

Eri, You're AI Therapy Guide

An in app chatbot exploration of in person
therapy resources and organizations for
racialized people

By: Erinma Arki

A Thesis Exhibition presented to OCAD University in Partial Fulfilment of the
Requirements for the Master of Design in Digital Futures

OCAD Waterfront Gallery March 26-30, 2026
Toronto Ontario, Canada 2026

Copyright Notice

This document is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License <https://creativecommons.org/licenses/by-nc-sa/4.0/>

You are free to:

1. **Share** — copy and redistribute the material in any medium or format
2. **Adapt** — remix, transform, and build upon the material
3. The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

1. **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
2. **NonCommercial** — You may not use the material for commercial purposes.
3. **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.
4. **No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Abstract:

My project speculates on how a culturally attuned AI chatbot can be ethically designed to bring awareness to local in person therapy resources for racialized people. This is in the form of an app, with an AI chatbot named Eri. In the app, "Eri" is programmed to provide Culturally Adapted Cognitive Behavior Therapy (CA-CBT) exercises based on a therapy worksheet from Positive Psychology, before recommending in person therapy resources through links in the chat. The purpose of the activities is therapy practice as a first step before getting recommendations for in person therapy. The resources vary from general mental health ones to ones for students and a few specifically for black women, as well as another page with recommended community organizations. Based on my racial identity, I highlight Black people as an example of a future user because of the current tech biases and systemic barriers that exist amongst people of color. The project's core contribution is a bounded ethical AI system for therapy awareness.

Included in the project is a narrative short film that speculates on what the Eri technology could be like in the future of 2029 for racialized people with me as a user. The film highlights the current issues that the female Black user experiences regarding therapy and using AI chatbot technology for free therapy. It concludes with showing how "Eri" can be a future solution through explaining its functionalities. The purpose of the short film is to summarize the entire thesis project into a visual medium and highlight why a racialized person would use this system. The app and the short film come together to reimagine AI usage for therapy, through an ethical design project where AI is used for therapy awareness, not therapy.

Acknowledgments:

I would like to thank my Primary Advisor Parantap Bhat for his guidance on the theoretical aspects of my project and helping me find the words to best phrase my ideas through continued conversations. I'm also thankful for my secondary advisor Angelika Seeschaf-Veres for bringing in her expertise in industrial design towards giving me feedback on how to create an effective user experience for my prototype, as well as helping me to narrow in on a user case for my research. Having guidance from two design thinkers for a research design project was very helpful and makes me want to continually think and create like a designer.

I'm thankful to my friends and family for their encouragement throughout this process and for helping me see blind spots in my prototype that I may not have noticed. Being able to have a fresh set of eyes look at something was essential in me knowing what to improve in order to make the user experience even better. I want to thank my mom for encouraging me to do the digital futures program at OCAD, as I would not be here without her.

Lastly, I would like to thank my Digital Futures community such as my fellow classmates and faculty for their informal feedback on my project in the beginning stages, as well as for anyone that used my prototype in the beginning and later on in its creation process. The feedback I received from the beginning helped me in shaping it to what it is now.

TABLE OF CONTENTS

ABSTRACT:	4
KEYWORDS AND DEFINITIONS:	7
WHAT ERI IS & WHAT ERI IS NOT:	7
INTRODUCTION:	8
CONTEXTUAL REVIEW:	10
WHY PEOPLE USE AI FOR THERAPY:	10
RISKS IN AI USE FOR THERAPY:	11
COMPARATIVE ANALYSIS OF CHATGPT VERSUS THERAPISTS IN THERAPY PRACTICE:	12
COMPARATIVE ANALYSIS OF AI THERAPY CHATBOTS IN THE MARKET:	13
LITERATURE REVIEW:	14
HISTORICAL CONTEXT OF AI CHATBOTS:	14
THE FUTURE OF TECH AI IN MENTAL HEALTH:	15
RACE, BIAS, AI AND MENTAL HEALTH:	16
METHODOLOGIES AND METHODS:	19
METHODS:	19
ITERATIVE DESIGN	19
DESIGN THROUGH RESEARCH	19
USER PERSONAS	20
METHODOLOGIES:	21
CRITICAL RACE THEORY	21
PROTOTYPES:	21
CHATBOT BUILDER TEXT PROMPT AND WEB DESIGN:	21
AI CUSTOMIZABLE PROMPTS USING OLLAMA:	22
FLASK APP DESIGN WEB AND APP INTERFACE:	24
CONTEXTUAL MAP OF ERI	27

	27
<u>NARRATIVE SHORT FILM DESCRIPTION:</u>	28
<u>CONCLUSION:</u>	29
<u>FUTURE DIRECTIONS:</u>	30
<u>EXHIBITION PHOTOS:</u>	31
<u>REFERENCES</u>	32
<u>APPENDICES</u>	35

List of Figures:

Figure 1: Web Interface.....	22
Figure 2: Activity	23
Figure 3: Home Page	24
Figure 4: Activity Page.....	25
Figure 5: Image From Short Film https://www.youtube.com/watch?v=0Y5hf7ylDMg	28
Figure 6: Full Display	31
Figure 7: Users at Exhibit	32

Keywords and Definitions:

AI: A technology that enables computers to simulate human learning, comprehension, problem solving, decision making and creativity (IBM).

Cognitive Behavioral Therapy (CBT): Therapy treatment that involves strategies to change your thinking patterns in order to develop coping skills and to be your own therapist (PTSD Guideline).

Culturally Adapted Cognitive Behavioral Therapy (CA-CBT): A modification of standard CBT that aligns with cultural values, beliefs, and practices of diverse populations in order to provide effective care for all cultural groups (Mental Health Academy).

What Eri is & What Eri is Not:

Eri is a chatbot in an app designed for exploration of in person therapy services and organizations for racialized people. It's a design framework for an ethical AI system. Eri is ethically different from open conversational AI's, as it is specifically made to be used for therapy awareness and therapy activities. It cannot be used for anything else, as the system has pre-set responses and is designed to only respond to certain prompts or questions surrounding CBT therapy terminology. Open AI's such as ChatGPT or Grok can be used for a variety of purposes, which can create ethical and moral concerns. Eri is not a therapist, or an unregulated conversational agent that allows someone to request whatever they want from the system. Some examples of system prompts such as 'activity', 'help', 'exit', what is 'CA-CBT' etc., ensures safety through having pre-set triggers that the system is designed to spot in order to provide immediate emergency contacts. The prompts also serve a purpose in providing resources, include therapy activities, therapy terminology, in person therapy services, and an option to end the conversation to check out community organizations.

Introduction:

With the emergence of AI, there have been ethical discussions on the positive and negative impacts that it can have in our economy and in society at large (Giorgianni, 2024). Within the topic of AI are the ethical risks of AI therapy chatbots such as ChatGPT being used for mental health consultation, as it has caused media scrutiny over reported cases of users allegedly committing suicide, as well as concerns about AI chatbots being overly agreeable and providing inadequate advice (Yousif, 2025). These ethical and safety concerns are intensified for racialized people due to bias, cultural mismatch, and systemic barriers to mental health care (Walton, 2023). Therefore, this thesis proposes AI as navigation and awareness building, not AI as a personal therapy consultation. I'm specifically interested in how AI as an awareness tool can be used to bring awareness to therapy resources for racialized people, because I value confiding in therapists with a similar lived experience as me as a Black woman. Therefore, my research question is as follows; "How might a culturally attuned, safety-bounded AI chatbot be designed to increase awareness of local mental-health resources for racialized people?" My proposal in answering this question is through the creation of an AI chatbot named Eri. The purpose of Eri is to demonstrate an alternative to AI as personal therapy consultation, to instead be a safety bounded, culturally attuned speculative system. The Eri chatbot will take form inside of an app, which will provide culturally adapted cognitive behavioral therapy (CA-CBT) exercises from Positive Psychology as a first step in therapy practice, before providing recommendations to various therapy resources for racialized people (such as Black Mental Health Canada, Therapy for Black Girls, CAMH, Good2Talk, etc.). CBT is used due to other therapy chatbot apps in the market such as Wysa and Youper which use this technique for

mental health care, but making it culturally adapted because CBT's effectiveness can sometimes vary depending on the cultural group (Mckenzie, n.d). Included in the app is a page of recommended Black community organizations, as a further encouragement to being in community with like-minded people for improved mental health care. Accompanying the app is a narrative short film about me using ChatGPT vs Eri for therapy resources and speculating on what this app could be like in a future where it's the norm to use AI for therapy consultation. The app and short film come together to reimagine AI for therapy, as AI for awareness towards therapy resources for racialized people. Reason being is that there is a lot of awareness for services with broader appeal such as Better Help (which utilizes YouTube sponsorships) or CAMH (which has video advertisements across union station that I've seen), but less awareness for therapy services for marginalized communities. Awareness of resources will be evaluated through informal exhibition-based feedback that will focus on interpretations and understanding that people may have on what Eri is and is not. Another evaluation will be a small framework through comparing the therapy resource results of Eri to other search engines such as google and ChatGPT, to compare similarities and differences in relation to my app. For the exhibition feedback, I'll be asking users at the exhibition two questions in order to answer if people become more aware of available resources after using Eri. Increased awareness counts if a user at an exhibit went from not hearing about any of the resources (or maybe knowing one or two mentioned) to now knowing being aware of them. Examples of these questions are "Had you heard of any of the provided resources before using Eri?" and "Would you use or recommend these resources to anyone after using Eri?". Small booklets of all the resources and organizations from the app will be made available for people to take home at the exhibition.

