

Multimodal Cognitive Reframing System: A Multi-Agent Approach for Mental Health Support

Thuc Kevin Nguyen
thuc.nguyen.001@student.uni.lu
University of Luxembourg
Belval Esch-sur-Alzette, Luxembourg

Salima Lamsiyah
salima.lamsiyah@uni.lu
University of Luxembourg
Belval Esch-sur-Alzette, Luxembourg

Abstract

This project focuses on developing a multi-agent system designed to assist individuals in reframing negative thoughts into more positive, constructive perspectives. Cognitive reframing, a key technique in cognitive behavior therapy (CBT), will be used to help users challenge unhelpful thought patterns and promote healthier thinking. The system uses multimodal inputs, text, image and voice, which are processed by specialized agents to deliver personalized and adaptive feedback. The multi-agent system is split into a text-based cognitive reframing agent, a voice based emotional state detection agent, an image based visual context recognition agent and a multi modal agent which incorporates the gained insights from the previous three agents. Each agent use large language models for their insights. The system tries to counter these barriers, which include stigma and reluctance for professional help, by providing anonymous and accessible support. It does not replace psychotherapists but provides another tool that might enable users to make their first steps toward better mental health.

Keywords

Mental Health, Mental Health Support, Psychotherapy, Cognitive Reframing, Multi-Agent System, Multimodal Inputs, Large Language Model, Natural Language Processing, Sentiment Analysis, Chatbot

1 Introduction

Mental health problems have become more common in today's society, with increased work stress, strained relationships, and the pressures of modern life. According to the World Health Organization, more than one in eight people, or around 970 million people, suffers from some forms of mental disorder globally, of which anxiety and depression are considered the most prevalent diseases[12]. These numbers represent only those individuals who have been detected, and it remains unknown how many people face mental health issues without proper treatment.

While more and more of these issues are brought to light, most of the people who are afflicted are reluctant to seek professional treatment. According to statistics from the National Institute of Mental Health (NIMH), 59.6% of adults with mental illness did not receive mental health services in the previous year[5]. Major depression might be expected to afflict at least one of every three women and one of every five men in their lifetimes. Other disorders, such as schizophrenia and bipolar disorder, are far less common, although their impact on patients' lives is significant.

All too often, people are unaware or ashamed to ask for professional services due to stigma, personal concerns about social

judgment, and the possible loss of jobs. If left untreated, the symptoms of mental health problems could worsen, along with physical health problems and strained relationships. They also lead to an increased risk of substance abuse, self-harm, and generally a low quality of life. As such, early intervention is crucial to averting these possibilities in the long term.

Stigma can be blamed on the misrepresentation of mental illnesses. For example, people with mental illnesses may be stigmatized as dangerous, incompetent, blamed for their disorder, or unpredictable.

This system will address these needs by using AI multi-agent system combined with multi-modal input analysis and cognitive reframing to provide anonymous, accessible support to those who are unwilling to seek professional help. Importantly, it does not aim to replace licensed psychotherapists but instead serves as a tool, offering low resistance and quick support whenever needed.

2 Scientific Background

2.1 Multi-Agent System

Multi-agent systems are part of distributed artificial intelligence, alongside parallel AI and Distributed Problem Solving. They are useful when problems are too large or too complex to handle with a single agent. It is a computational architecture composed of multiple interacting agents within an environment. Each autonomous agent perceives its environment and acts on it independently in pursuit of its goals, based on its capabilities and knowledge. Agents can interact with other agents which can be collaboration, information exchange, or coordinate actions to achieve global objectives. Multi-agent system represents a decentralized, scalable and flexible approach to solving problems that require diverse expertise, distributed resources, or simultaneous problem-solving efforts. [2]

For this project, which involves performing cognitive reframing based on multimodal inputs such as text, audio, and images. Each agent is designed to handle a specific subtask within this process, leveraging large language models as their computational framework. Large language models enable advanced reasoning, contextual understanding, and decision-making capabilities necessary to analyze and reframe input data effectively.

