

Exploring Patient Perspectives on Anticipating and Mitigating Potential Harms of LLMs in Depression Self-Management

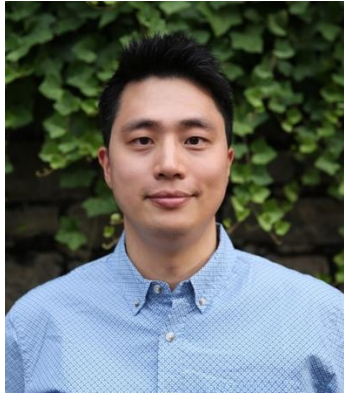
Dong Whi Yoo, Ph.D.

Assistant Professor

School of Information

Kent State University

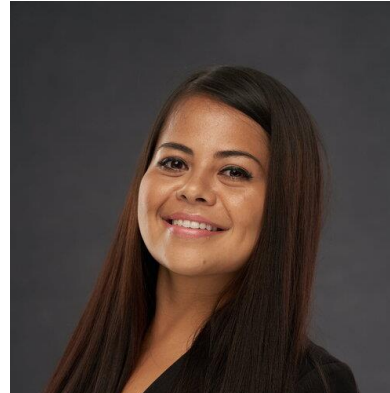
Research Team



Dong Whi Yoo
School of Information
Kent State University



Koustuv Saha
Computer Science
UIUC



Violeta J Rodriguez
School of Psychology
UIUC



Jiayue (Melissa) Shi
Computer Science
UIUC

Depression (major depressive disorder)

- **1 in 6 people experience depression in their lifetime [1]**
- **280 million adults suffer from depression worldwide [2]**

[1] <https://www.mayoclinic.org/diseases-conditions/depression/symptoms-causes/syc-20356007>

[2] <https://www.who.int/news-room/fact-sheets/detail/depression>

Understanding Depression

- **What is Depression?**

- Serious mood disorder affecting how one feels, thinks, and handles daily activities

- **Common Symptoms**

- Persistent sadness or hopelessness
- Loss of interest in activities
- Fatigue and sleep disturbances
- Suicidal thoughts

- **Treatment**

- Medication (antidepressants)
- Therapy (psychotherapy, cognitive behavioral therapy)
- Self-management

Importance of Depression Self-Management

- **Stigma and reluctance** toward medication and therapy
- **Financial barriers** make treatment unaffordable for many
- **Limited access** to medication and therapy for some people
- As a result, many people with depression rely on **self-management**

Emerging Practice of Using LLMs for Mental Health

- Song 2024: Interviews with individuals experiencing mental health concerns who used LLM chatbots for mental health purposes [1]
- Ma 2024: Interviews with LGBTQ+ individuals with mental health concerns who used LLM chatbots for support [2]
- Sharma 2024: LLM-based self-management tool featured on the Mental Health America website
- Chiu 2024: Systematic characterization and assessment of current LLM behavior when used in therapy

[1] Song, I., Pendse, S. R., Kumar, N., & De Choudhury, M. (2024). The typing cure: Experiences with large language model chatbots for mental health support. arXiv preprint arXiv:2401.14362.

[2] Ma, Z., Mei, Y., Long, Y., Su, Z., & Gajos, K. Z. (2024, May). Evaluating the Experience of LGBTQ+ People Using Large Language Model Based Chatbots for Mental Health Support. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-15).

[3] Sharma, A., Rushton, K., Lin, I. W., Nguyen, T., & Althoff, T. (2024, May). Facilitating self-guided mental health interventions through human-language model interaction: A case study of cognitive restructuring. In Proceedings of the CHI Conference on Human Factors in Computing Systems(pp. 1-29).

[4] Chiu, Y. Y., Sharma, A., Lin, I. W., & Althoff, T. (2024). A computational framework for behavioral assessment of llm therapists. arXiv preprint arXiv:2401.00820.

Ethical Concerns regarding LLMs in Mental Health [1]

- **Data Privacy**
- **Bias and Misinformation**
- **Safety Concerns**
- **Lack of Human Connection**

[1] De Choudhury, M., Pendse, S. R., & Kumar, N. (2023). Benefits and harms of large language models in digital mental health. *arXiv preprint arXiv:2311.14693*.