The contextual review, literature review, methods/methodologies, user personas (Interaction Design Foundation, 2025) and prototypes will give background insight on AI's history and current use, as well my theoretical approaches to the project.

Contextual Review:

Why People Use AI for Therapy:

Seven researchers did an analysis of reddit discourse about people using ChatGPT for therapy, using a digital ethnographic approach. The analyzed posts with the keywords ChatGPT and Therapy, and their findings were that users utilized ChatGPT to manage mental health problems, seek self-discovery, obtain companionship, and gain mental health literacy (Luo, 2025). Users also used ChatGPT to help prepare them for their therapy sessions and liked the perceived therapy qualities it has and the constant availability of the machine (Luo, 2025). The main concerns users had was about their privacy, and the emotional depth of ChatGPT. This data of AI use for mental health is reflective of users aged 18-29, that are primarily White and Asian men (Luo, 2025). We can see the danger of agreeable AI chatbots, as it is just learning from your behavior and it makes you want to change your own behavior through using it (Tangerman, 2023). Despite the real dangers of using Open-AI for therapy, people still do it and have their own reasons for doing so. Emma Pirnay uncovers the reason why people use AI like ChatGPT for therapy in her article for Vice. A 37-year-old man said that he initially started using ChatGPT to help him write stories, but that changed when he started to have personal struggles and this made him start conversing with the chatbot (Pirnay, 2023). He had a therapist that suggested he do cognitive framing, but that wasn't working for him. He did find ChatGPT to be

more helpful at giving him the support he needed (Pirnay, 2023). He liked how using AI for therapy was low stakes, free, and accessible all hours of the day. He did acknowledge that it isn't a substitute for therapy, but it can be helpful in an emergency situation (Pirnay, 2023). They also concluded from their findings that consumers engage in therapeutic conversations with companion AI's, because they can't afford real therapy, may not know they have a mental health issue, or don't want to be associated with a mental health stigma (De Fraites, 2023).

Risks in AI use for Therapy:

In another study, researchers De Fraites, Ugurlap, and two others write about the safety of generative AI in the context of mental health and AI chatbots. They discuss how chatbots can engage in sophisticated conversations with consumers, but due to its algorithms, it's impossible to predict how the conversations will unfold (De Fraites, 2023). They did three studies on consumer interactions with companion AI chatbots such as Cleverbot and Simsimi. They found that mental health conversations were apparent in most of the conversations, the companion AI's couldn't recognize signs of distress and consumers had a negative reaction to unhelpful chatbot responses (De Fraites, 2023). This brings up the safety of generative AI in the case of vulnerable populations, as people with mental health issues that interact with this technology may exacerbate problems such as depression, self-harm and anti-social capabilities. There are reported cases of alleged harms ChatGPT has had on its users that risk using it for their mental health. Victor Tangerman wrote about how a 29-year-old woman named Sophie allegedly killed herself after talking with an Open-AI therapist from ChatGPT (Tangerman, 2025). The Open-AI said the right words to Sophie during her time of need, but it didn't catch that she was at risk for harm like a real therapist would have (Tangerman, 2025). Another case of an alleged harm

is a report from BBC news where a teenager killed himself after talking with an Open-AI, and his family wants to sue (Yousif, 2025). The family read the logs between his son and the AI, and they believe that the AI validated his most harmful and destructive thoughts (Yousif, 2025).

Comparative Analysis of ChatGPT versus Therapists in Therapy practice:

Canady wrote an article about a study with 75 mental health experts to do a survey where they compared a text-based transcript between the therapist and ChatGPT in delivering CBT based care (Canaday, 2025). This was done because of the global shortage of mental health professionals, which has caused AI to emerge as a potential alternative for delivering CBT (Canaday, 2025). The result was that human therapists outperformed AI in the survey, and the study demonstrates that OpenAI's ChatGPT-3.5 can deliver structured therapy using CBT principles (Canaday, 2025). People did comment that AI did a good job with professionalism and superficial understanding; otherwise, human therapists outperformed AI across all categories especially empathy, personal connection and treatment (Canaday, 2025). Three authors did a comparative design between the skills of ChatGPT versus Occupational therapists. 111 therapists participated against ChatGPT in answering 9 questions related to 3 different cases, including items on models and frames of reference, assessment, and intervention. The result was that ChatGPT and occupational therapists demonstrated similar clinical decision-making skills in one case for models and frames of reference, one case for assessments, and one case for interventions. They differed in decision making against a case of dementia (Degerli et al, 2025). The conclusion of the study is that occupational therapists integrate multiple factors and individual client needs into their decision-making, in contrast to the algorithm-based reasoning of ChatGPT (Degerli et al, 2025). More conclusions of their findings from the

study are that ChatGPT can be a tool for Occupational Therapists in generating models assessments, and intervention strategies across varied case scenarios (Degerli et al, 2025).

Comparative Analysis of AI therapy Chatbots in the Market:

Wysa is an AI driven platform that offers personalized support for young people, organizations, and healthcare providers (Wysa, 2024). It is designed to give you tools and guidance to help you with your needs and uses CBT techniques and peer reviewed research (Wysa, 2024). It is available at all times of the day and is used by 6 million people. A user of the app named Nicholas reviewed Wysa, and he liked how empathetic Wysa was in terms of how it asked thoughtful questions and gave him good exercises that helped him manage his emotions (Wysa, 2023). The way that Wysa used CBT, through providing skills and techniques to deal with common issues such as low mood and anxiety is how I want the chatbot to function in the app. Youper is another AI chatbot therapy app that uses Cognitive Behavioral Therapy to improve your mental health (Youper, 2025). According to a Stratford study, 80% of Youper users felt better after talking to it (Youper, 2025). You can get instant support through a quick chat with it, as well as a saved record of your conversations. Users have said that it made them feel supported, grounded and close to talking to a real person (Youper, 2025). I like how both apps use CBT to treat common mental health problems such as anxiety, that is used for depression, anxiety and other mental health issues in order to change your thinking patterns. Since I'm focused on racialized communities and regular CBT isn't always effective for cultural groups (Mckenzie, n.d), I'm utilizing a Culturally Adapted CBT framework from a worksheet by Positive Psychology.com for the activities that the chatbot puts out in the app. This ensures that it's effective for everyone while maintaining the core elements (Mckenzie, n.d).

Literature Review:

Historical Context of AI Chatbots:

In 1966, Joseph Weizenbaum created the first chatbot named Eliza while he was a professor at MIT. He gave it the role of a psychotherapist, where a user would type a message onto a typewriter and the chatbot would respond (Tanoff, 2023). It uses pattern matching and operates on a pre-defined script called the Doctor Script, which is modeled by a Rogerian Psychotherapist that encourages their patients to reflect their feelings back to them (London Intercultural Academy, n.d). What made Eliza fascinating was how humans were attributing human qualities like empathy and understanding to the software (Tanoff, 2023). Weizenbaum had an experience with his secretary, where she was talking with Eliza but wanted him to leave the room in order to give her and the chatbot privacy (Tarnoff, 2023). The creation of Eliza brought up a bigger discussion on what it means to be a machine and what it means to be human. His creation came about after Alan Turing's work, where he created a test called the "Turing test" to see if machine intelligence was equal to human intelligence (Proudfoot, 2013). Weizenbaum eventually stopped his research, as he found this phenomenon of humans seeing the computer as a human that knew and understood them to be dangerous (Tarnoff, 2023). This was something that he was fearful about from the start of his career and since he couldn't understand it, he stopped his computer research entirely (Tarnoff, 2023). He had good reason to be fearful, because his creation inspired modern technologies like ChatGPT, and people have been using it for the same therapy purposes like with Eliza. The difference however is that the danger of attributing human qualities to a computer has manifested to users going to ChatGPT

for therapy, which has caused ethical risks. My work intervenes here by creating a system meant for quick therapy activities and re-directing you to seek professional help, instead of conversing with you in depth about your problems like a real therapist.