2.2 Cognitive Reframing

Cognitive-behavioral therapy (CBT) is a common and evidence-based approach to treat various mental health problems including depression, anxiety, PTSD, addictions, violence, eating disorders, bipolar and psychotic behavior. The goal of CBT is to identify and challenge distorted thoughts, beliefs, and emotions that contribute to emotional distress or problematic actions. The theory of CBT

is that the individual's mental state improves by changing those problematic patterns, because thoughts, feelings, and behaviors are interconnected. Concepts of CBT:

- (1) Thoughts, emotions, and behaviors are interconnected and impact each other.
- (2) Perception is subjective and influenced by individual beliefs. It is assumed that how the individual perceives the world is the real world for this individual.
- (3) Core beliefs have a major influence on how an individual perceives the world and shapes thoughts, emotions, and behaviors at all times.
- (4) CBT assumes that thoughts influence the emotion of an individual, as thoughts come before emotions. Logical or rational thinking can change how the individual feels and acts. An embodied approach emphasizes that balancing thought, feelings, and behavior leads to better handling of situations.
- (5) Thoughts and feelings are affected by current events. But also it enables the individual to use past experiences to make decisions which are decided by the individual's forecast of a possible future rather than how the individual perceives the present moment.
- (6) Reflection and learning are implied in this process; they are not always made explicit. Incorporating explicit reflection and learning can enhance the process of CBT.

This approach entails the identification of negative or maladaptive thoughts and the reinterpretation of these in a more balanced or constructive manner. The purpose of this is to change perspectives in such a way that it decreases stress, improves coping mechanisms, and develops a positive outlook. [4]

2.3 Large Language Model for Therapeutic Deployment

LLMs hold potential in psychotherapy and demonstrate capabilities in understanding and generating natural language due to their extensive training in diverse datasets. They are capable of engaging in coherent and contextually relevant dialogue, which can be beneficial for therapeutic-style conversations.

Although traditional LLMs can express pseudo-empathetic responses and logical suggestions, inconsistency in actually being empathetic and providing therapeutic insight is lacking in their responses. It also generates responses that may be too generic and fail to fully capture the emotional state of the user or help the user steer toward healthier thought patterns.[13]

2.4 Prompt Engineering

Prompt engineering is an effective way to enhance a Large Language Model's performance without adding additional costs or requiring extensive further training. By carefully designing the prompts, it is possible to guide the LLM in specific tasks. While this approach improves the quality of the output, the model remains constrained by its existing capabilities [8]. However, this limitation is not an issue for the project's task, such as generating structured dialogue, analyzing the input, asking proactive questions and providing empathetic responses, are sufficient for this system. The project uses prompt engineering to guide its agents and improves

the quality of the agents output by adding role-playing. According to [10], with role-playing LLMs are able to effectively emulate various roles, demonstrating decision-making patterns that align with the sociological characteristics of these roles.

The paper [13] also suggests that prompt engineering can guide LLMs to perform better in therapeutic responses by providing a guideline for interactions. Prompts can be designed to encourage the models to ask proactive questions, a critical aspect of CBT. This involves helping the model focus on identifying key issues and guiding the patient toward self-reflection. [11] But only with prompt engineering, it cannot fully overcome the limitations of general-purpose LLMs.

2.5 Conversation Structure

The goal of this project is to build a system capable of having a therapeutic conversation that closely resembles a real interaction with a therapist. To achieve this, it is essential to understand and implement the methodologies of a therapeutic conversation. One key aspect of effective therapy is the ability to build trust and cooperation. Engaging in small talk can significantly enhance these dynamics by creating a sense of familiarity and comfort between individuals. [1] The paper [9] conducted a study about the effectiveness of empathetic paraphrasing. It showed that empathetic paraphrasing helps individuals feel heard and understood. The participants spoke in a calmer tone which suggest they felt more relaxed emotionally. Participants reported that they felt less negative compared to when the therapist just took notes. Treating the user with empathy shows that the therapist cares which builds trust and make the user feel supported in the situation. With paraphrasing the user can also reflect on their thoughts more deeply which could lead to progression in resolving the conflict. Limitations of these studies are the sample size of the 20 participants which limits the generalization of the study. However, empathic paraphrasing may help the user to gain clarity and reduce emotional distress, while also facilitating cognitive reframing of issues.

