on compliance in conversations. Each time the LLM has a conversation with a user (participant), we will use a rule-based method to rate protocol compliance (engagement time, reminiscence-induction frequency, and user vocabulary). The ratings will be used to train a reward model that guides the LLMs to generate protocol-compliant conversations via Reinforcement Learning<sup>11,34</sup>.

### C.4 Approach for Feasibility Study.

We will test the chatbot in older adults with MCI diagnosis or normal cognition recruited from the Memory Division at Mass General Brigham (MGB, formally known as the Massachusetts General Hospital and Brigham Women's Hospital) and the Massachusetts Alzheimer's Disease Research Center (MADRC) (**see Dr. Gomez-Isla's letter of support**). 40 older adults (20 MCI, 20 with normal cognition) will be recruited to complete feasibility/satisfaction surveys on their overall experience in terms of adherence, happiness, emotional lifting, empathy, and trustworthiness (e.g., reporting harmful or hateful biases). The survey will also provide insights into how to design the chatbot to engage humans and increase social interactions.

Each participant will receive 15-min conversations with our chatbot for 24 sessions in 6 weeks (4 sessions per week). A staff member will conduct weekly 10-min check-in phone calls to assess emotional status and potential risks or any technical challenges in using AI. In 6 weeks, we will collect both objective data and subjective survey responses to assess feasibility and satisfaction.

To isolate the effects of social engagement from confounding factors and enhance cognitive stimulation, the study

incorporates three key design elements. (1) First, we employ persona rotation (via prompting LLMs) across participants to enhance interaction variability and increase conversation novelty. Personas will be drawn from a pool initiated from I-CONECT interviewer profiles. (2) Our intuitive interface and natural-sounding voice capabilities enable participants with limited Internet or webcam proficiency to participate in conversations without additional cognitive burden from learning new technology. (3) Third, we leverage the standardized daily themes and image stimuli from I-CONECT to support participants in organizing and expressing their ideas within conversations<sup>10</sup>.

**Participant Recruitment.** We will recruit a total of **40 older adults with limited social interactions aged  $\geq 75$  years** from the MGH Memory Division and MADRC. The cohort will consist of two balanced groups: **20 individuals with mild cognitive impairment (MCI)** and **20 cognitively normal (NC)** participants. We adopt the same operational inclusion/exclusion criteria used in the original NIH-funded I-CONECT study.

**Inclusion criteria:** (1) Aged 75+ (In the original I-CONECT study, this age group was selected due to an increased risk of reduced social interaction. We will use the same age group for this pilot feasibility study.); (2) Cognitive status assessed via the Telephone Interview for Cognitive Status (TICS)<sup>35</sup> is MCI or NC (1:1 ratio); (3) Identified as socially isolated by one of: (a) Score  $\leq 12$  on the 6-item Lubben Social Network Scale (LSNS-6)<sup>36</sup>; (b) Engages in conversations lasting 30 minutes or longer no more than twice per week, per participant self-report; (c) Answering “Often” to at least one question of the three<sup>37</sup>: “How often do you feel that you lack companionship?”, “How often do you feel left out?”, “How often do you feel isolated from others?” (4) GDS-15 (15-item Geriatric Depression Scale) at or below 9 (not severely depressed)<sup>38</sup>; (6) Ability to understand the research consent form.

**Exclusion criteria:** (1) Diagnosed with dementia such as AD, ischemic vascular dementia, normal pressure hydrocephalus, or Parkinson’s disease. (2) Schizophrenia or other major psychiatric disorder defined by DSM-IV criteria<sup>39</sup>. (3) Medications: frequent use of high doses of analgesics; use of sedative medications except for those used occasionally for sleep ( $\leq 2$  per week); unstable dosing of cholinesterase inhibitors (must be stable for 2 months).

## D. DATA PROCESSING AND ANALYSES

**Aim 1:** We will use I-CONECT recorded conversations to optimize AI for better compliance with the I-CONECT protocol.

**Method: Processing I-CONECT Data and Constructing Virtual Users.** We will use the recorded I-CONECT audio conversations to build virtual users of the corresponding participants. All data will be processed on HIPAA-compliant servers with personal identifiable information removed. The whole process includes these steps: (1) *Transcription*: Audio will be transcribed using a speaker-aware ASR optimized for older adult speech, separating interviewer and participant utterances. (2) *Feature extraction*: We will extract lexical features (e.g., word counts, lexical diversity) from transcripts. (3) *Virtual user construction*: We will create LLM-based “virtual users” of older adult participants by fine-tuning models on transcribed conversations to simulate their language patterns and engagement behaviors. The fidelity of the virtual users will be validated by comparing simulated conversations to real ones, ensuring that key lexical features are preserved.