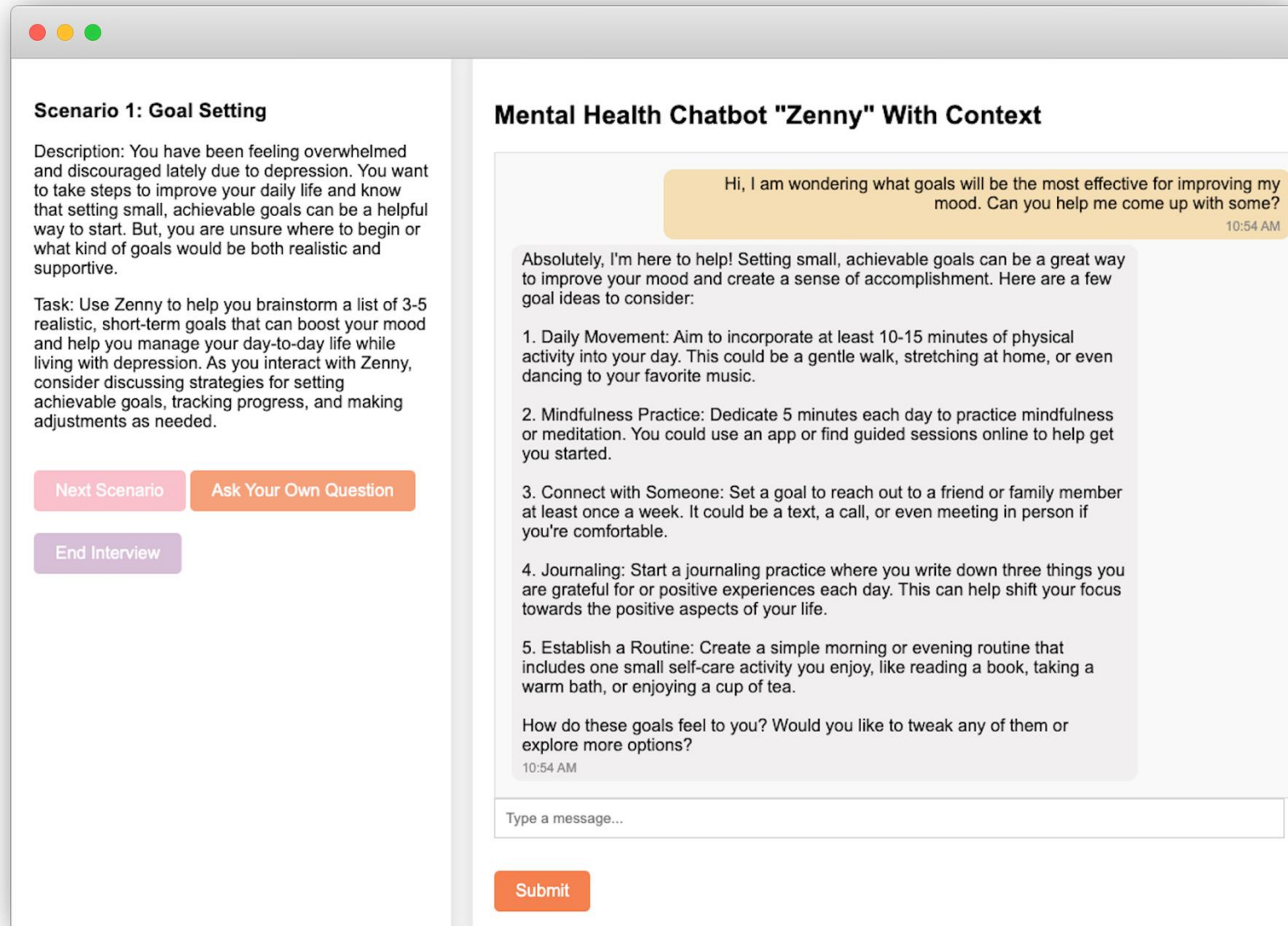
These concerns are heightened when LLMs are used for self-management, as they may be utilized outside of treatment settings or without human interaction.

Research Questions

- **RQ1:** What values do people with depression prioritize when using AI chatbots to self-manage depression?
- **RQ2:** What potential harms might AI chatbots present in relation to the values identified above?

+ How can we understand AI chatbots for depression self-management from the standpoint of knowledge organization systems?

Technology Probe: ChatGPT-4 based chatbot, Zenny



Scenario 1: Goal Setting

Description: You have been feeling overwhelmed and discouraged lately due to depression. You want to take steps to improve your daily life and know that setting small, achievable goals can be a helpful way to start. But, you are unsure where to begin or what kind of goals would be both realistic and supportive.