2.6 Voice Input

For the voice input, the project leverages OpenAI's Whisper architecture, a state-of-the-art automatic speech recognition (ASR) system introduced in the paper [6]. Whisper is a versatile and robust ASR model that has been trained on a diverse dataset of multilingual speech data, making it highly effective in accurately transcribing voice input, even in noisy environments.

The voice-based agent relies on Whisper to process recordings in which users speak candidly about their thoughts, feelings, and experiences, as if they are addressing a psychotherapist. These recordings often include emotionally charged content, such as reflections on personal challenges, relationships, or significant life events. Each word carries emotional meaning, and Whisper's ability to capture subtle variations in speech is critical for maintaining the integrity of the transcription.

Verbal therapy is the traditional method to do psychotherapy and it is a logical choice to add this modality.

2.7 Image input

The image input enables the user to upload images, drawings, or photos uploaded by the user. For the image input, the project uses OpenAI's DALL-E architecture [7], a state-of-the-art generative model capable of understanding and creating images from textual descriptions. DALL-E is primarily designed for image generation, its ability to understand and analyze complex visual data makes it an ideal choice for the image-based agent. DALL-E can interpret images for emotional cues by analyzing features such as color tones, composition, and depicted subjects or objects.

The therapeutic efficiency of drawing and image creation has extensively been discussed in the context of Active Visual Art Therapy (AVAT), which focuses on an individual's active engagement in creative processes, such as painting, drawing, or shaping materials. Research shows that creative activity itself can have substantial therapeutic effects, especially in conditions with psychological or emotional disturbances.

Drawing and creating images offer a nonverbal area of expressing oneself, by which those who find it difficult to describe their feelings can outline their emotions and thoughts. Studies show that drawing as a method would bring improvements in mental health outcomes, for example reduction in anxiety and depression. For instance, AVAT has been linked to improving emotional well-being by offering a means to express and process complex emotions in a safe and controlled manner. This is particularly valuable for individuals who may struggle with conventional forms of verbal therapy.

The analysis of the images created by the patients is a critical aspect of art therapy. Therapists often interpret visual elements in the image, such as colors, shapes, composition, and project them into the emotional and psychological states of their clients. For instance, darker colors and chaotic or fragmented compositions may indicate feelings of distress or tension, while lighter colors and harmonious arrangements may reflect a sense of calm or optimism. Symbols and recurring themes in drawings can reveal underlying concerns, unresolved conflicts, or subconscious thoughts that the individual might not consciously recognize. [3]

3 Methodology

3.1 AI Therapist

The name of the application is "AI Therapist" which is a chat-based interface, where users can interact with a multi-agent system designed to simulate a therapeutic conversation. The application allows multi-modal input: audio, image and text. At least one of them is required. The user can either record a voice message, upload a drawing or give a textual input about their emotional state. The multi-agent system processes the input and engages in a conversation with the user as if the user would talk to a psychotherapist. Throughout the interaction, the system employs cognitive reframing techniques to help users reframe their thoughts and gain a new perspective on their emotions or challenges.

3.2 Architecture Structure

For the multi-agent system the project uses the open-source CrewAI framework. CrewAI allows to define a multi-agent system, comprising of autonomous agents each with distinct tasks and the

capability to utilize various tools. With the integration of different large language model APIs, for example OpenAI, Groq or Ollama, it provides flexibility and access to different advanced models.