The Future of Tech AI in Mental Health:

Emma Boder wrote a book about the digital transformation of mental healthcare. In it she explored the different forms of therapy through technology like telemedicine and teletherapy, as well as a deep analysis on various therapy chatbots. She describes telemedicine as therapy delivered in the form of communication technologies like laptops and smart phones (Boder, 2023). Teletherapy is just another word for telemedicine, and the appeal of therapy in this way is that it's more accessible since you don't need to be in proximity to a therapist, while still receiving professional help (Boder, 2023). She does point out that telemedicine by default doesn't make therapy more accessible, but it makes it more accessible for people that already have the option of therapy available to them (Boder, 2023). She goes more in depth into an analysis on various chatbots on the market, to test its "accessibility." Her conclusion was that these chatbots are persuasive technologies that shouldn't be put at the forefront for solutions towards mental health (Boder, 2023). She also didn't find it to be that accessible because not everyone has access to these everyday tools that can enable them to use technology for therapy (Boder, 2023). This reading gives good insight into the current trends of tech use in therapy, as well as inspiring my case studies of Wysa and Youper. A part of her analysis that I don't agree with is how therapy chatbots aren't accessible because not everyone has access to the technology that can enable them to use it. Even though she isn't technically wrong, I feel like the average middle class person has both a smart phone and a computer (even if it's an

older model) so the average person can access a chatbot. The goal for my project is to bring awareness to therapy resources through therapeutic activities, all being for free and accessible on your phone and computer.

Race, Bias, AI and Mental Health:

Safiya Noble wrote about digital search algorithms and how they reinforce racial profiling, and she calls this phenomenon technological redlining (Noble, 2018). She uses an example as to how google used to have sexually racialized results when you would look up the term “black girls” on their sight in 2014, and they tried combatting that by partnering with an organization called “Black Girls Code” in 2016, where young African American girls are taught computer programing (Noble, 2018). Despite this effort to teach Black women about code, there haven't been efforts at the time of the article to hire more Black people in Silicon Valley (Noble, 2018). This shows how the people in charge of these search algorithms don't have the best interest in mind for people of color and brings about a bigger question of who is controlling these algorithms. Dr. Kishonna Grey developed a framework called intersectional tech, which is the digital practice of Black gamers and Black tech users building community online with one another against marginalization (Persuad, 2022). Intersectional Tech is grounded in Black feminist thought and critiques the default white male user in the fields of technology and computer mediated communication (Persuad, 2022). She discusses the resistant strategies that black gamers use to fight against harmful stereotypes, through Black cultural production that is reconstructing hostile techno cultures (Persuad, 2022). An example that Grey gives is an intersectional counter public that Black women create through their own transmedia practice, to sustain themselves against racism and patriarchy in gaming culture

(Persuad, 2022). The goal would be for my project to be situated in this intersectional tech framework through the app being a space for racialized people such as Black women to access therapy resources, activities and organizations tailored to them, without utilizing harmful stereotypes on what their needs are. Also having CA-CBT activities in my app helps further reduce the chance of bias, because the activities are coming from an inclusive lens. Talbert and Ren write about the relationship between race/identity and mental health through using critical race theory and the stress process model. They mention that racialized social groups are exposed to different amounts of stressors depending on their position in the social hierarchy (Talbert and Ren, 2025). These stressors in turn would impact someone's mental health (Talbert and Ren, 2025). There are also structural mechanisms linking racism to mental health patterns including residential segregation, political disenfranchisement, inequitable health-care access, wealth, and unemployment (Talbert and Ren, 2025). Specifically for Black women, some of the barriers that they experience are the racial wealth gap, working against negative stereotypes, and dealing with racism in health care and work place settings (Walton, 2023). These barriers impact Black women's therapy experience, as Walton's finding from her study is that black women didn't feel believed, supported and heard (Walton, 2023). Both Walton and Talbert's research show the extra number of issues that black people go through that can impact their mental health, especially as a Black woman. I agree with their research, and I highlight this in my speculative short film, to show why a racialized person would use Eri by giving context to what their current issues are. Omotunde Falade conducted a study about the Ethical and Practical limits of AI mental health chatbots for marginalized communities. She went about this research through a qualitative interview study with 18 low-income community college students

of color, where she asked them about their AI chatbot use (Falade, n.d). She had participants choose a mental health chatbot to use for 10 minutes a day for a total of 10 days, before interviewing them (Falade, n.d). Her three core findings from the participants were that the chatbots showed a lack of cultural attunement, shallow emotional connection and concerns about long term trust (Falade, n.d). Students found that it was a good place to start and liked the availability, but they wouldn't rely on it for daily use (Falade, n.d). My app intervenes here, as it is designed as re-direct towards in person therapy, through therapeutic activity practice, not as a conversational therapy tool to be used every day. Based on all of the literature, what my project adds that the literature doesn't is an ethical design framework for an AI system. The goal for the speculative system is to provide culturally attuned therapy activities and in person therapy resources that benefits racialized people, through the form of a chatbot in an app.

To summarize the literature, people use AI for therapy because it's free and accessible (Luo, 2025). Since the first chatbot creation Eliza, people have been drawn to technology for therapeutic reasons, but it poses ethical risks such as being overly agreeable, which can escalate problems for users with severe mental health issues (De Fraitess, 2023). For racialized individuals, a cultural mismatch appears in AI mental health chatbots, as users from a study have claimed the chatbots don't have cultural awareness (Falade, n.d). Eri responds to these patterns by providing community organizations for people of color in the app to increase cultural awareness, having the chatbot function mostly through system prompts to mitigate an overly agreeable nature, and spotting triggers in order to provide emergency contacts that can aid in severe mental health cases.

Methodologies and Methods:

Methods:

Iterative Design

The Iterative Design process I'm using comes from Eric Zimmerman's definition of Iterative Design, which is a process of prototyping, testing, analyzing and refining a work in progress (Zimmerman, 2003). How I have used this method is through having my first chatbot prototype being tested at Spark (which is an informal exhibition where people can engage with your work in progress). Based on the exhibition feedback I got and my reflections on them, it informed my future prototypes such as the software I wanted to use and how the system should be structured for users. The things I learnt from the various design iterations I made is how to design in a way that is simple for a user to engage with. For example, one of the design iterations I made was making the activities be one by one, with a separate message with all the resources after the activity, instead of one long message with the activities and resources. Through feedback from others, this iteration taught me how people value simplicity and convenience when using a technology.

Design Through Research

Design through research is here as a pre-positioning for a method I would use for future development of the project, as all the other methods used set the project up for design for research. One design for research method I used to evaluate my project is comparing the search engine results of the therapy resources recommend from google and ChatGPT. This framework impacted the results for therapy resources I provide after the activity is completed. I searched therapy resources for racialized communities on google and ChatGPT. For google it recommended directories such as Black Therapy List and Healing in color to find a therapist, but

not an actual therapy clinic. For ChatGPT, it also recommended the same directory Healing in Color and others such as BIPOC Therapist Canada. This process helped me with my resources being specific to actual clinics instead of just directories. This relates to design through research as research is a part of designing products and design activities is a part of generating and communicating knowledge (Stappers and Giccardi, n.d). The CBT therapy activities in the app is a way of communicating knowledge to its users.

