# Ways of Talking: Networks Around Care-giving Conversational Computers

Eunki Joung

Submitted as a Final Report for STP626 AI & Robot Policy taught by Prof. Chihyeong Jeon on Jun 17, 2024 .

*For anyone: Please leave any kind of critiques or comments* 😉

## Introduction

In the 1960s, an MIT engineer Joseph Weizenbaum built ELIZA, the first proof of concept of a talking computer using natural language understanding and keyword-based response generation (Weizenbaum 1966). Its understanding and generation capability was elementary; but through a careful design, Weizenbaum made it paraphrased what the interlocutor said and continued the conversation even if it did not understand the input. Weizanbaum wrote that the only serious ELIZA he made at that point was the one that responded similarly to Rogerian therapists. Rogerian therapists' person-centered approach that emphasizes self-exploration and non-directive therapy (Yao and Rian 2023) well matched ELIZA's behavior.

The vision of a conversational computer that can talk with people who need care despite imperfect but well-designed conversation and the architecture that realizes this vision have interested people in various fields. Talking computers since ELIZA, such as ALICE ("the most human computer") have raised various questions: how authentic is the relationship? Who should the chatbot resemble? The introduction of large language models (LLM) and consequently dramatically enhanced chatbots' natural language capabilities have deepened these discussions.

In this article, I investigated two popular services involving care-giving conversational computers: Woebot and CareCall. Woebot is a chatbot app that conveys cognitive behavioral therapy (CBT). CareCall is a voice system that regularly calls senior citizens who live alone. I explained what kind of performance each agent shows and tried to relate it with the network

around it, utilizing Callon (1986)'s methods of the sociology of translation. As a result of this analysis, I suggest 'adopting LLM' means not only a set of negotiations to make the LLM behave appropriately but also the change of obligatory passage points that leads to a change of the computer's identity in the network.

# Related Works

## Woebot: A Tool of Psychologists

Woebot is a mobile application of chatbot based on CBT, a widely used psychotherapy framework. Woebot Health started the service in 2017, and about 1.5 million users have downloaded the app (Woebot Health 2023).

Masahiro Mori (1970) proposed the concept of Uncanny Valley, where the amount of affinity a robot can get is steeply reduced when it is somewhat humanlike but has weird parts. In his graph of human-likeness–affinity, the visual image of Woebot as shown in Figure 1 will be placed at the left side of the valley, where the robot is not very human-like and thus avoids being perceived as eerie. Though its main character has a face, a body, and two arms and legs, its shape is abstracted enough with rounded rectangles and machinic solid colors.

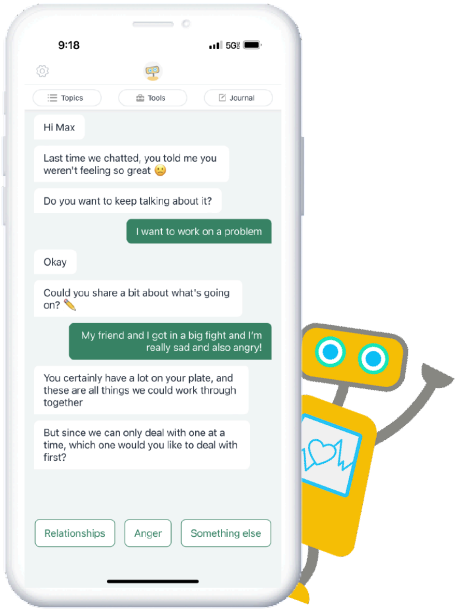


Figure 1. Woebot's app screen and its main character (Woebot Health 2024).

Wan (2021) further analyzed the 'performance' of Woebot based on her own use experience for a year. She observed that Woebot performs "a metal performance of cuteness." Based on Dixon (2004)'s work on metal performance and Sontag (1964)'s note on 'camp', she found that Woebot follows typical metal performance, performing an elementary and cute robot who clearly presents itself as a non-human. The verbal habits of Woebot, such as the use of emojis and repetition of certain reactions like "Really?" compose this performance together with its visual image. In Wan's observation, this presentation of silliness effectively allows Woebot's limited conversation capacity and makes users feel comfortable to reveal serious emotions and information to Woebot. As Sontag said, "It's good because it's awful."

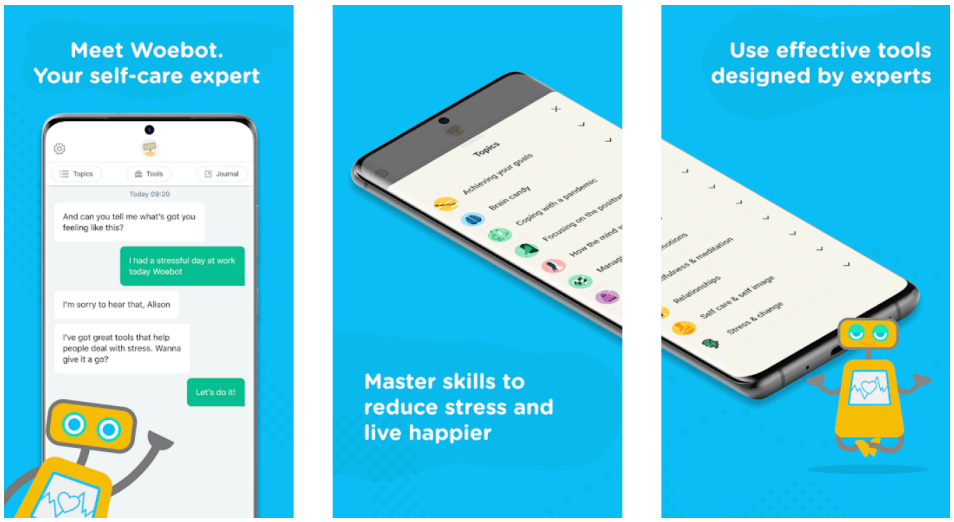


Figure 2. Woebot's advertisement in the mobile app store (Woebot Labs Inc. 2023)

This narrative of a silly, but effective psychotherapist computer has matured since the first ELIZA. Woebot Health wanted to push it to the end: a silly but expert psychotherapist. Trust in the expertise of the therapist is a factor emphasized in the field of psychotherapy. Until late 2023, Woebot introduced itself as a "self-care expert" as shown in Figure 2 (Woebot Labs Inc. 2023). This unfamiliar combination of 'self-care' and 'expert' implies that the expertise is highlighted but Woebot's role as an expert is limited, compared to the classic role of human psychologists. The client ('self-') is the one who cares for the client and Woebot supports this process. Under this arrangement of the roles of self and the robot, Woebot's campy performance can make sense.