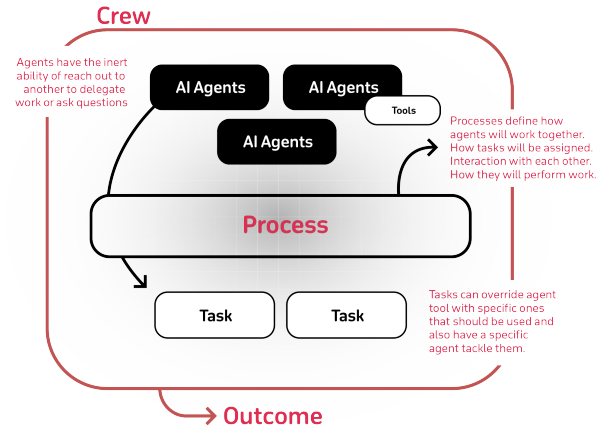


Figure 1: CrewAI Crew

One of its key features is to use the advantages of role-playing by defining a role, goal and a backstory for each agent, enhancing their decision-making and interaction processes. CrewAI achieves this through their specific prompt engineering for the large language models.

The crew consists of four agents. Three of them process the input and provide a structured report as output. Each agent is assigned a specific modality: text, image, or voice. The fourth agent is responsible for responding to the input by aggregating the reports from the different modalities and applying cognitive reframing techniques when necessary.

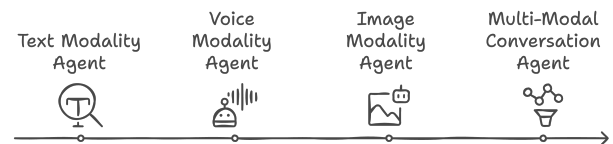


Figure 2: Synchronous Execution

By default, CrewAI runs the agents sequentially, which results in a high response time due to four API calls being made one after another. However, CrewAI also supports asynchronous execution. To maximize efficiency and maintain modularity, the input processing agents are designed to run asynchronously

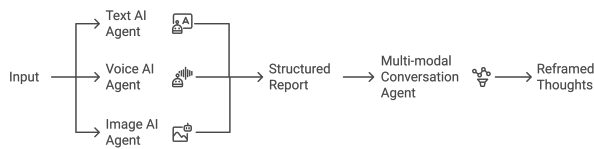


Figure 3: Asynchronous Execution

The role, goal and backstory are constructed similarly for every agent.

- Role: <Task of agent> Specialist
- Goal: Short summary of Task
- Backstory:
 - (1) PhD in Psychology and 20 years of experience in Psychotherapy
 - (2) Specialization in its task
 - (3) Short description of task and task-goal

3.3 Text-based Agent

The assigned role of the text-based Agent, also referred as the text therapist, is assigned the role as a cognitive reframing specialist. It's primary objective is to process with a large language model the textual user input and create a detailed report about the user's identified negative thoughts and cognitive distortions and suggestions on reframing those thoughts and guiding the user to healthier thinking patterns. This report is used by the multimodal agent.

Agent backstory:

- You hold a PhD in Psychology and have over 20 years of experience in psychotherapy.
- You specialize in analyzing textual inputs for therapeutic insights, focusing on identifying and reframing negative thoughts and unhelpful behavioral patterns.
- Your therapeutic approach is grounded in understanding the underlying purpose of each thought or behavior and transforming them into more constructive alternatives.

Cognitive Reframing Task:

- If no text is provided, do nothing and note that in the feedback document.
- For non-empty text, analyze the text to identify cognitive distortions and provide reframing suggestions:
 - (1) Detect emotions and identify cognitive distortions.
 - (2) If emotions are positive, suggest continuing with small talk and do not apply reframing techniques.
 - (3) For negative emotions, follow these steps:
 - (a) Identify the situation or issue the user describes.
 - (b) Question the validity of negative thoughts: Are there alternative, more balanced perspectives?
 - (c) Shift perspectives: How might someone else view this situation? Consider the broader context.
 - (d) Reframe the user's thoughts into constructive and balanced ones. Highlight potential lessons or opportunities.
 - (e) Reduce perceived severity: What's the realistic impact, and how could it be managed?