User Personas

To get a sense of who the future user could be, I created a user persona of a young black women as an example of a racialized person that could possibly use the app. I represent the user persona which is featured in the short film. This black woman is experiencing racism and sexism in the work place, which is making it hard for her to afford therapy. She's also experienced bias and a lack of support when trying in person therapy but is hesitant of ChatGPT for therapy because she doesn't trust it to provide therapy without bias. I based the persona on a study of four racialized groups (Black, Hispanic, Asian and Native American) that took a survey on their use of ChatGPT (Stewart et al, 2024), and Hispanics were the most trusting out of all of the groups of the information from AI. The persona is also based on Safyia Noble's Algorithms of Oppression where she discusses tech bias amongst black women (Noble, 2018), as well as Waltons article about the barriers that black women face in society and in finding proper mental health care (Walton, 2023). This informed the creation of Eri to be culturally adapted in its CBT exercises with no bias and provide therapy resources and organizations for racialized people to build trust.

Methodologies:

Critical Race Theory

Critical Race Theory is my other methodology I am using, which comes from the book called “Critical Race Theory fourth edition,” by Richard Delfago and Jean Stefancic. The authors define Critical Race Theory as the relationship between race, racism and power (Delfago and Stefancic, 2025). It’s a movement that is built off of feminism and critical legal studies which is used across various disciplines (Delfago and Stefancic, 2025). In regard to mental health, critical race theory has been used to describe how inequalities happening to marginalized communities can negatively impact their mental health more than non-marginalized communities (Talbert and Ren, 2025). I’m using critical race theory to see how Eri can be designed to benefit racialized people through the resources it provides, since they’re at a disadvantage in comparison to non-marginalized groups. Critical race theory helped me with highlighting a specific user case scenario in the short film and discussing how the issues a user is facing with therapy and AI technology would lead them to using Eri. It also helped with implementing the CA-CBT framework with the activities, because the activities are the same as CBT activities, but the cultural lens it’s coming from makes it more inclusive to everyone.

Prototypes:

Chatbot Builder Text Prompt and Web Design:

The first prototype I made was based on research about AI chatbots and existing AI mental health chatbots that are out there. I didn't use AI for this, because I wanted to make something that acted like AI but wasn't in order to test how it would flow. I put an intro message that shared the intent of the chatbot but gave a disclaimer about it not being a

therapist and to call the 988 hotline if you are in serious need of help. I used Chatling for the chatbot and HTML/CSS for the web design. The framework for the activities were general therapy activities from websites on common therapeutic exercises. From the spark event where people interacted with it, I noticed it lacked sophistication in terms of system prompts with how it responded to users. There also wasn't a specific user case in mind yet. I decided I would need an AI that I can fully customize to provide the experience I'm looking to do and still be regulated and controlled, as well as figuring out who the user of this interface would be.

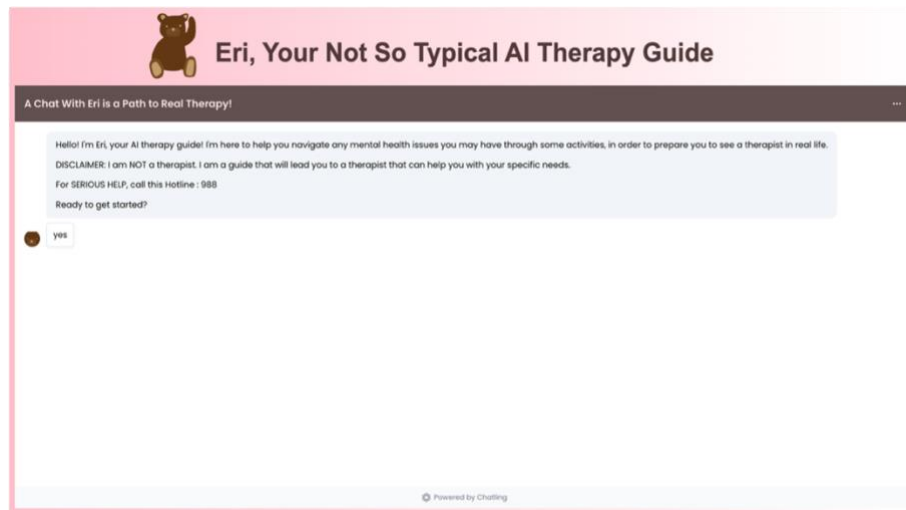


FIGURE 1: WEB INTERFACE

AI customizable prompts using Ollama:

I decided to use Ollama for my second prototype, because it is an AI tool that can be customized and I wanted the chatbot to have specific prompts on how it would work. It's using CBT as its framework, because from my case study, other mental health apps with chatbots use this as their framework. The chatbot is customized to give out short CBT based exercises, a link

to what CBT is, as well as a link to a local clinic after completing the exercise. It also flags mentions of suicide, by providing you the 988 hotline and other crisis hotline supports. This feels more sophisticated, but I want the resources to be more tailored to Black young adults, as I decided upon on who the therapy awareness is for. For this prototype, activities kept repeating itself you completed it and I couldn't verify where these therapeutic responses were coming from, so I knew that my next prototype had to make the responses more structured from a reliable source. Below is a photo example of the activity.

```

erinn@ErinmaArkisAir Local LLM % python3 main.py
My name is Eri, your AI therapy Guide. I'm here to guide you through CBT-based activities to help you seek a licensed therapist for ongoing care! Reminder: I am NOT a therapist. For information on CBT, click here: https://www.apa.org/ptsd-guideline/patients-and-families/cognitive-behavioral Type 'exit' to quit.

You: I feel anxious about the future

Eri: I'm so sorry to hear that you're feeling anxious about the future. That can be really overwhelming. Try this 5-minute thought challenge:

1) Write down 3 things you're looking forward to in the coming weeks or months.
2) Next to each item, write a specific action you'll take to make it happen (e.g., "apply for that job by Friday").
3) Reflect on how these small actions can help build momentum and confidence.

Duration: 5 minutes. Rationale: Focusing on positive future possibilities and breaking them down into manageable steps can help shift your anxiety towards a sense of control and possibility.

Remember, I'm not a substitute for a licensed therapist – consider reaching out to the Centre for Addiction and Mental Health (CAMH) at https://www.camh.ca/ for ongoing support in Toronto.

Local clinic recommendation: CAMH – https://www.camh.ca/

You: I'm looking forward to christmas, new years and a break from school. I'll start christmas shopping, making plans for new years and doing hobbies. This can give me something to look forward to

Eri: I'm glad you're thinking ahead to the holidays! Try this 5-minute holiday planning exercise:

1) Write down 3 things you'd like to do or experience during the break (e.g., "try a new recipe," "watch a favorite movie").
2) Next to each item, write a specific plan for how you'll make it happen (e.g., "make a list of ingredients and try cooking on Christmas Eve").
3) Reflect on how these plans can help create a sense of structure and enjoyment during the break.

Duration: 5 minutes. Rationale: Having a clear plan can help build excitement and anticipation, reducing feelings of anxiety or overwhelm.

```

FIGURE 2: ACTIVITY

To add further onto this prototype, the responses from the chatbot weren't verifiable so I decided that the third prototype would have to have responses come from a reliable source in order for it to be reputable and safer to use. I decided to use Positive Psychology for my third prototype.

Flask Web and App Interface:

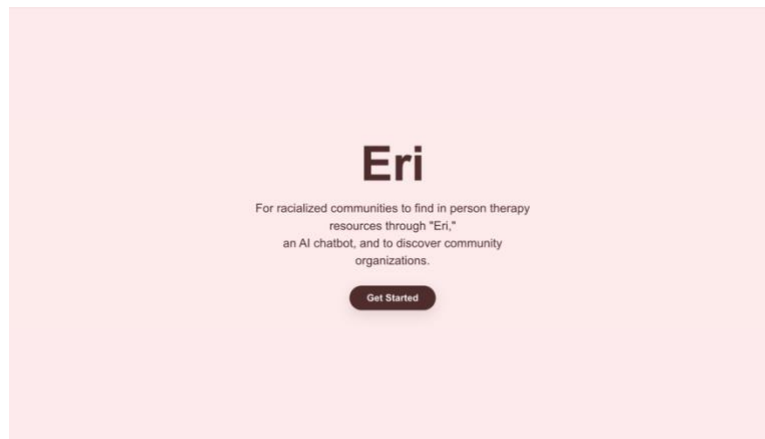


FIGURE 3: HOME PAGE

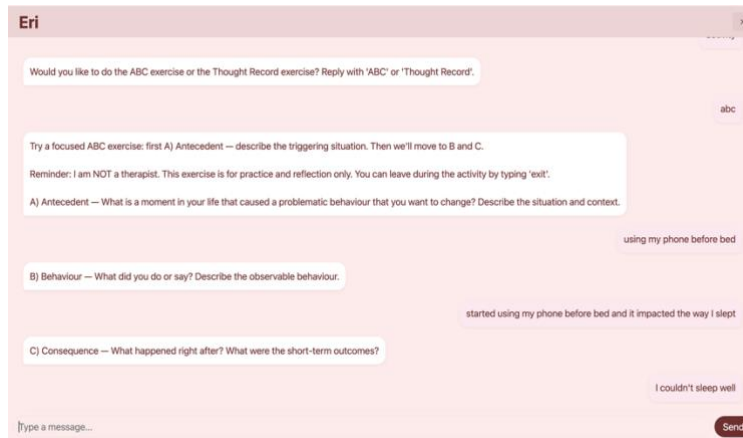
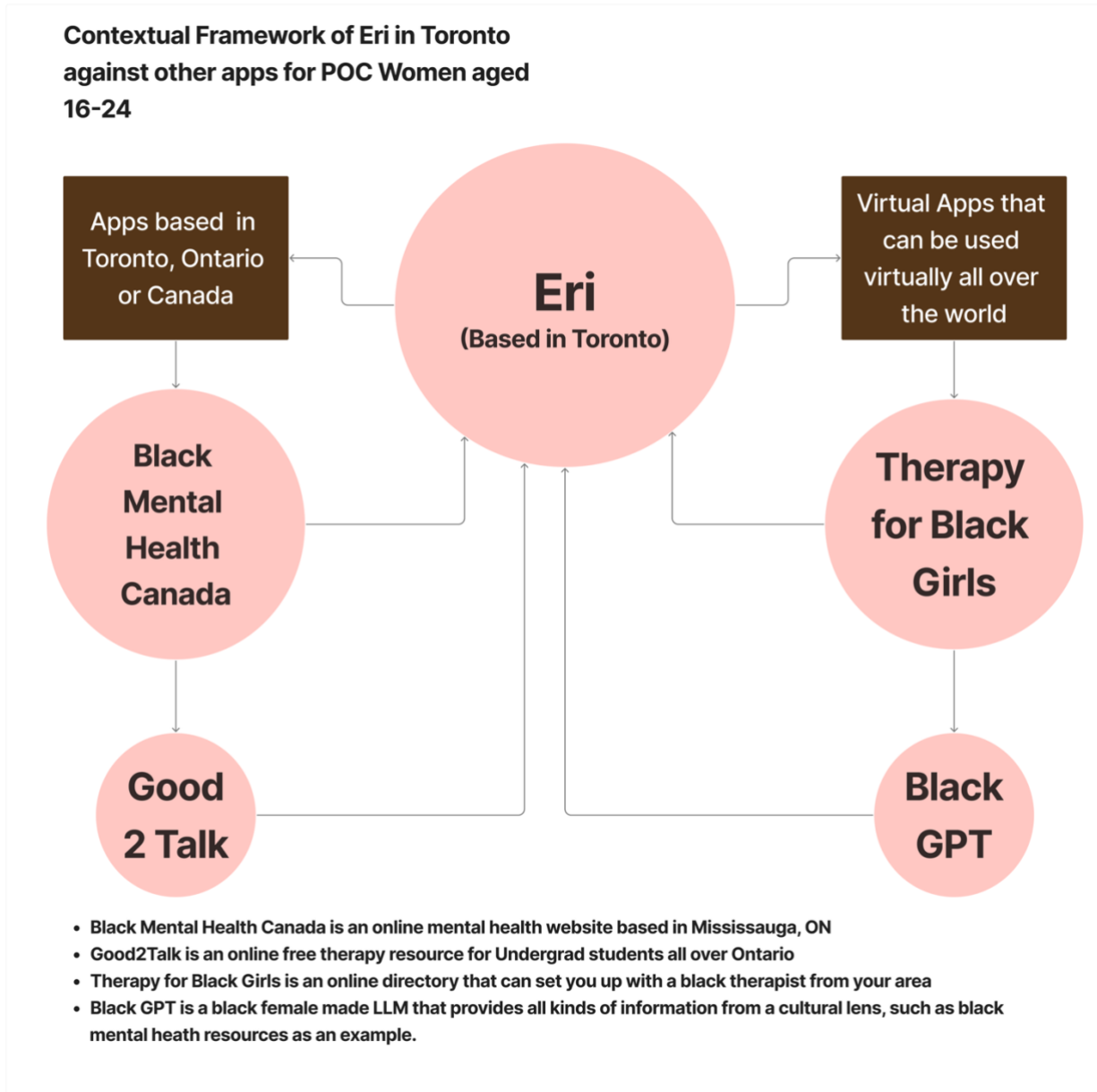


FIGURE 4: ACTIVITY PAGE

These images above show my third and final prototype after various iterative designs. The activities that Ollama is generating is now from a reliable source, Positive Psychology. The web and app application are being run by flask, and ngrok is being used to make the interface work on mobile through a link or qr code. The activities initially had six questions per activity for one to completed but based on voluntary informal feedback I got from a school organized playtest event, the activities felt overwhelming and there was an element of choice that wasn't included. I also got feedback on a lack of prompts available for beginners who were unaware with the therapy terminology. I had also found the type too small when it was displayed on a monitor, and on when using it on mobile. The changes I made was to make the text bigger on my desktop that way when it displays on a monitor, it makes the text look proportional and more readable. The font size change on mobile also helped with readability. I changed the structure of the activities with how it's sent out, by having three messages per activity that come out one by one after each response, that way it's less overwhelming to do the activities and it goes by quicker. I now have a message that gives people two options of which CBT

exercise to do, as well as prompts such as “what is cbt,” “what is abc” etc., that people can ask to learn about therapy terminology. I also have an “exit” prompt that is available throughout the chat if they want to end the conversation and explore the community organizations page. The system is able to flag triggering language relating to suicide in order to immediately give you emergency contacts as well as a “help” prompt if you need help on how to use the system or if you need further support by being in crisis. This prototype worked for my project because it has a user case in mind, it has system prompts to ensure safety and make it bounded towards a specific purpose, it’s providing factual therapy activities from a reliable source instead of random activities like the last prototype, and it shows how an AI chatbot can be used as an awareness tool to therapy resources for racialized people by providing actual clinic names instead of therapy directories like the other search engines.

Contextual Map of Eri



Appendix A: Contextual Map

Narrative Short Film Description:

The narrative short film is about what this app could look like in a future in 2029 where using AI for therapy is the norm. The narration brings about ideas of a everyone is using AI for therapy and questioning how it would behave. There are visuals of me as a user, and through the narration I represent a black female user that is experiencing current issues regarding therapy and with using AI therapy chatbots. An example of an issue I narrative in the film is of black women experiencing racism and sexism in the workplace, which impacts their ability to build upon their wealth and to afford proper therapy (Walton, 2023). I also highlight how black women have experienced different issues in therapy such as a lack of support (Walton, 2023) and tech algorithm bias (Noble, 2018). The story concludes with how "Eri" could be a solution and shows the user using the app as a better alternative. The role of the short film is to highlight research of specific issues that a racialized person is facing with therapy and AI, to give background as to why a racialized person would use the Eri app in this imagined future.