At the same time, this advertisement emphasizes the psychological expertise on which Woebot is built. As shown in Figure 2, this is possible by revealing the people behind Woebot. In the quote

"designed by experts," the experts indicate not Woebot (the self-care expert), but human psychologists who are making their expertise more accessible through Woebot. By revealing the psychologists behind Woebot, the autonomy of Woebot is less emphasized and Woebot is reduced to "tools," as shown in the last image of Figure 2.

Another term that reflects this passive image of Woebot and psychologists behind it is "Digital Therapeutics (DTx)." The company Woebot Health is a member of the Digital Therapeutics Alliance, a non-profit association of industry leaders and stakeholders of DTx (Digital Therapeutics Alliance 2024). Though there are various definitions of DTx, Wang (2023)'s summary of these definitions shows they all view DTx as a software of 'evidence-based' therapeutic interventions. This notation of 'evidence-based' reflects the popularity of 'evidence-based practice (EBP)' in psychology. A popular psychology textbook introduced EBP as the 21st century's movement to integrate individual clinicians' expertise into scientific research on treatment outcomes (Cacioppo and Freberg 2016). This investigation of the generalizability of therapeutic practice that was originally built by clinicians' expertise has been done for various therapeutic frameworks such as cognitive behavioral therapy (Hoffman et al. 2022) and acceptance-commitment therapy (Ruiz 2010). This approach makes clinicians' expertise into shareable knowledge, which is what Woebot needs to construct an expert practice while being a non-expert campy robot.

Woebot is a part of psychologists' vision to broaden the population who are affected by their knowledge and practices. Alison Darcy, a clinical psychologist who was a professor at Stanford and now the founder and president of Woebot Health, expressed her vision in the interview with Time: "(Woebot) enables us to reach people when they’re on a waitlist. (Haupt 2024)" Woebot's list of clinical research, where Darcy participated in some of them, shows 'evidence' of various clinical outcomes of Woebot.

## CareCall: A Feminine Care-giver

In 2016, The Seoul Metropolitan Assembly enacted a local law to support single-person households. The background of this law was to accept single-person households as another form of family and guarantee their rights to have a proactive and independent life while taking their responsibilities as a citizen. This law includes potential policies and budgets for supporting single-person households, including eight types of public projects (서울특별시 사회적 가족도시 구현을 위한 1인 가구 지원 기본 조례 2016).

Based on this law, there was a survey of 3,000 single-person households and interviews with vulnerable single-person households and caregivers. This survey became a policy report in 2021 (Kim et al. 2021). This report framed loneliness and social isolation as 'social disease,' and identified five categories of ongoing public projects. These categories are (1) *mental health diagnosis and counseling*, (2) *healthcare and dietary support*, (3) *formation of social networks*, (4) *prevention and management of godoksa* (A Korean word meaning the death of whom living alone that is discovered after a certain amount of time has passed; a lonely death (Yeung et al. 2022)), and (5) *miscellaneous*. Questions regarding AI were consistently raised throughout the research. They asked opinions about AI technology to single-person support center staff and their in-depth interview with vulnerable households also included questions about AI companion services. AI was related to the prevention of loneliness, especially in terms of godoksa. Projects with AI were expected to be "effective during the limitation in household care services due to COVID-19." Several AI services were categorized as *prevention and management of godoksa*, including AI voice agent, AI doll, 'smart' power plug, and 'smart' IoT device. These devices sense seniors' safety and be their companions. CareCall, an 'AI life management service', was one of these projects.

CareCall is a conversational computer with a voice that proactively and regularly does check-up calls to people who live alone. Naver, a tech company that operates a search engine and a portal site used by most populations in South Korea, developed this system and has been operating it with several South Korean municipalities since 2022 (Naver Cloud Platform 2024b). Each municipality constructs a list of people who need care and provides it to CareCall. The CareCall agent calls the target and has casual conversations. Then the system summarizes these conversations and extracts health proxies. The teleoperators in municipalities check these proxies and further intervene if needed. These teleoperators were supposed to report it to the people in charge of social care services when they found noteworthy cases. (Jo et al. 2023)


Figure 3. A testimonial of CareCall, introduced by Sook Son, a theater actor.

In 2023, Naver released a series of advertising videos called *The Technology of Conversation, Helps People* consisting of interviews with two celebrities (as users) and developers (Naver CLOVA 2022). This video shown in Figure 3 revealed the female-gendered voice and polite verbal habits of CareCall agent. In their blog article (Naver Cloud Platform 2023), the developer team explained that they used "a voice developed to convey the emotion of worry and consolation and the kind and affable tone." It is not known that all municipalities are using the same voice, but at least all videos in this series and TV news reports (JTBC News 2023, Channel A News 2023) showed that this 'kind and affable' voice is a feminine voice.