**Aim 2:** To evaluate the feasibility of using AI as an alternative to human-led delivery, we will recruit a total of **40 older adults with limited social interaction aged  $\geq 75$  years**, including **20 individuals with mild cognitive impairment (MCI)** and **20 cognitively normal (NC)** participants.

**(1) Adherence.** The primary feasibility outcome is **adherence rate** to the intervention, defined at the participant level as completion of  $\geq 80\%$  of scheduled AIMI-CONECT sessions. A total sample size of N=40 allows estimation of the overall adherence ratio with adequate precision. For example, assuming an expected adherence rate of approximately 80%, N=40 yields a 95% confidence interval with a half-width of approximately  $\pm 12\%$ , which is sufficient to determine whether adherence meets the predefined feasibility threshold.

In addition, **recruitment ratio**, ratio of contacted to enrolled participants, will be collected during the recruitment process, reflecting people’s willingness to use AI. **User engagement level** will be measured by the participant’s

word ratio in a conversation, which is the proportion of words spoken by the participant relative to the total words in the conversation. This metric provides insight into how actively participants are engaging with the chatbot during sessions. For example, in the original I-CONECT study, interviewers were trained and standardized to ensure that participants contributed at least 30% of the total words spoken.

(2) **Acceptability** of the AIMI-CONECT intervention will be evaluated using the **Client Satisfaction Questionnaire-8 (CSQ-8)**<sup>14</sup>, a validated measure of user satisfaction for health and behavioral interventions, with total scores ranging from 8 to 32 and higher scores indicating greater satisfaction. Acceptability will be defined a priori using explicit, quantitative criteria, such that the intervention will be considered well accepted if the **mean CSQ-8 score is  $\geq 24$ , at least 70% of participants achieve CSQ-8 scores  $\geq 24$ , and no more than 15% of participants score  $\leq 20$** , indicating the absence of systematic dissatisfaction. A total sample size of N=40 is sufficient for this purpose because it provides adequate precision. Specifically, if approximately 70% of participants meet the satisfaction threshold ( $\text{CSQ-8} \geq 24$ ), a sample of 40 yields a 95% confidence interval with a half-width of approximately  $\pm 14\%$ , allowing us to *distinguish between unacceptably low satisfaction* (e.g., < 55–60%) and satisfaction levels consistent with good acceptability.

We will also collect **descriptive surveys** to understand participants' **concerns and preferences** for the different components and security in the system design, to inform future design.

**Exploratory Outcomes.** In addition to the score in the whole study group, the **subgroups** (20 MCI and 20 Normal Condition participants) ensures that feasibility (adherence and acceptability) can be assessed across two cognitively distinct but relevant subpopulations, allowing identification of group-specific implementation challenges (e.g., differential session completion, technology burden, or need for support).

Other outcomes will be collected as supplementary information to inform future design and trials. Cognitive assessments, obtained from the Memory Division of MGB and MADRC using the date closest to study enrollment, will be collected for exploratory and descriptive purposes only.

## E. POTENTIAL RISKS AND ALTERNATIVE STRATEGIES

We prepared alternative strategies to address potential risks that may arise during the project: (1) **Poor environmental conditions.** In real-world home conditions (e.g., intermittent connectivity, background noise, device microphone/speaker failures, or latency that disrupts turn-taking), usability and session completion could be reduced. As alternatives, we will implement a dynamic conversation flow that can operate with lower bandwidth (longer pauses between turns allowing more time for caching data). If the environment is not suitable (diagnosed automatically by a self-troubleshooting in our software), we will provide visual reminder to guide the participants to find a workaround (e.g., move to a quieter room). (2) **Device malfunctioning.** Devices could be lost, damaged, or malfunctioning during the study, which could lead to missing data and reduced adherence. As backup, we will maintain spare devices for swap-out, provide rapid remote troubleshooting, and implement regular device health checks to identify issues early. (3) **LLM/API server downtime or rate limiting.** Dependence on third-party API services could lead to interruptions in chatbot availability due to server downtime or rate limiting, which could affect session completion and user experience. As alternatives, we will monitor API status continuously, implement caching strategies to handle temporary outages, and maintain a contingency plan to switch to alternative services.

## F. CONCLUSION

In summary, this R21 will translate the evidence-based I-CONECT conversational protocol into an AI-delivered, multimodal system designed for older adults with limited social interaction. Led by an interdisciplinary team spanning neurology, biostatistics, gerontology, and computer science, we will develop the AI system and establish feasibility to inform future scale-up and dissemination. This feasibility study represents a necessary preliminary step before clinical trials evaluating the clinical efficacy of AIMI-CONECT.