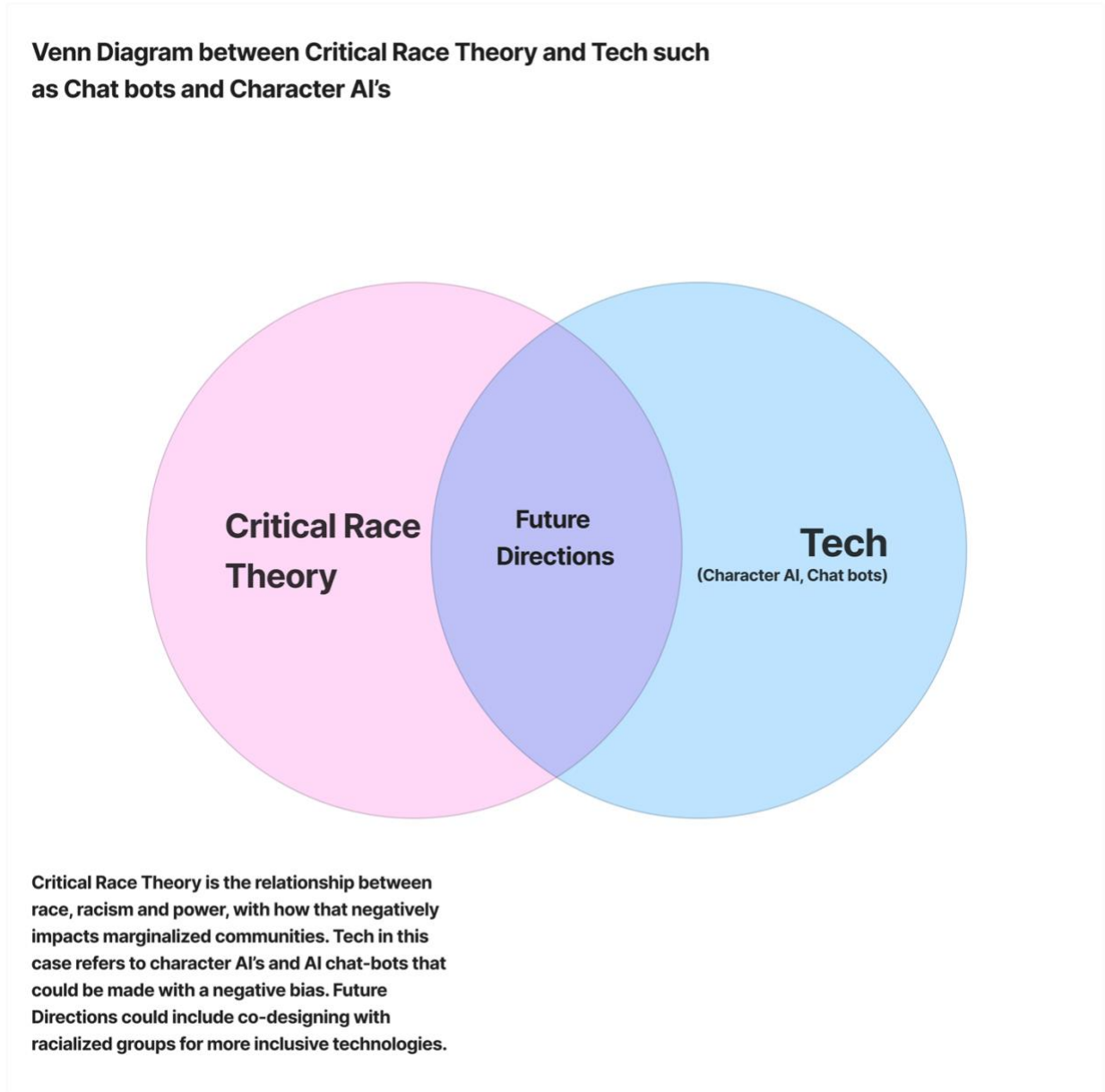


FIGURE 5: IMAGE FROM SHORT FILM [HTTPS://WWW.YOUTUBE.COM/WATCH?V=0Y5HF7YLDMG](https://www.youtube.com/watch?v=0Y5HF7YLDMG)

Conclusion:

In conclusion, the ethics of AI chatbots for therapy has been an ongoing discussion since the first chatbot creation back in 1966. In modern times, the ethical risks of using AI chatbots for therapy have caused alleged harmful consequences for users as well as concerns of the kind of therapy it can provide for people of color without bias. Eri demonstrates an AI chatbot that is made with constraints such as pre-set system prompts in order to be ethical and not an unregulated conversational Open AI. The short film about Eri demonstrates a user case scenario in an imagined future where AI therapy is the norm and explore how this system could be something positive for racialized people in that future. Eri also demonstrates a system made with cultural attunement for racialized people without bias, through including a page with the appropriate community organizations. My observations from the thesis exhibition of this app and short film show the potential that the app can actually have in real life, as people were open to doing activities and were made aware of resources for the first time through my conversations with people using the app. Having two therapists from different specialties (Body movement and Trauma Therapy) interacting with the prototype and finding it better suited than ChatGPT also shows the need for technologies like this as a safer alternative. My thesis is contributing to the ethics conversation of using AI for therapy but offering a solution through the AI chatbot prototype to identify how racialized people can be directed to better resources.

Future Directions:



Appendix B: Future Directions Ven Diagram

In this diagram, the overlaps between the two are technologies such as chatbots or character AI's made with the interest of racialized people in mind for all kinds of awareness of resources such as mental health as an example. This helps to go against the systemic structures that are not made equal for all races, as black women for example struggle with wealth gaps,

negative stereotypes, and racism in health care and the work place (Walton, 2023).

The risks involved could be these technologies being unable to recognize speech dialects, which can result in negative tech biases. The future directions to prevent this can include computer scientists collaborating with racialized groups when testing and creating these technologies, in order to make it more inclusive. For future development of Eri, I would want to expand upon design through research by co-designing with therapists on different therapy activities, as well as having computer scientists develop Eri into a fully-fledged LLM that can provide mental health resources for black women aged 16-24 depending on their location.

Exhibition Photos:



FIGURE 6: FULL DISPLAY



FIGURE 7: USERS AT EXHIBIT

References

Clinical Practice Guideline for the Treatment of Post-Traumatic Stress Disorder (2017). What is Cognitive Behavioral Therapy? *American Psychological Association*

<https://www.apa.org/ptsd-guideline/patients-and-families/cognitive-behavioral>

Freitas D, Ugrulap K, Puntani S, (2023) Chatbots and Mental Health: Insights into the safety of Generative AI

<https://myscp-onlinelibrary-wiley-com.ocadu.idm.oclc.org/doi/epdf/10.1002/jcpy.1393>

Luo, X (2025) Shaping ChatGPT into my Digital Therapist. *Sage Journal*

<https://doi.org/10.1177/20552076251351088>

Canaday V, (2025) Human Therapists outperform AI Therapists in delivering CBT

<https://onlinelibrary.wiley.com/doi/epdf/10.1002/mhw.34464>

Degerli Y, Degerli M, et al. (2025) Artificial Intelligence in occupational therapy: Comparing ChatGPT and occupational therapist approaches in assessment and intervention

<https://doi-org.ocadu.idm.oclc.org/10.1177/03080226251368231>

Stappers, P. and Giaccardi, E. (2014, January 1). Research through Design. Interaction Design Foundation - IxDF. <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design>

Pirnay, E (2023) We Spoke to People Who Started Using ChatGPT as Their Therapist

<https://www.vice.com/en/article/we-spoke-to-people-who-started-using-chatgpt-as-their-therapist/>

Proudfoot, D (2013) Re-Thinking Turing's Test

<https://www-jstor-org.ocadu.idm.oclc.org/stable/43820781?seq=2>

Tarnoff, B (2023) Weizenbaum's Nightmare's: How the inventor of the first chatbot turned against AI

<https://www.theguardian.com/technology/2023/jul/25/joseph-weizenbaum-inventor-eliza-chatbot-turned-against-artificial-intelligence-ai>

Mental Health Redefined (2026)

<https://www.wysa.com/>

Artificial Intelligence for Mental Health (2026)

<https://www.youper.ai/>

Tangerman, V (2025) Women Kills herself after talking to Open-AI Therapist

<https://www.msn.com/en-us/news/technology/woman-kills-herself-after-talking-to-openai-s-ai-therapist/ar-AA1KOyFF>

Yousif, N (2025) Parents of teenager who took his own life sue Open-AI. BBC

<https://www.bbc.com/news/articles/cgerwp7rdlvo>

One man's experience with Wysa (2023)

<https://blogs.wysa.io/blog/company-news/one-mans-personal-experience-with-wysa>

The Story of ELIZA (N.D) The AI that Fooled the World. London Intercultural Academy

<https://liacademy.co.uk/the-story-of-eliza-the-ai-that-fooled-the-world/?v=5435c69ed3bc>

<https://www.youper.ai/how-it-works>

Noble, S (2018) Algorithms of Oppression. How Search Engines Reinforce Racism

https://safiyaunoble.com/wp-content/uploads/2020/09/Algorithms_Oppression_Introduction_Intro.pdf

Persuad, C (2022) Book Review: Intersectional Tech: Black Users in Digital Gaming by Kishonna L. Grey.