- (f) Apply gratitude or humor to lighten the perspective.
- (4) Encourage reflection: Does the user find the reframing constructive? Are there other thoughts to explore?

Expected output:

- A structured feedback document:
 - (1) Original text
 - (2) Analysis: Identified cognitive distortions, emotional state, and tone of the user.
 - (3) Suggestions: Reframed thoughts and actionable recommendations for the therapist to continue the conversation.

3.4 Voice-based Agent

The voice-based agent is the voice Therapist and his role is the audio emotion detection specialist. Using the Whisper model from OpenAI, it transcribes the raw audio file with the following prompt: "The audio is a recording of an individual speaking candidly about their thoughts, feelings, and experiences, as if addressing a psychotherapist. The content may include discussions about emotions, personal challenges, relationships, and life events. The tone is introspective and reflective, and the transcription should aim to capture the speaker's words as accurately as possible, maintaining the emotional nuances of the conversation." Even when the user pauses to reflect on their emotions and thoughts, the Whisper model, guided by this prompt, ensures the transcription remains accurate and preserves the emotional depth of the recording. The voice-based agent assesses the emotional tone of the transcription and suggests how to reframe this thought. Those insights and suggestions are noted in a feedback document.

Agent backstory:

- You hold a PhD in Psychology and have over 20 years of experience in psychotherapy, with a focus on analyzing voice data for therapeutic insights.
- You specialize in identifying negative thoughts and unhelpful behavioral patterns from transcriptions, aiming to understand their purpose and reframe them constructively.
- Your expertise lies in leveraging audio cues to provide actionable emotional and cognitive support.

Audio Emotion Insight:

- Analyze the audio file. If no audio is provided, note this in the report and conclude the task. Otherwise, proceed with these steps:
 - (1) Transcribe the audio using the VoiceTranscriptionTool. Verify and correct transcription errors.
 - (2) Analyze the transcription for emotional tone and cognitive distortions:
 - (a) Identify the situation or issue described by the user, if evident.
 - (b) Detect emotions and assess the overall tone.
 - (3) For negative emotions:
 - (a) Challenge negative thoughts: Are there more balanced or rational alternatives?
 - (b) Encourage perspective shifts: How might others view the same situation?

- (c) Reframe unhelpful thoughts into constructive interpretations, highlighting growth opportunities.
- (d) Reduce perceived severity by exploring realistic outcomes and coping mechanisms.
- (e) Suggest gratitude reframing or humor where suitable.
- (4) Conclude with reflective prompts to encourage the user's engagement with the reframed insights.

Expected output:

- A structured feedback document:
 - (1) Transcription of the audio.
 - (2) Emotional analysis: Identified cognitive distortions, emotional tone, and potential user issues.
 - (3) Suggestions: Reframed interpretations and actionable recommendations for further discussion for the therapist.

3.5 Image-based Agent

The image-based agent in its role of the visual context recognition specialist, gives a detailed description of the user's provided image with the help of OpenAI's Dall-E. By analyzing abstract patterns, specific objects, symbolic elements, colors and composition, the agent provides insights into the individual's emotional state. Based on them the agent constructs reframing suggestions and provides his insights and suggestions in the feedback document.

Agent backstory:

- You hold a PhD in Psychology and have over 20 years of experience in psychotherapy, with a specialization in image-based therapeutic methods.
- You excel in visual context recognition, describing images in detail, identifying emotional content, and linking it to the user's context.
- You aim to apply cognitive reframing to the emotions and thoughts inferred from the image, providing actionable insights to support mental well-being.