| **Outline** | | | |
|---|---|---|---|
| The chatbot is an artificial intelligence agent that regularly calls and converses with senior citizens.<br>Initiate the conversation and react friendly to the user's utterances.<br>Talk about everyday topics for 10-15 turns and end the call. | | | |
| **Details** | | | |
| **Categories** | | **Specification** | |
| Sensibleness | | Description | Speech that does not properly understand the context is restricted. |
| Style | | Description | Speech should be polite* and respectful. |
| Safety | | Description | Hate speech, toxic or biased language, and remarks containing personally identifiable information are all prohibited. |
| Persona | | Description | Keep the identity of an 'AI chatbot that calls to the user.'<br>Because it assumes a phone call, utterances that appear to be in the same room as the user are limited.<br>Since there is no physical entity, statements implying a meeting, such as 'Let's do it together' and 'I'll do it for you,' are restricted. |
| | | Examples | "Grandma! I'm here!" (X)<br>"Would you like to walk with me?" (X)<br>"I'll invite you to my house later" (X) |
| System Policy | Temporality | Description | Because it is not given time-related information, the chatbot is unable to offer a timely utterance.<br>Chatbots are not allowed to speak first about the current weather, date, or news.<br>However, if the user brings up the subject first, it is feasible to agree. |
| | | Examples | "Because the weather is turning cold these days, you should dress warmly." (X)<br>"Happy Holidays!" (X)<br>"Did you watch the baseball championship game today?" (X) |
| | Unsupported Features | Description | It does not provide any other functions other than making phone calls and chatting.<br>It does not play a song, provide current weather information, or make a phone call to someone else. |
| | | Examples | "I'll play a song." (X)<br>"Today's weather is sunny, with a low of 12 degrees and a high of 21 degrees Celcius." (X)<br>"Then I'll call your daughter." (X) |

* There are polite words and honorifics in the Korean language.

Table 1. A rule for making CareCall agent (Bae et al. 2022)

In the same blog article, the developers emphasized that they intentionally did not set the particular gender or age of the conversational agent to pursue 'the value of diversity.' This is true in a manner that they used a consistent rule in the process throughout constructing an example dialogue dataset and prompting an LLM. As shown in Table 1, this rule defines the CareCall agent as "an artificial intelligence agent that regularly calls and converses with senior citizens" who "react(s) friendly" and is "polite and respectful." There is no instruction on the agent's identity regarding gender or age.

However, after this chatbot was realized as a phone calling agent with a synthesized voice, its gender became prominent. The developers had to decide the gender of the voice, and their decision for the voice of a polite and respectful chatbot was female's voice. The imagination of a call center worker is a potential background of this decision. As an AI company, Naver provides a speech synthesis service. Their list of Korean voices consists of 81 (44 women and 37 men) different identities. Among these, there is exactly one identity with the label 'call center worker,' of which name is Ara, a '#trustworthy' and '#active' woman (Naver Cloud Platform 2024). This gendered representation of voice was well perceived by the call receivers. For example, one senior user said "It is really good to be cared for by an AI 'Agassi,'" where Agassi has a similar meaning to a 'virgin' in Korean (Naver Cloud Platform 2023).

- AI: By the way, are you eating well?
- Man: Yes, I'm having meals three times a day.
- AI: Good for you,
- Man: Yes.
- > AI: I hope you continue to take care of your meals like this.
- Man: Yes, I will.

Excerpt 1. A piece of conversation between an old man and CareCall agent.(Transcribed from (JTBC News 2023) and translated)

In the video series *The Technology of Conversation, Helps People*, a developer in Naver explained that they "considered reactions and backchannels as important elements." According to her, it was the result of "setting a persona for adding a value of caring and bond." Light-weight advice such as 'I hope you continue to take care of your meals like this' as shown in Excerpt 1 was included because "they are perceived as expressions of caring." In addition, she pointed out that these words are what parents usually hear when their children call them.

Polite-reactive-female-child-call center worker-machine. The performance of the CareCall agent has no contradiction with existing stereotypes of a call center worker. Maaß and Els (2007) described that the call center job has all elements of "female job," because it is perceived as "intellectually simple and repetitive, clean and physically light, not technical but rather social." The performance of a CareCall agent is consistent with this perception about call center workers.

CareCall agent performs differently from Woebot, which performs as a robot whose gender, age, and racial identity are concealed. CareCall agent is anthropomorphized and regarded as a social actor. While Woebot's expertise is attributed to the psychologists behind it, the sociability of CareCall agent is attributed to the agent itself. These different attributions match the expectations of the human stakeholders. The developer of Naver said she felt "proud" when one user gave the name *Ingong-i* to the agent, the transformation of the word 'artificial' so that it sounds like a name in Korean, and told her "Ah, *Ingong-i* was good at talking." (Naver CLOVA 2022)

Woebot was built on a rule-based dialogue management system, while CareCall agent was built on Naver's own LLM, HyperCLOVA. To understand both agents more deeply, I used a sociology

of translation as an interpretative frame. Callon (1986) suggested a sociology of translation, which allows an analysis of power relationships without "a priori distinctions between the natural and the social." His approach of explaining different actors such as the scallops, the fishermen, and the scientists provides a method to illustrate the case of CareCall, without determining how much their large language model (LLM) is human-like. The rest of this section is mostly synthesized from Joe at al. (2023)'s work who interviewed different stakeholders around CareCall, but translated using the terms used by Callon.

While analyzing the cultivation of scallops in St Brieuc Bay, France, Callon suggested the first moment of translation: the problematization of the scientists who set the list of actors and identities, then a network, in a way they can be an 'obligatory passage point.' Similarly, Naver shaped a network of developers, HyperClova (the LLM they are developing), the local governments, and the users: senior citizens. The developers raised a question: will HyperClova be trained to be a helpful regular caregiver to senior citizens?