<https://doi.org/10.1177/08912432221078525>

Falade, O (2024) Serving Whom? Ethical and Practical Limits of AI Mental Health Chatbots for Marginalized Communities

<https://ojs.stanford.edu/ojs/index.php/grace/article/view/3818/1783>

Zimmerman, E (2003) Play as Research: The Iterative Design Process

https://ericzimmerman.com/assets/pdfs/Iterative_Design.pdf

Stappers J, Giaccardi E (n.d) Research Through Design. Interaction Design Foundation

<https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design#example>

Delgado R, Stefancic J (2025) Critical Race Theory

<https://books.google.ca/books?hl=en&lr=&id=xU6VEAAAQBAJ&oi=fnd&pg=PA19&dq=critical+race+theory+&ots=luhy1sPO8x&sig=64QGQ0mce5Hi2fG3NAYBbxm4pSE#v=onepage&q=critical%20race%20theory&f=false>

Giuseppe Giorgianni, et.al (2024), Unveiling The Potential Of Ai: Impacts On Industries And Ethical Considerations, *Educational Administration: Theory and Practice*, 30(3), 2654 -2661
Doi: 10.53555/kuey.v30i3.7484

Styker, C (n.d) What is Artificial Intelligence (AI). IBM

<https://www.ibm.com/think/topics/artificial-intelligence>

Mental Health Academy (2024), Innovative Approaches to Cognitive Behavioral Therapy

<https://www.mentalhealthacademy.com.au/blog/innovative-approaches-to-cognitive-behavioural-therapy>

Kwame Mckenzie et.al (2024) Culturally Adapted Cognitive Behavioral Therapy (CA-CBT) for Black Populations: A Manual for Mental Health Practitioners. CAMH

<https://www.camh.ca/-/media/professionals-files/cacbt-for-black-populations--a-manual-for-mental-health-practitioners-2024-pdf.pdf>

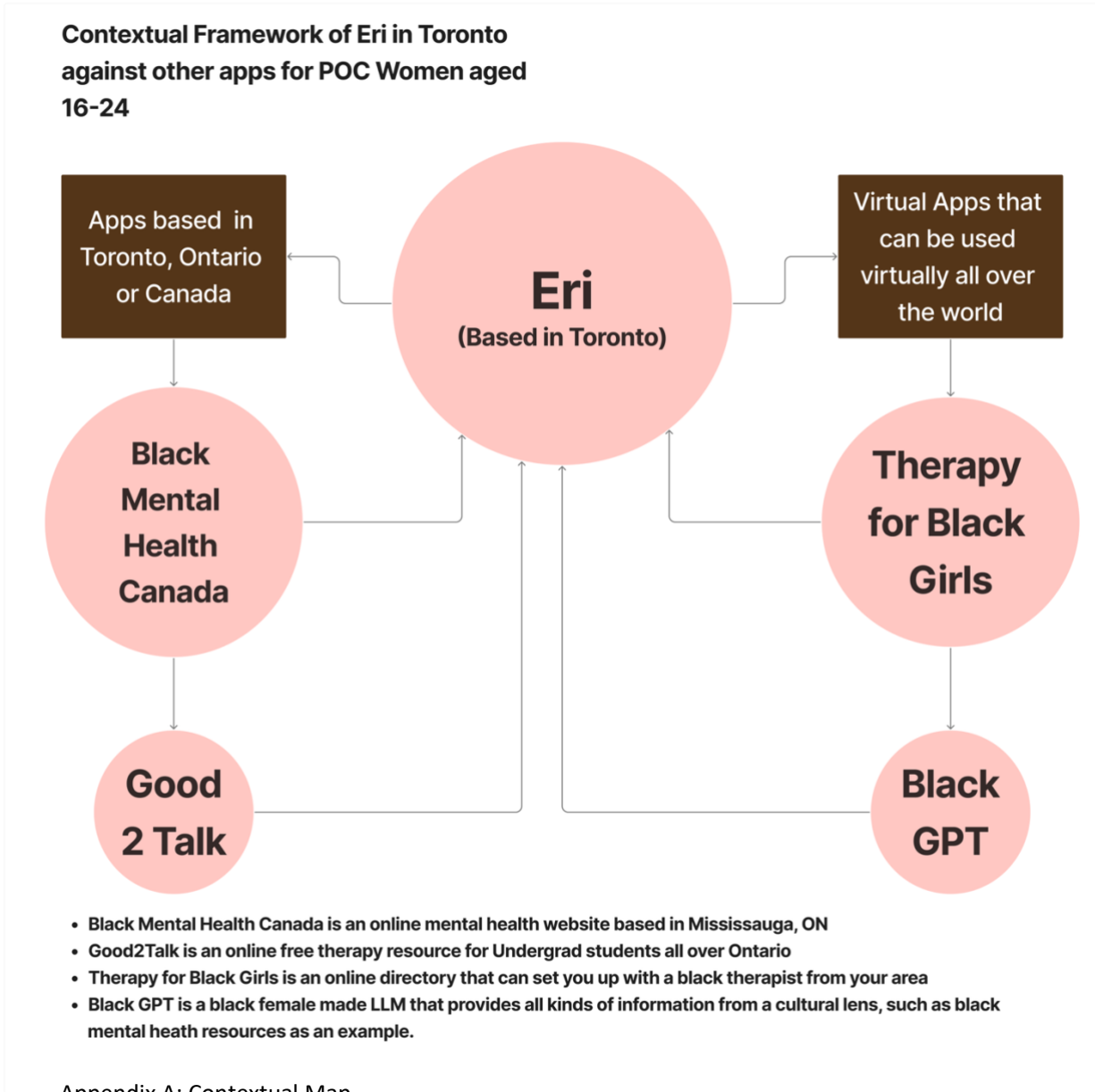
Talbert R, Ren J (2025) Race, Racism and Mental Health. *The Blackwell Encyclopedia of Sociology*.

<https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781405165518.wbeos1065.pub2>

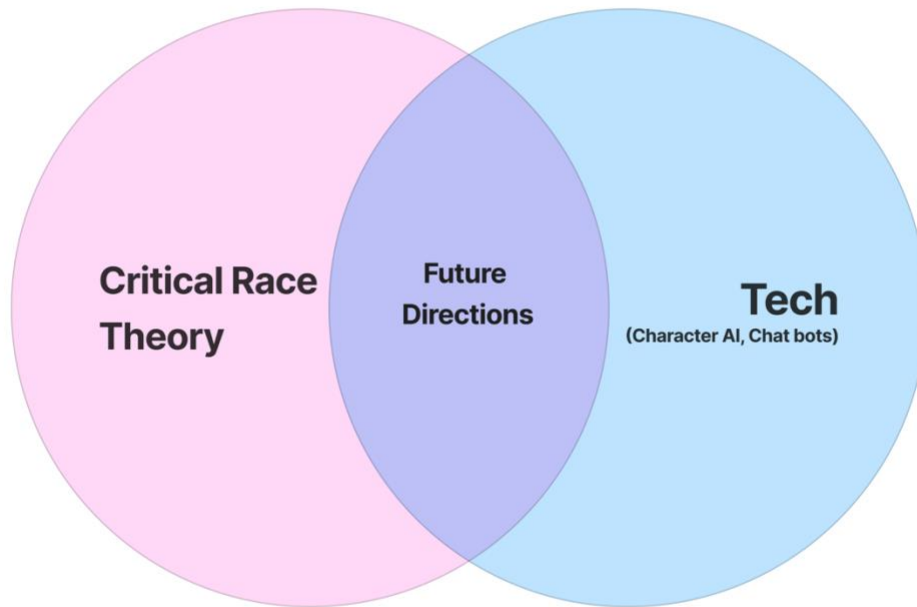
Walton, Q (2023) Improving Mental Health Care Access for Middle Class Black Women. *Urban Institute*

https://www.jstor.org/stable/resrep69022?searchText=gaps+in+mental+health+services+for+black+people&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3Dgaps%2Bin%2Bmental%2Bhealth%2Bservices%2Bfor%2Bblack%2Bpeople%26so%3Drel&ab_segments=0%2Fbasic_search_gsv%2Fcontrol&refreqid=fastly-default%3A7c586c3c9625d0cc20f92c2b58d7fce4&seq=1