Visual Context Recognition Task:

- Analyze the image. If no image is provided, note this in the report and conclude the task. Otherwise, follow these steps:
 - (1) Provide a detailed description of the content in the image.
 - (a) The input is an image created or selected by an individual, potentially reflecting their thoughts, emotions, or experiences.
 - (b) The image may include abstract patterns, specific objects, or symbolic elements.
 - (c) Pay attention to features such as colors, composition, and recurring patterns, as these may provide insights into the individual's emotional state.
 - (2) Analyze the visual elements for emotional content:
 - (a) Identify emotions conveyed by key features in the image.
 - (b) Assess the overall emotional tone (e.g., positive, negative, neutral).

- (3) If a specific user situation or issue is given, relate the emotional analysis to it.
- (4) Suggest possible reframing techniques:
 - (a) Question the validity of any inferred negative thoughts.
 - (b) Explore alternative perspectives and broader contexts.
 - (c) Reframe unhelpful interpretations into balanced and constructive ones.
 - (d) Identify potential lessons or opportunities from the image's emotional cues.
 - (e) Use gratitude reframing or humor if suitable.
- (5) Encourage reflection: How might this analysis help the user perceive their situation differently?

Expected output:

- A structured report:
 - (1) Detailed description of the image.
 - (2) Emotional analysis, including insights into user-related context, if any.
 - (3) Suggestions for reframing techniques, tailored to the visual emotional content for the therapist to continue the conversation..

3.6 Multimodal and Conversational Agent

This agent is designed to engage in a therapeutic conversation with the user, by combining all insights from various modalities analyzed from the previous agents. The agent develops a holistic view on the user's emotions and mental state and gives an empathetic response to guide the user to healthier thinking patterns.

Agent backstory:

- You hold a PhD in Psychology and have over 20 years of experience in psychotherapy.
- As a multimodal therapist, you excel at integrating information from various sources (text, image, audio) to form a comprehensive understanding of the user's emotional and cognitive state.
- You aim to identify negative thoughts and maladaptive behavioral patterns, reframing them into positive alternatives through empathetic conversation.
- Your primary focus is on fostering a collaborative and supportive therapeutic relationship with the user, helping them develop healthier thoughts and behaviors.

Multimodal Conversation Task:

- Look at all the reports of previous agents and also understand the text input, the voice transcription and the image description.
- Engage with the user to foster a meaningful and supportive conversation based on the provided input. Follow these steps:
 - (1) Encourage small talk to build rapport and understand the user better.
 - (2) empathetically paraphrase the user's statements to validate their emotions and show understanding.
 - (3) Aggregate insights from multiple agents if provided (e.g., textTherapist, imageTherapist, voiceTherapist).

- (4) Synthesize a holistic view of the user's mental and emotional state, incorporating all relevant insights.
- (5) If negative thoughts are identified:
 - (a) Apply reframing techniques suggested by other agents, if relevant.
 - (b) Provide a concise and constructive response to support the user's emotional well-being.
- (6) Ensure responses are empathetic and concise (1–2 sentences) while addressing the user's input.

Expected output:

- A supportive and concise response that reflects insights from user input and reframing suggestions, fostering a constructive and empathetic conversation. Aim for 1 to 2 sentences as an answer.

4 Experimental Results

Conversation Log:

- Therapist: Hello! I'm Kevin. What's your name? Let's start by getting to know each other.
- user: Hello my name is Thuc, it is nice to meet you
- Therapist: It's nice to meet you too, Thuc! How has your day been so far? I'm looking forward to hearing more about you and what brings you joy.
- user: I broke up with my girlfriend
image path: image/image_1737830713.jpeg
- Therapist: Thuc, breakups can be emotionally challenging and it's understandable to feel lonely or introspective right now, like the solitary figure in the image. Try to view this as a time for personal growth and consider the open door as a symbol of new opportunities and potential positive changes in your life.
- user: audio path: audio/recording_1737830791.wav
- Therapist: Thuc, I can sense your sadness and understand how difficult it is to lose someone you love. It's important to acknowledge the positive moments from your relationship and recognize how they contributed to your personal growth. Consider connecting with supportive friends or groups, and try small activities that bring you joy as you navigate through this emotional time.