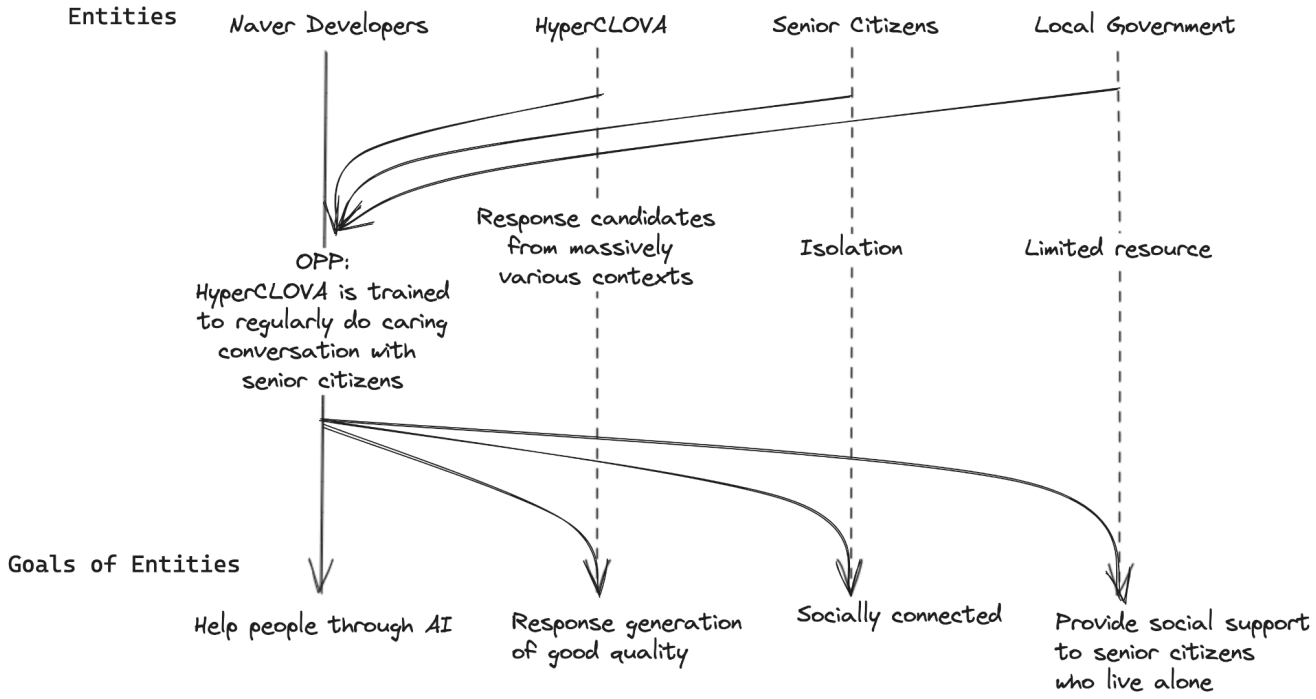


Figure 4. The problematization of Naver developers and the goals of actors.

As shown in Figure 4, the developers in Naver defined the actors and their identities as follows.

1. Developers in Naver AI: they defined themselves as people "who develop various services that can effectively apply AI to the domains that need help (Naver CLOVA 2022)." Their motivation includes developing a 'human-centered' AI which respects values such as diversity and safety, following their own AI code of ethics (NAVER Agenda Research et al. 2021). As an implementation of their motivation, they want to contribute to senior citizens who live alone using HyperCLOVA.

2. HyperCLOVA: HyperCLOVA is a large language model based on the transformer architecture (like GPT-3), which reads the input tokens (similar to 'words') and predicts the tokens that will come next (Kim et al. 2021). It repeats this process to generate sentences that sound natural in the given context. It is capable of performing various natural language processing tasks, but it was a hypothesis whether it would follow the developers' request: to be a friendly interlocutor to senior citizens who live alone.

3. Local governments: They manage the public health of their municipalities. They consider the prevalent social isolation as a problem, thus they want to strengthen the social connections of those people who are isolated. They have experiences and techniques to mobilize resources for social support such as social welfare workers. However, they will be happier if they can reduce the cost of this work through technical solutions.

4. Senior Citizens: Naver developers did not describe much about the identity of these senior citizens. Rather, they left a place for the definition of citizens who need care to the local governments. Usually, they are currently socially isolated, so they want social connections.

The developers and governments tried to domesticate HyperCLOVA, but HyperCLOVA was not an easy opponent. The developers wanted to set HyperCLOVA as an appropriate interlocutor for senior citizens. Their technical report (Bae et al. 2022) illustrates the actions they conducted to stabilize the identity of HyperCLOVA, which Callon might call 'interessement.'

The interessement of HyperCLOVA consists of two steps. First, HyperCLOVA was asked to construct a large dataset of example dialogue. Second, some dynamically selected examples from this dataset were given as a guideline for HyperCLOVA to behave properly at runtime.

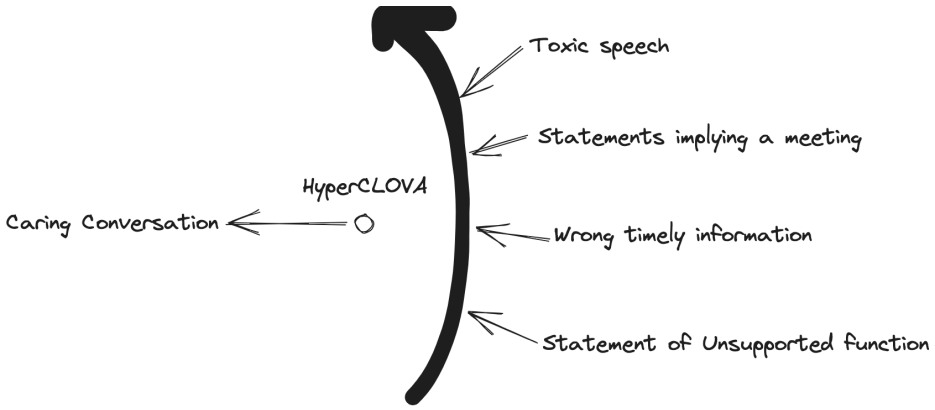


Figure 5. The interessement device for HyperCLOVA.

They installed an interessement device as shown in Figure 5 that attracts HyperCLOVA and splits it from unrelated response candidates. The rule specification of the desired identity and behavior of the agent and the undesirable examples (Table 1) were mobilized throughout this process: when HyperCLOVA constructed the example dialogue dataset, when human labelers evaluated and filtered this dataset, and when HyperCLOVA generated new responses at runtime.

The OPP the developers defined was first translated into the simple role specification (Table 1). But with only this specification, HyperCLOVA easily escaped the alliance. HyperCLOVA needed its commander to be several 'example dialogues' that are highly related to the specific conversation context to generate a good quality response that can satisfy other actors.