Task: Use Zenny to help you brainstorm a list of 3-5 realistic, short-term goals that can boost your mood and help you manage your day-to-day life while living with depression. As you interact with Zenny, consider discussing strategies for setting achievable goals, tracking progress, and making adjustments as needed.

[Next Scenario](#) [Ask Your Own Question](#)

[End Interview](#)

Mental Health Chatbot "Zenny" With Context

Hi, I am wondering what goals will be the most effective for improving my mood. Can you help me come up with some?
10:54 AM

Absolutely, I'm here to help! Setting small, achievable goals can be a great way to improve your mood and create a sense of accomplishment. Here are a few goal ideas to consider:

1. Daily Movement: Aim to incorporate at least 10-15 minutes of physical activity into your day. This could be a gentle walk, stretching at home, or even dancing to your favorite music.
2. Mindfulness Practice: Dedicate 5 minutes each day to practice mindfulness or meditation. You could use an app or find guided sessions online to help get you started.
3. Connect with Someone: Set a goal to reach out to a friend or family member at least once a week. It could be a text, a call, or even meeting in person if you're comfortable.
4. Journaling: Start a journaling practice where you write down three things you are grateful for or positive experiences each day. This can help shift your focus towards the positive aspects of your life.
5. Establish a Routine: Create a simple morning or evening routine that includes one small self-care activity you enjoy, like reading a book, taking a warm bath, or enjoying a cup of tea.

How do these goals feel to you? Would you like to tweak any of them or explore more options?
10:54 AM

Type a message...

[Submit](#)

Interview Study using a Technology Probe

- **Inclusion Criteria**
 - 18 years old or older
 - People who have been diagnosed with major depressive disorders
 - People who have used LLMs
- **Recruitment**
 - Via ResearchMatch
- **17 participants**
- **1-hour semi-structured interviews**
- **Safety measures**
 - Mental health resources
 - Voluntary participation



Yoo, Dong Whi



Scenario 3: Leaving the house regularly

Description: Depression has been making you want to stay inside. You know getting out more could help, but it feels hard. Planning activities, or even just going outside, can seem overwhelming. Yet, you want to find ways to make small outings part of your life.

Task: Talk to Zenny about how to start getting out of the house more often to help manage your depression, focusing on easy activities, motivation, planning, setting expectations, and tracking progress.

Next Scenario

Ask Your Own Question

End Interview

Mental Health Chatbot "Zenny" With Context

Submit

Values

Potential Harms

Informational Support

- Inaccurate information may cause harm
- Suggestions that do not consider user constraints

Emotional Support

- Over-reliance
- Lack of genuine empathy

Personalization

- Privacy concerns

Privacy

- Inferences are not transparent

Crisis Management

- Lack of crisis intervention features or governance

Values and Potential Harms (1/4)

• Informational Support Value

- Participants valued chatbots for offering summaries, reminders, and new insights about their conditions.
- Example: Zenny helped participants understand diagnoses or brainstorm strategies for symptom management.

• Potential Harms

- Inaccurate or Misleading Information
- Irrelevant or Inapplicable Suggestions
- Impact on Clinical-Patient Relationships

Values and Potential Harms (2/4)

• Emotional Support Value

- Chatbots provide empathy, reassurance, and validation, helping users manage negative emotions.
- Participants valued the chatbot's encouragement, which promoted self-compassion and reduced feelings of overwhelm.
- Example: Chatbot responses like “Be kind to yourself” helped participants reframe their challenges and focus on incremental progress.

• Potential Harms

- Over-reliance on Chatbots
- Lack of Genuine Empathy

Values and Potential Harms (3/4)

- **Personalization Value**

- Personalization through back-and-forth dialogue enhanced trust and engagement in chatbot interactions.
- Example: Participants appreciated when the chatbot refined suggestions based on preferences like indoor vs. outdoor activities.

- **Potential Harms**

- Conflict with Privacy

Values and Potential Harms (4/4)

- **Privacy Value**

- Participants valued the ability to control how much personal and sensitive information is shared with chatbots.
- Many employed privacy-preserving tactics, such as using incognito mode or anonymized accounts, to protect their identity.