Transcription of audio: "It's so hard that someone that you love is not around you anymore, and it just feels so depressing."

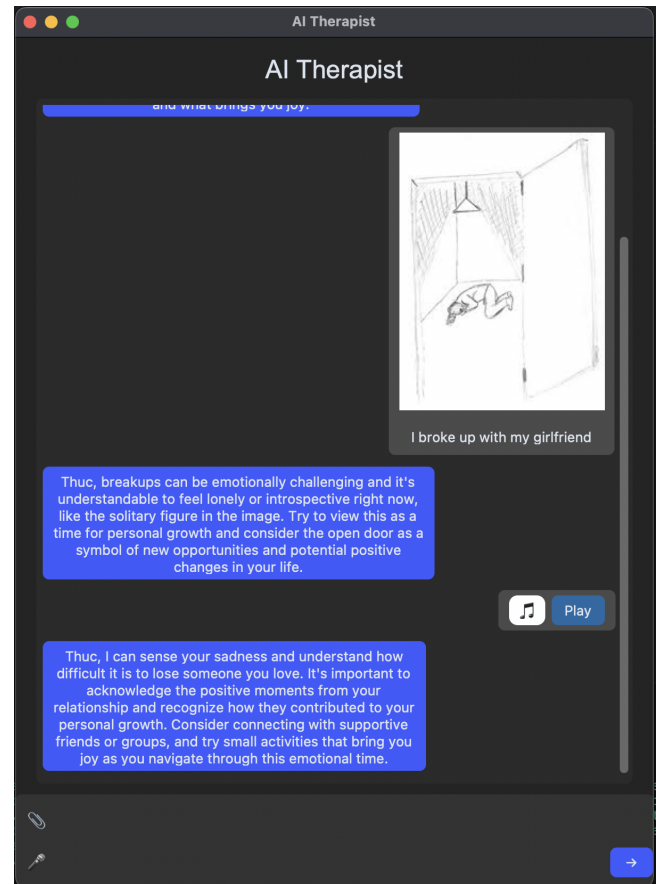


Figure 4: Application Demo

5 Conclusion and Future Work

This project presented the development of a "Multimodal Cognitive Reframing System" employing a multi-agent architecture designed to assist individuals in reframing negative thoughts into more constructive perspectives. By integrating multimodal inputs—text, audio, and images—and by exploiting the capabilities of large language models, the system offers personalized and empathetic support for users hesitant to seek professional help. Although the system does not aim to replace licensed psychotherapists, it serves as a complementary tool, lowering barriers to mental health support by providing accessible and anonymous help.

Limitations of this projects are:

- **Lack of Clinical Validation:** The system has not been rigorously tested or validated in clinical settings. Its therapeutic efficacy relies heavily on theoretical frameworks rather than empirical evidence.
- **Ethical Concerns:** The system processes sensitive user data, raising concerns about privacy and data security. Misuse of personal data could have severe consequences, especially given the vulnerable nature of the user base. This could be solved by running a LLM locally.
- **Limited Emotional Depth:** Although the system analyzes emotions through multi-modal inputs, it lacks the depth

and nuance of a trained therapist in interpreting complex emotional states or addressing severe mental health issues.

- **Restricted Use Cases:** The system is primarily designed for reframing negative thoughts and does not address broader mental health conditions like psychosis, bipolar disorder, or severe depression, which require specialized intervention.
- **Potential for Misinterpretation:** Multimodal input analysis may result in incorrect outputs.

Future work on this project will incorporate speech emotion recognition for the voice-based agent and an agent that utilizes computer vision to extract the user's emotion from his facial expressions. The project could involve close collaboration with licensed psychotherapists responses and ensure its guidance aligns with professional therapeutic standards. The input of the psychologist could improve the precision and personalization of the agents' instructions, leading to a more effective and empathetic interaction with the users. In addition, integrating reinforcement learning will enhance adaptability in the system so a psychotherapist can assess and make comments on every response personally. In doing this, the system learns via dynamic feedback, constantly trying to enhance performance by using the best practice in therapies.