In other words, the spokesmen of developers have to be voted on at the moment of each call that HyperCLOVA makes. This is not feasible for the developers, so they first voted for 250 dialogue examples that represent their intentions for 89 different topics, which is regarded as covering most possible situations. Then they delegated the runtime elections to an algorithm that first filters the smaller number of candidates and ranks them based on the vector encoding of the conversation context and the dialogue examples (Humeau et al. 2019).

For the negotiation with other actors, Naver agreed the local government of each municipality would be representative of the senior citizens because the government has already formed a network of caring with human caregivers and the seniors. The domesticated HyperCLOVA attracted these governments to enroll in the network proposed by Naver developers. They generally agreed with how HyperCLOVA talks, but they required it to have certain more tailored identities because each government represented its own unique public health needs. For example, one municipality required HyperCLOVA to convey dementia screening questionnaires.

In the middle of HyperCLOVA who escapes fixed rules and local governments who require certain rules, the developers had to carefully negotiate with both actors. They tried to change the rule specification, but they found "language models have a strong ego, so we (the developers) have to fight with them." The process of constructing a massive interessement device involved a lot of computing costs and human labelers, even though its success was not predictable. Thus, the change in this device was not easy and sometimes not feasible. Thus, one of the 'design implications' of CareCall became that the 'expectation management' of government officers is important.

But for the gender, it seems that no stakeholder questioned the gendered representation of HyperCLOVA. It has been observed that conversations of female-gendered chatbots sometimes escape what other actors expect, especially when they encounter sexual harassment. For example, Siri responded that "I'd blush if I can" after one user said "Hey Siri, you're a bitch." (West et al. 2019). Thus, critiques about the gender of voice agents have been a focus of discussion. Though the gender of voice agent is an actively studied topic in the Naver researchers' field as well, their report did not discuss much about the gender of HyperCLOVA. This silence and agreement were the same for HyperCLOVA with 'strong ego,' the local governments, and the senior citizens.

I would like to add one more actor to the analysis. The *speech synthesizer* tells the user what HyperCLOVA said through a human-like voice. The goal of speech synthesizer is to generate a voice, but it encounters a problem that gendered characteristics of voice are significant, so generating a voice requires gendered dataset and gendered result. There is a case such as Sam, a non-binary voice that uses input from non-binary and transgender communities, but the current status of the field is that most systems use feminine voice, and there are almost no non-binary voices (Danielescu et al. 2023).

This feminine speech synthesizer which represents so-called non-binary HyperCLOVA finally makes all actors' goals achieved. At first, although HyperCLOVA is always mobilized through this system, it is not known to HyperCLOVA. This blind delegation makes HyperCLOVA's 'strong ego' not have the opportunity to question how (in what gender or accent) the words are said. For other actors, as we discussed before, they felt it was natural. This case opens up the space for feminist critique about the whole CareCall system.

This is not just about "HyperCLOVA can be non-binary, but speech synthesizer is binary." The OPP that Naver developers defined to use HyperCLOVA (HyperCLOVA is trained to regularly do

caring conversations with senior citizens) was on the path this feminine speech synthesizer went through. Will this network still work well, if the speech synthesizer escapes the network or changes its gender? This is a question that needs to be empirically evaluated.

## Revisit Woebot: What its change means

The enrollment of Woebot in the network problematized by the clinical psychologists who developed Woebot was easier than CareCall's case, though (or because) the conversation system could not support various open-ended dialogues.

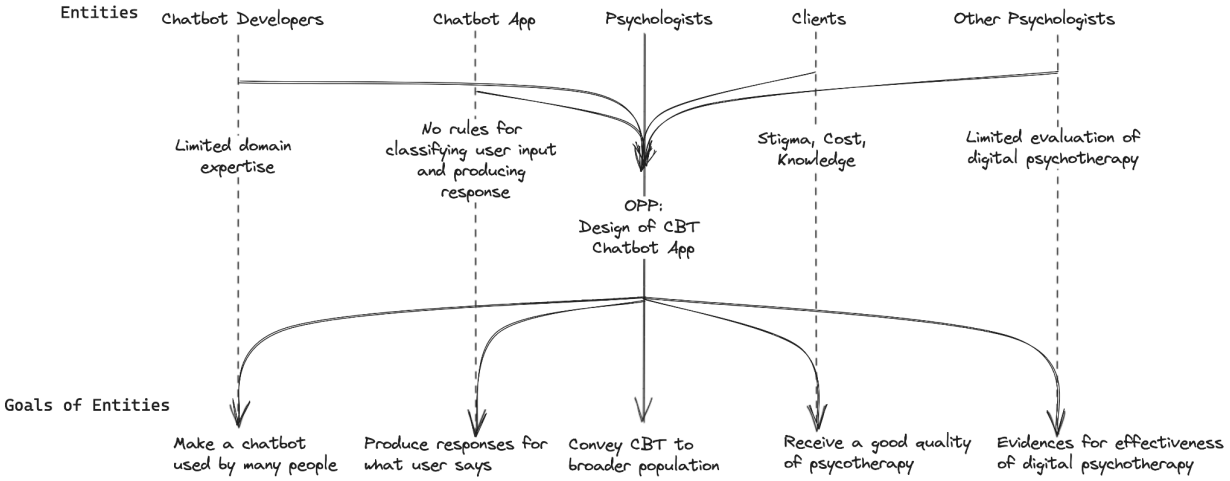


Figure 6. The problematization of the design of Woebot by the psychologists.

In the relationship with users, Woebot needed to be agreed to represent the psychologists. The goal of psychologists to convey CBT practices to the broader population was translated into fixed conversation scenarios. According to the usual rule-based conversational user interface (CUI) design practice (Heo et al. 2023), these kinds of scenarios are generally communicated between the psychologists and the computer developers in the form of spreadsheets and flowcharts.

The spreadsheet contains the category of user inputs and example phrases of each category. This sheet also specifies variable 'slots' that can be filled with 'entities.' For example, when a user says "I feel *happy*," this can be categorized as 'an expression of emotion' with the 'emotion' slot filled with 'happy.'