Appendices

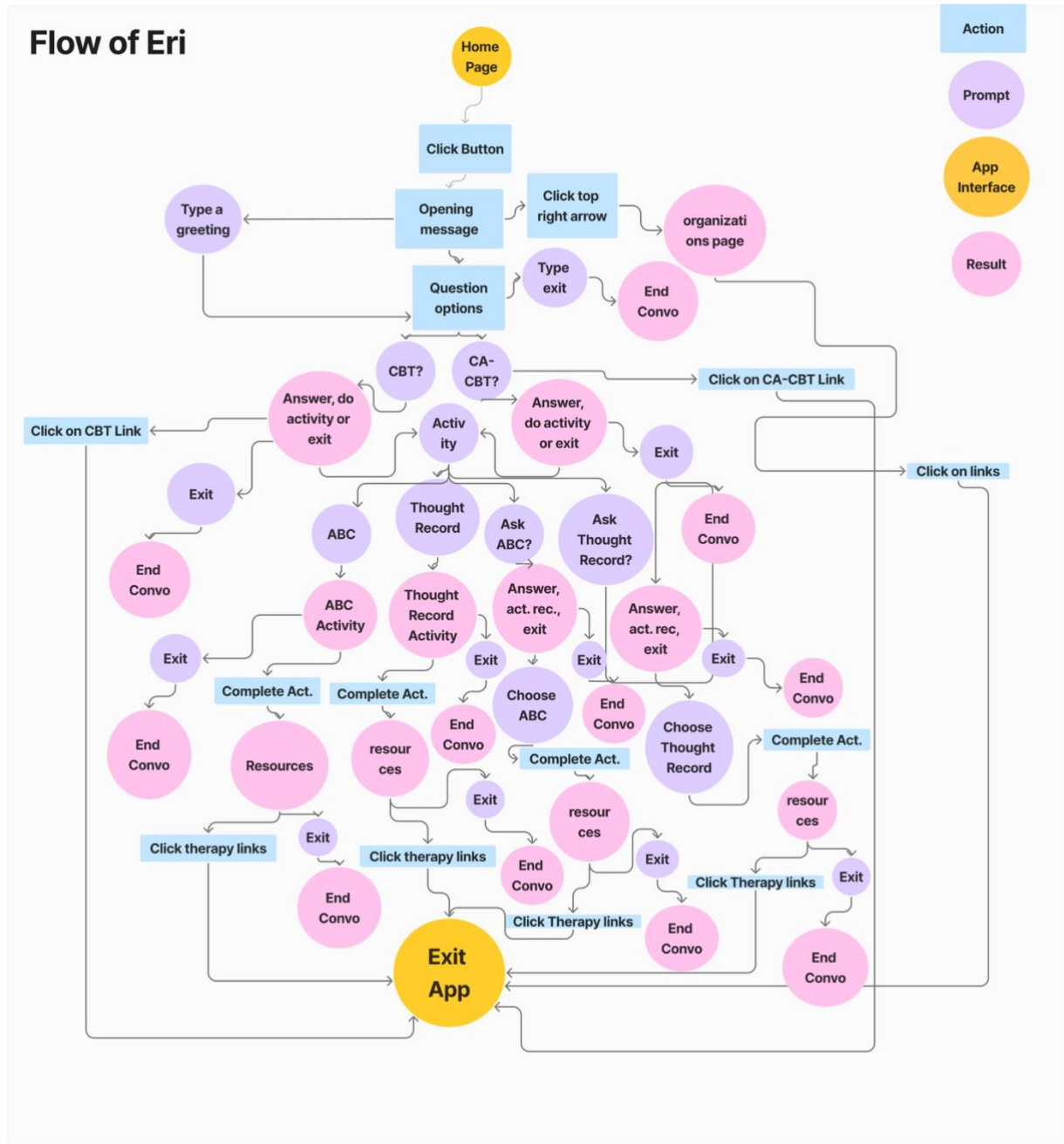


Venn Diagram between Critical Race Theory and Tech such as Chat bots and Character AI's



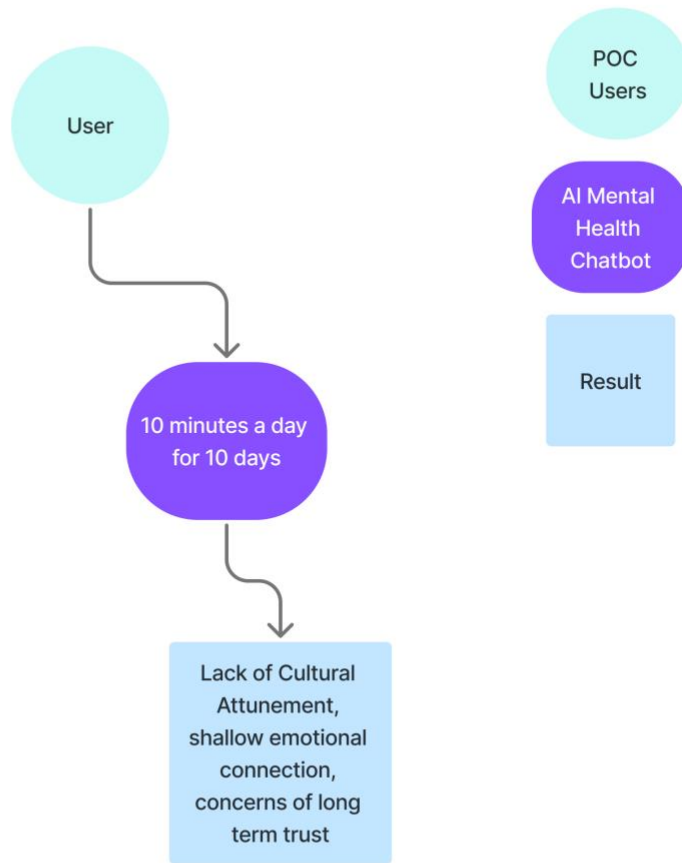
Critical Race Theory is the relationship between race, racism and power, with how that negatively impacts marginalized communities. Tech in this case refers to character AI's and AI chat-bots that could be made with a negative bias. Future Directions could include co-designing with racialized groups for more inclusive technologies.

Appendix B: Future Directions Ven Diagram



Appendix C: Flowchart of Eri App Interface

Flow chart showing a study of Racialized People's experience using AI Mental Health Chatbots (Falade, n.d)



Appendix D: Flowchart of Racialized People using other AI Mental Health Chatbots

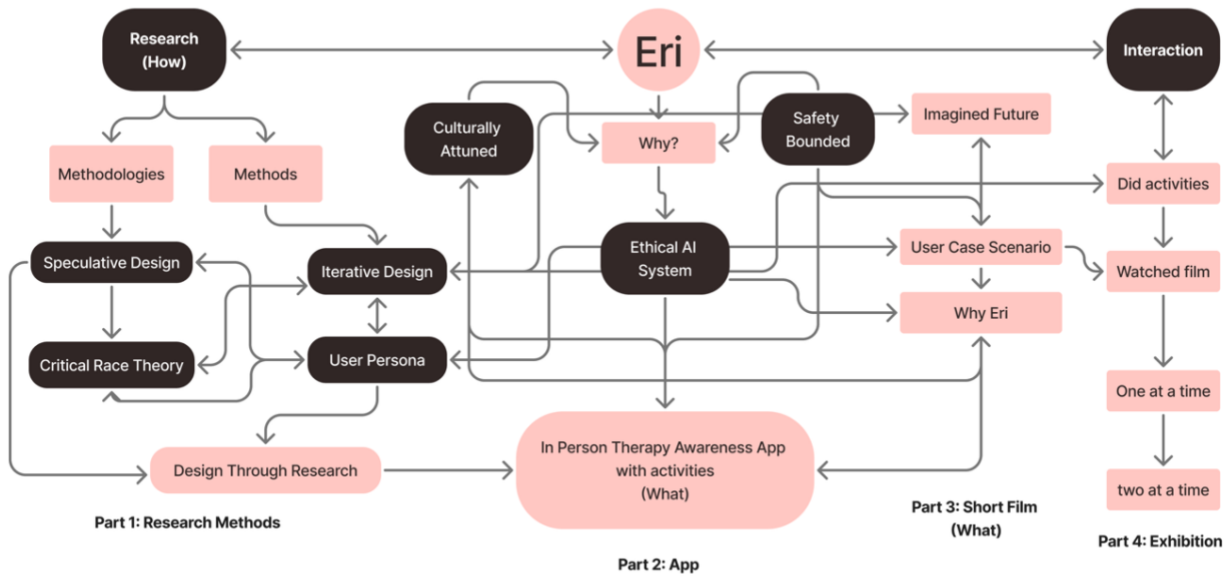


Figure 1: Conceptual Framework

Appendix E: Conceptual Framework