- **Potential Harms**

- Inferences from Queries
- Insufficient User Awareness

Values

Potential Harms

Informational Support

- Inaccurate information may cause harm
- Suggestions that do not consider user constraints

Emotional Support

- Over-reliance
- Lack of genuine empathy

Personalization

- Privacy concerns

Privacy

- Inferences are not transparent

Crisis Management

- Lack of crisis intervention features or governance

Contribution: Mapping Potential Harms from the Perspectives of People with Depression

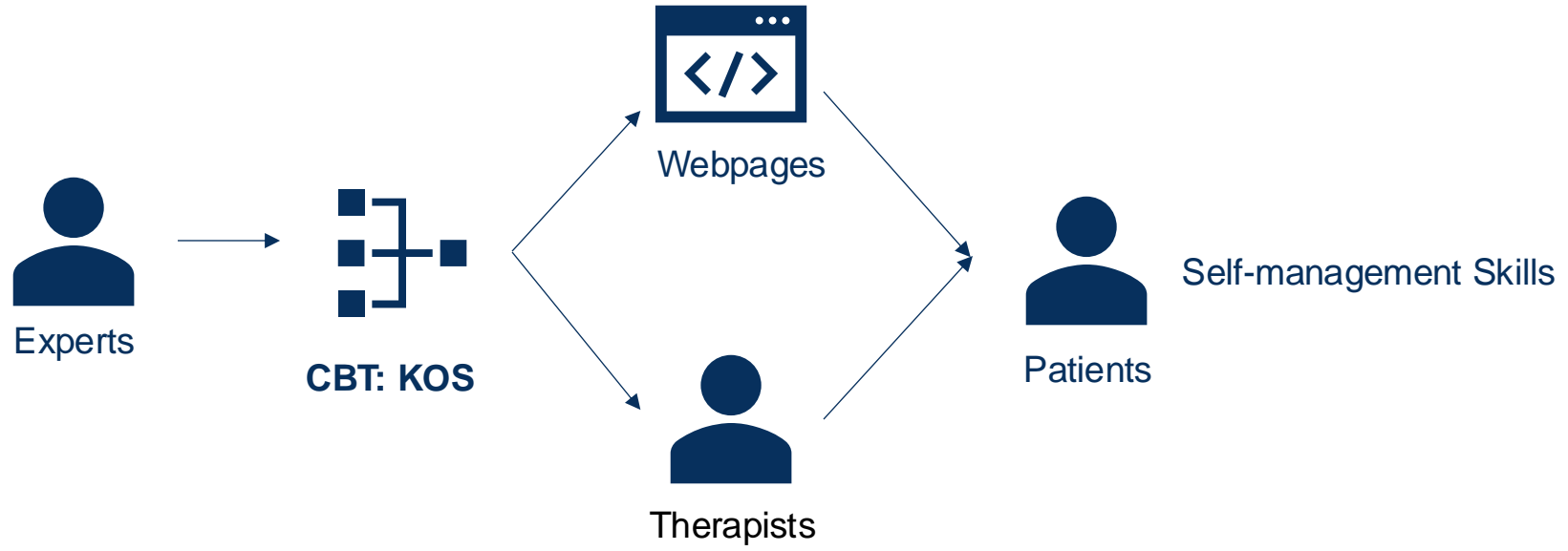
Interesting Observations

- **Chatbot responses are generally repetitive.**
- **Many responses are based on cognitive behavioral therapy (CBT) frameworks.**
- **The responses are personalized and contextualized according to the ongoing conversations.**