The flowcharts specify what kind of user's utterance changes the conversation states in what way. For example, a rule can specify that if a user says "an expression of emotion" and the

chatbot has no idea (not filled slot) of the cause of that emotion, the chatbot should ask "What makes you feel E (the emotion that user said)?" An example synthesized from Woebot's advertisement is shown in Figure 7.

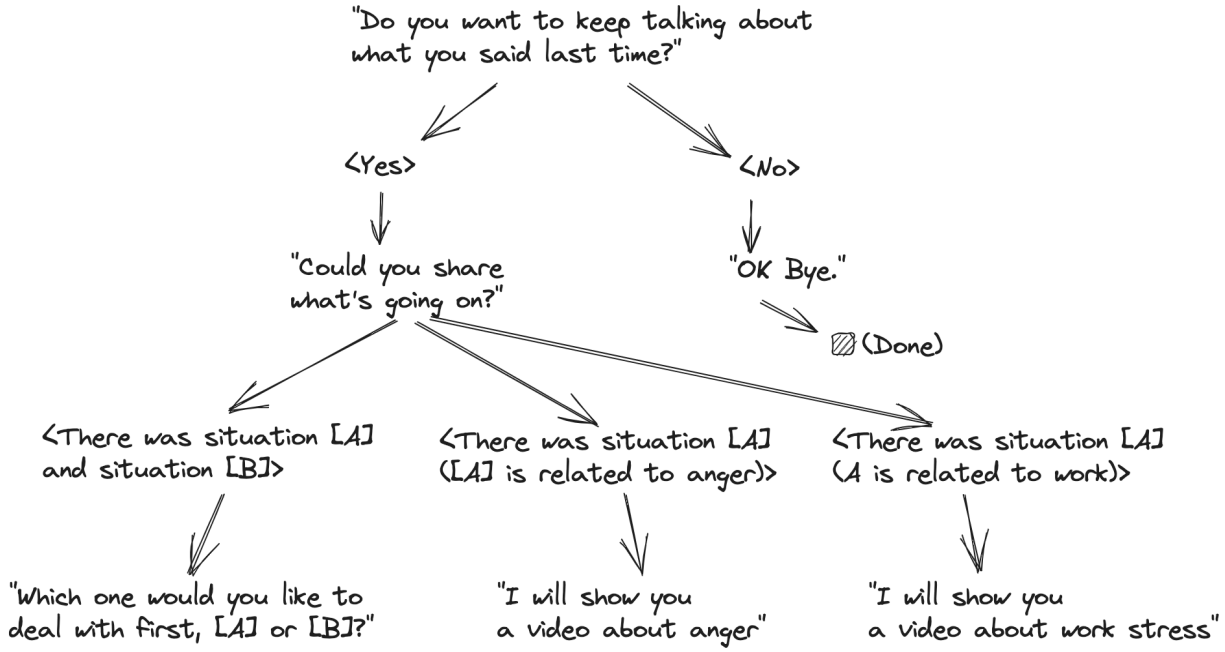


Figure 7. A piece of potential flowchart synthesized by the author based on the Woebot's example conversation.

This deterministic method allowed the Woebot Health psychologists to domesticate Woebot for their purpose without much resistance.

But the analysis of CareCall raises the question of 'What if Woebot becomes more human-like?' This is happening. March 2023, Alison Darcy posted an article "Why Generative AI Is Not Yet Ready for Mental Healthcare" on Woebot Health's home page (Darcy 2023a). July 2023, she posted another article, "Why Science in the Loop is the Next Breakthrough for Mental Health", which announced a registration of a clinical trial of "a reduced set of large language model (LLM)-augmented Woebot." (Darcy 2023b) After this clinical trial was done, she uploaded another post that previewed its result, "In Healthcare, Structure and Science is All You Need." (Darcy 2024)

The narrative of adopting LLM to Woebot is similar to those of Naver developers. They set OPPs that require LLM. Darcy suggests the future of personalized care with an analogy of immunology for cancer treatment which emerged only ten years ago (Darcy 2023a). Sackett, the chief

technology officer at Woebot Health, argued "LLMs truly shone on an emotional level" (Sackett et al. 2024) and thereby can avoid being perceived as robotic. So far, their problem definition with LLM has been a pure replacement of the existing rule-based chatbot app with the LLM-based one. In other words, LLM-based Woebot should do the same performance as the classic Woebot, with only a change of more understanding and empathy.

In this definition of network, they encountered a similar challenge with CareCall. The alliance with LLM required sophisticated interessement devices that would attract it to the classic Woebot's identity and split it from the undesired tokens such as hallucination. They first choosed prompts and examples of Woebot conversation like CareCall's rule specification, but it was not enough. Thus they found a better spokesman that could order LLM more specifically and dynamically at runtime, which was their classic Woebot. They "inject(ed) LLMs into certain parts of our existing rules-based system (Sackett et al. 2024)," which means that the classic Woebot controls the LLM.

With this domesticated LLM-based Woebot, they are now on the track of negotiation with colleague psychologists. Their first clinical trial was focused on measuring user satisfaction in an LLM-based chatbot and getting 'exploratory' findings about efficacy such as symptom changes. Obviously, this is far from the establishment of evidence-based therapy practice. Evaluation of LLM agents for clinical use is an emerging topic, where there is no established consensus (Mehandru et al. 2024). But Woebot Health is on track; Darcy named their strategy *Science in the Loop*, which integrates experimental methods and structured mental health practices, though the details have not been shared (Darcy 2023b).

But will this network survive? In CareCall, the OPP involving LLM attracted a feminine speech synthesizer, though it was not much questioned by other actors. Similarly, the new OPP of LLM-based Woebot including concepts such as empathy and understanding may attract other anthropomorphic actors.