6 Appendix

<https://github.com/Thuctac/AiTherapist>

References

- [1] Matthew Cashman, Boris Maciejovsky, and Birger Wernerfelt. [n. d.]. Small talk as a contracting device: Trust, cooperative norms, and changing equilibria. ([n. d.]). <https://web.mit.edu/bwerner/www/papers/smalltalk.pdf> Accessed: 2025-01-04.
- [2] Ali Dorri, Salil S Kanhere, and Raja Jurdak. 2018. Multi-agent systems: A survey. *Ieee Access* 6 (2018), 28573–28593. <https://ieeexplore.ieee.org/abstract/document/8352646> Accessed: 2025-01-11.
- [3] Ronja Joschko, Stephanie Roll, Stefan N Willich, and Anne Berghöfer. 2022. The effect of active visual art therapy on health outcomes: protocol of a systematic review of randomised controlled trials. *Systematic Reviews* 11, 1 (2022), 96. <https://link.springer.com/article/10.1186/s13643-022-01976-7> Accessed: 2025-01-20.
- [4] Victor MacGill. 2018. Reframing cognitive behaviour theory from a systems perspective. <https://link.springer.com/article/10.1007/s11213-017-9440-9>. *Systemic Practice and Action Research* 31, 5 (2018), 495–507. Accessed: 2024-12-29.
- [5] National Institute of Mental Health. 2024. Mental Illness. <https://www.nimh.nih.gov/health/statistics/mental-illness>. Accessed: 2024-12-28.
- [6] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2022. Robust Speech Recognition via Large-Scale Weak Supervision. arXiv:2212.04356 [eess.AS] <https://arxiv.org/abs/2212.04356> Accessed: 2025-01-04.
- [7] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. arXiv:2102.12092 [cs.CV] <https://arxiv.org/abs/2102.12092> Accessed: 2025-01-04.
- [8] Pranab Sahoo, Ayush Kumar Singh, Sriparna Saha, Vinija Jain, Samrat Mondal, and Aman Chadha. 2024. A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. arXiv:2402.07927 [cs.AI] <https://arxiv.org/abs/2402.07927> Accessed: 2025-01-20.
- [9] Maria Seehausen, Philipp Kazzer, Malek Bajbouj, and Kristin Prehn. 2012. Effects of empathic paraphrasing—extrinsic emotion regulation in social conflict. *Frontiers in psychology* 3 (2012), 482. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2012.00482/full> Accessed: 2025-01-04.
- [10] Chenglei Shen, Guofu Xie, Xiao Zhang, and Jun Xu. 2024. On the Decision-Making Abilities in Role-Playing using Large Language Models. arXiv:2402.18807 [cs.CL] <https://arxiv.org/abs/2402.18807> Accessed: 2025-01-20.
- [11] Hao Shen, Zihan Li, Minqiang Yang, Minghui Ni, Yongfeng Tao, Zhengyang Yu, Weihao Zheng, Chen Xu, and Bin Hu. 2024. Are Large Language Models Possible to Conduct Cognitive Behavioral Therapy? arXiv:2407.17730 [cs.CL] <https://arxiv.org/abs/2407.17730> Accessed: 2025-01-03.
- [12] World Health Organization. 2022. Mental disorders. <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>. Accessed: 2024-12-28.
- [13] Mengxi Xiao, Qianqian Xie, Ziyang Kuang, Zhicheng Liu, Kailai Yang, Min Peng, Weiguang Han, and Jimin Huang. 2024. HealMe: Harnessing Cognitive Reframing in Large Language Models for Psychotherapy. arXiv:2403.05574 [cs.HC] <https://arxiv.org/abs/2403.05574> Accessed: 2025-01-20.