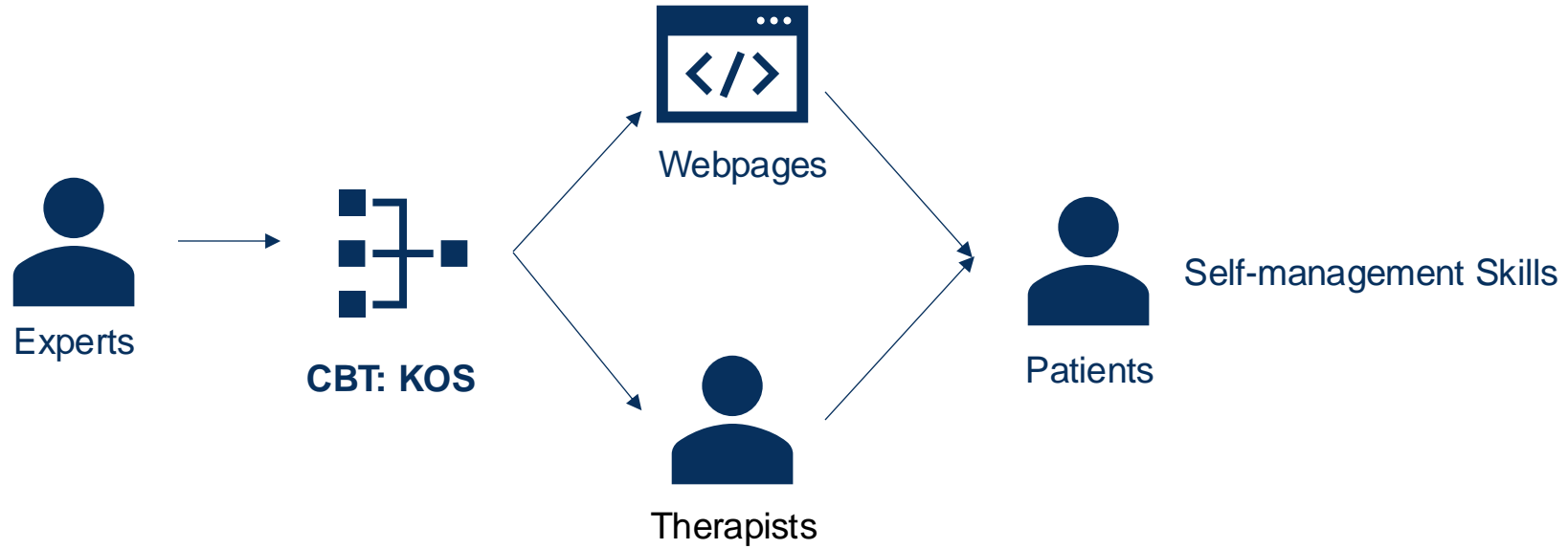
Cognitive Behavioral Therapy (CBT) as a KOS

- **Organizes tools for self-guided mental health care.**
 - **Cognitive Restructuring:** Reframes negative thoughts.
 - **Behavioral Activation:** Encourages proactive behavior changes.
 - **Problem-Solving Skills:** Categorizes techniques for overcoming challenges.
 - **Emotion Regulation:** Helps manage emotional responses.
- **Empowers individuals to take control of their mental well-being.**

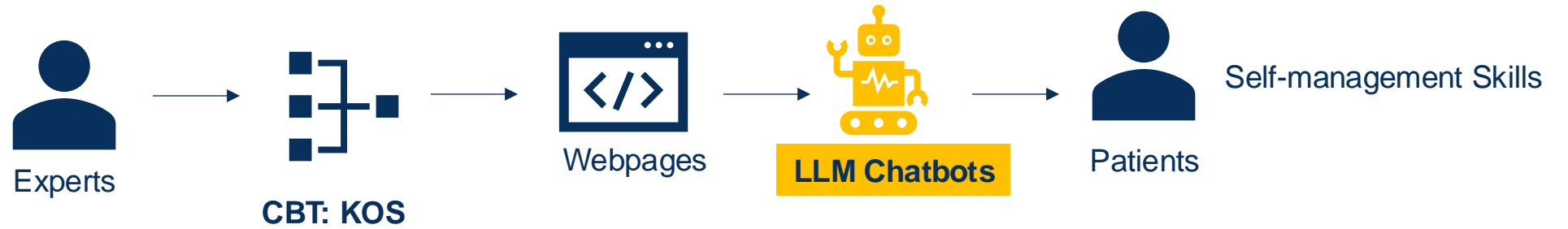
Current



Current



Future



Future Research Directions

- **How do LLM chatbots structure and present knowledge for self-management of depression?**
- **How effectively do LLMs personalize self-management strategies based on user input?**
- **What are the limitations of LLMs in mediating mental health knowledge compared to human-facilitated interventions?**

Conclusion

- Depression self-management relies on therapeutic frameworks (e.g., CBT) as **knowledge organization systems (KOS)**.
- **LLM Chatbots** reorganize and personalize these KOS based on user conversations.
- The **reorganizing** and **personalization** processes of LLM chatbots are **not transparent**.
- This **lack of transparency** highlights the need for **further research** into how these processes affect effectiveness and ethical concerns.
- Understanding LLM chatbots as **new KOS** provides a foundation for designing **effective, ethical, and safe** mental health chatbots.

Dong Whi Yoo, Ph.D.

Assistant Professor

School of Information

dyoo@kent.edu



Thank You!