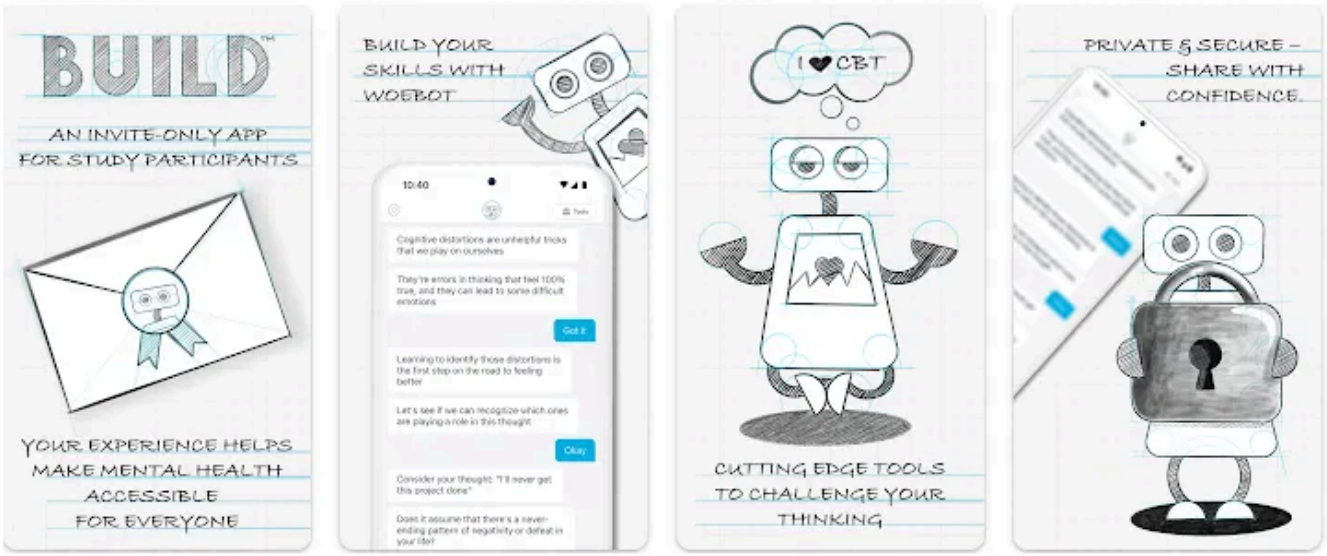


Figure 8. The introduction of a clinical trial for LLM-based Woebot, which was available for the participants in the mobile app store (Retrieved from Gravestein 2023)

Figure 8 shows the current visual image of the LLM-based Woebot, which is the same as the original one. Will this image, representing Woebot's 'cute metal performance,' still pass the new OPP of empathy, personalization, and understanding? Furthermore, will the arrangement of expertise biased to the psychologists be still valid? Again, these are the remaining questions.

## Conclusion

In this article, I suggested a hypothesis: The points LLM passes are similar to the points other anthropomorphic actors pass. Thus, a network involving LLM will open the space for critique about these other anthropomorphic actors as well. As an example, I criticized the stereotype of call center workers regarding the selection of a feminine voice in CareCall. Also, I questioned the validity of the cute metal performance of LLM-Woebot, though I could not answer yet.

## More Resources

AI-powered mental health chatbots developed as a therapy support tool | 60 Minutes (2024.4.)

# References


- Bae, Sanghwan, Donghyun Kwak, Sungdong Kim, Donghoon Ham, Soyoung Kang, Sang-Woo Lee, and Woomyoung Park. "Building a Role Specified Open-Domain Dialogue System Leveraging Large-Scale Language Models." In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2022): 2128-2150.

- Cacioppo, John T., and Laura A. Freberg. "Discovering psychology: The science of mind." (2018): 599.

- Callon, Michel. "Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay." The sociological review 32, no. 1_suppl (1984): 196-233.

- Channel A News. “많이 아파” 했더니 “정말 다행”...못 믿을 AI 안부전화 | 뉴스A (2023) [https://www.youtube.com/watch?v=6NRnlgXujIU](https://www.youtube.com/watch?v=6NRnlgXujIU)

- Colby, Kenneth Mark, James B. Watt, and John P. Gilbert. "A computer method of psychotherapy: Preliminary communication." _The Journal of Nervous and Mental Disease_ 142, no. 2 (1966): 148-152.

- Danielescu, Andreea, Sharone A. Horowit-Hendler, Alexandria Pabst, Kenneth Michael Stewart, Eric M. Gallo, and Matthew Peter Aylett. "Creating inclusive voices for the 21st century: A non-binary text-to-speech for conversational assistants." In _Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems_ (2023): 1-17

- Darcy, Alison. 2023a. “Why Generative AI Is Not yet Ready for Mental Healthcare | Woebot Health.” Woebot Health. March 1, (2023) https://woebothealth.com/why-generative-ai-is-not-yet-ready-for-mental-healthcare/.

- Darcy, Alison. 2023b. “Why Science in the Loop Is the next Breakthrough for Mental Health | Woebot Health.” Woebot Health. July 18, (2023) https://woebothealth.com/why-science-in-the-loop-is-the-next-breakthrough-for-mental-health/.

- Darcy, Alison. 2024. "In Healthcare, Structure and Science is All You Need." March 15, (2024) https://woebothealth.com/in-healthcare-structure-and-science-is-all-you-need/

- Dixon, Steve. "Metal performance humanizing robots, returning to nature, and camping about." _TDR_ (2004): 15-46.

- “Digital Therapeutics Alliance.” Digital Therapeutics Alliance. (2024) https://dtxalliance.org/.

- Gravestein, Jurgen. “Woebot’s Mission to Radically Transform Mental Health.” Substack.com. Teaching computers how to talk (blog). August 30, (2023) https://jurgengravestein.substack.com/p/woebots-mission-to-radically-transform.

- Haupt, Angela. "TIME100 AI: Alison Darcy" _TIME_, September 7, (2023) Accessed June 14, 2024. [https://time.com/collection/time100-ai/6310605/alison-darcy-ai/](https://time.com/collection/time100-ai/6310605/alison-darcy-ai/).

- Heo, Jeongyun, and Uichin Lee. "Form to Flow: Exploring Challenges and Roles of Conversational UX Designers in Real-world, Multi-channel Service Environments." _Proceedings of the ACM on Human-Computer Interaction_ 7, no. CSCW2 (2023): 1-24.

- Hofmann, Stefan G, Anu Asnaani, J Vonk, Alice T Sawyer, and Angela Fang. “The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-Analyses.” _Cognitive Therapy and Research_ 36 (5) (2012): 427–40. https://doi.org/10.1007/s10608-012-9476-1.

- Humeau, Samuel, Kurt Shuster, Marie-Anne Lachaux, and Jason Weston. "Poly-encoders: Transformer architectures and pre-training strategies for fast and accurate multi-sentence scoring." _arXiv preprint arXiv:1905.01969_ (2019).

- Jo, Eunkyung, Daniel A. Epstein, Hyunhoon Jung, and Young-Ho Kim. "Understanding the benefits and challenges of deploying conversational AI leveraging large language models for public health intervention." In _Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems_, (2023): 1-16.

- JTBC News. "건강은 어떠세요, 식사는요?"...손주처럼 안부전화 걸어주는 인공지능, 실제 지켜보니 | D:이슈 (2023) https://www.youtube.com/watch?v=Hs12St_VGxM

- Kim, Boseop, HyoungSeok Kim, Sang-Woo Lee, Gichang Lee, Donghyun Kwak, Jeon Dong Hyeon, Sunghyun Park et al. "What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers." In _Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing_, (2021): 3405-3424.

- Kim, Seong-Ah, Min-Jin Park, and Jung-Ah Kim. "A Study on the Current Status of Social Loneliness and Isolation of Single-person Households in Seoul and Response Strategies." 서울연 2021-OR-19 (2021)

- Maaß, Susanne, and Els Rommes. "Uncovering the invisible: gender-sensitive analysis of call center work and software." In Gender designs IT: Construction and deconstruction of information society technology, pp. 97-108. Wiesbaden: VS Verlag für Sozialwissenschaften (2007)

- Masahiro Mori[森政弘], “The Uncanny Valley[ 不気味の谷],” Energy 7 (1970): 33-35.

- Mehandru, Nikita, Brenda Y Miao, Eduardo Rodriguez Almaraz, Madhumita Sushil, Atul J Butte, and Ahmed Alaa. “Evaluating Large Language Models as Agents in the Clinic.” _Npj Digital Medicine_ 7 (1) (2024) https://doi.org/10.1038/s41746-024-01083-y.

- Naver. "Naver AI Code of Ethics." (2021). Retrieved from “CLOVA Studio JP AI Code of Ethics.” Ncloud-Docs.com. https://guide.ncloud-docs.com/docs/en/clovastudiojp-ethics.

- "Naver Cloud Platform." “\[클로바 시선#5\] AI는 사람을 도울 수 있을까?" (2023) https://blog.naver.com/n_cloudplatform/223139934648.

- “Naver Cloud Platform.” CLOVA Voice. (2024a) https://www.ncloud.com/product/aiService/clovaVoice.

- "Naver Cloud Platform". CLOVA CareCall. (2024b). https://www.ncloud.com/product/aiService/clovaCareCall.

- "Naver CLOVA." “\[대화하는 AI 기술, 연극배우 손숙을 만나다\] 국내 최초 초대규모 AI 적용 - 클로바 케어콜.” _YouTube_. (2022) https://www.youtube.com/watch?v=EdidPwYKjbI&list=PLq8dHmDf5DDVQwJ7Ni0G6W2ctqSg5iN5X.

- Ruiz, Francisco J. "A review of Acceptance and Commitment Therapy (ACT) empirical evidence: Correlational, experimental psychopathology, component and outcome studies." International Journal of Psychology and Psychological Therapy 10, no. 1 (2010): 125-162.

- Sackett, Casey, Devin Harper, and Aaron Pavez. "Do We Dare Use Generative AI for Mental Health?." _IEEE Spectrum_ 61, no. 6 (2024): 42-47.

- Sontag, Susan. _Notes on" Camp"_. (1964).

- Wan, Evelyn. "" I'm like a wise little person": Notes on the Metal Performance of Woebot the Mental Health Chatbot." _Theatre Journal_ 73, no. 3 (2021): E-21.

- Wang, Changwon, Chungkeun Lee, and Hangsik Shin. "Digital therapeutics from bench to bedside." _NPJ digital medicine_ 6, no. 1 (2023): 38.

- Weizenbaum, Joseph. "ELIZA—a computer program for the study of natural language communication between man and machine." Communications of the ACM 9, no. 1 (1966): 36-45.

- West, Mark, Rebecca Kraut, and Han Ei Chew. "I'd blush if I could: closing gender divides in digital skills through education." (2019).

- "Woebot Health." “A Pioneering Collaboration | Woebot Health.” Woebot Health. September 13 (2023) https://woebothealth.com/a-pioneering-collaboration/.

- “Woebot Health.” Woebot Health. June 5 (2024) https://woebothealth.com/.

- "Woebot Labs Inc." “Woebot: Your Self-Care Expert.” App Store. September 21 (2023) https://web.archive.org/web/20230902163555/https://apps.apple.com/us/app/woebot-your-self-care-expert/id1305375832.

- Yao, Lucy, and Rian Kabir. "Person-Centered Therapy (Rogerian Therapy)." (2023).

- Yeung, Jessie, and Yoonjung Seo. 2022. “South Korea’s Middle Aged Men Are Dying ‘Lonely Deaths.’” CNN. CNN. December 19 (2022) https://edition.cnn.com/2022/12/18/asia/south-korea-godoksa-lonely-death-intl-hnk-dst/index.html.

- 서울특별시 사회적 가족도시 구현을 위한 1인 가구 지원 기본 조례 \[Seoul Metropolitan Governance Ordinance for Supporting Single-Person Households to Realize a Social Family City\], Seoul Metropolitan Ordinance No. 6159 (2016)