# ERICA: An Empathetic Android Companion for Covid-19 Quarantine

**Etsuko Ishii[1], Genta Indra Winata[1], Samuel Cahyawijaya[1], Divesh Lala[2],**
**Tatsuya Kawahara[2], Pascale Fung[1]**
[1]Center for Artificial Intelligence Research (CAiRE), HKUST
[2]Graduate School of Informatics, Kyoto University
`{eishii, giwinata, scahyawijaya}@connect.ust.hk`

## Abstract

Over the past year, research in various domains, including Natural Language Processing (NLP), has been accelerated to fight against the COVID-19 pandemic, yet such research has just started on dialogue systems. In this paper, we introduce an end-to-end dialogue system which aims to ease the isolation of people under self-quarantine. We conduct a control simulation experiment to assess the effects of the user interface, a web-based virtual agent called Nora vs. the android ERICA via a video call. The experimental results show that the android offers a more valuable user experience by giving the impression of being more empathetic and engaging in the conversation due to its nonverbal information, such as facial expressions and body gestures. Demo video available at `https://youtu.be/PLPEBXLeKJI`.

## 1 Introduction

To combat the COVID-19 pandemic, lockdowns have been imposed around the world, leading many to experience social isolation. Many people have also undergone weeks of mandatory self-quarantine as they crossed a border or had close contact with a patient. The resulting social loneliness can affect people's mental state, and mental support for those under isolation is suggested (Choi et al., 2020; Zhao et al., 2020). For more than half a century, dialogue systems have played the role of therapist, psychologist or counselor (Vaidyam et al., 2019), and many were designed to help people with a specific concern (Rizzo et al., 2011; DeVault et al., 2014). Hence, dialogue systems have a role in helping curb the effects of social isolation arising from the pandemic.

To meet the emerging needs arising from the pandemic, we extend the idea of Nora, an empathetic dialogue system which mimics a conversation with a psychologist (Winata et al., 2017, 2021),

to specifically mentally support people under self-quarantine, and we install her dialogue system into the autonomous android ERICA (Glas et al., 2016). We utilize ERICA's nonverbal features, which are not offered by Nora, to improve the user interface (UI), because it is well-accepted that the nonverbal behavior of clinicians and therapists affects the outcome of patients (Foley and Gentile, 2010; Beck et al., 2002). During the conversation session, our system asks a set of questions to screen for stress and depression as well as health conditions such as body temperature or shortness of breath. We conduct a comparative study of the virtual agents between the web-based Nora and android ERICA, and we design a dialogue flow particularly for quarantined users based on Nora's graphical UI.

The experimental results show that nonverbal information actually enhances the quality of the user experience during the session by giving the user the impression he or she is being empathized with and listened to. This suggests the importance of the design of nonverbal behavior in dialogue agents, especially for those in the mental health care domain.

## 2 Conversational Agents

Here we describe the end-to-end system for the Nora web-based virtual agent and the android ERICA, whose architectures are depicted in Figure 1.

### 2.1 Dialogue Manager

The dialogue manager consists of three sub-modules: language understanding, response generation, and facial expression prediction. The language understanding module detects the user's intent and slot entities. The response generation module will then generate an appropriate response sentence according to the information from language understanding and empathy analysis. The system utterance generated from the response generation is
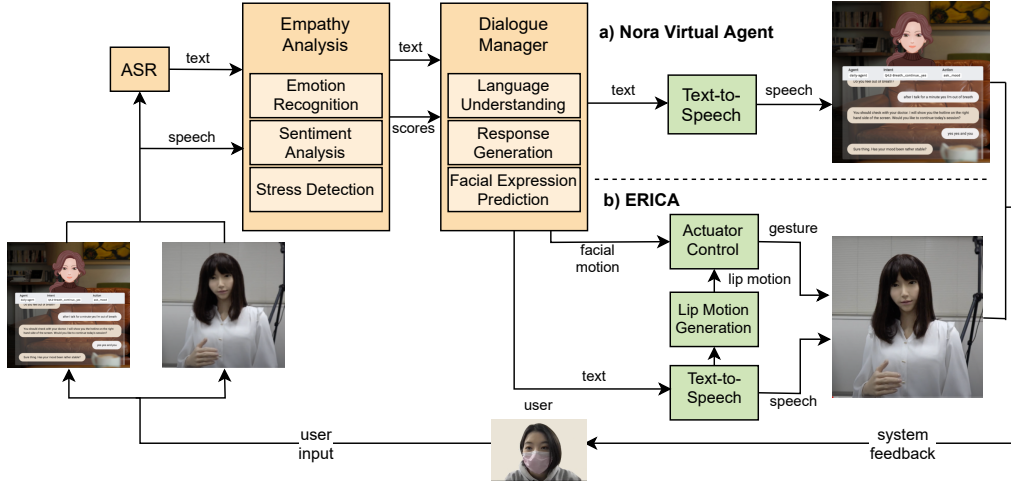
Figure 1: Architecture of a spoken dialogue system for the (a) Nora web-based virtual agent and (b) android ERICA. Note that modules coloured in orange are shared, while the others differ depending on the agent type.
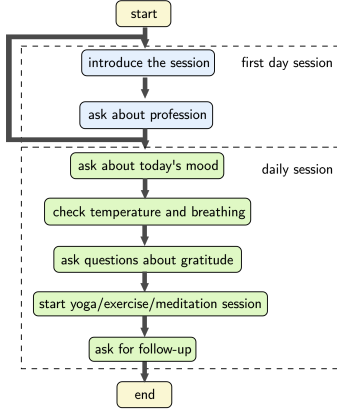


Figure 2: General dialogue flow of Nora.

then passed to the facial expression prediction module, which decides the appropriate facial expression to show. The facial expression is categorized into six distinct classes: happiness, sadness, anger, surprise, laughter, and neutral.

We design a dialogue flow that focuses on a conversation with users in quarantine. As shown in Figure 2, Nora's dialogue conversation is divided into two sessions, the first day session and daily session. In the first day session, the agent will introduce the session and ask about the user's profession. The agent will proceed in the daily sessions by asking about the user's mood and continue with a temperature and shortness of breath check. Afterward, the agent asks questions about gratitude and then recommends that the user enjoy activities such as yoga, exercise, and meditation. At the end of each activity, the agent will ask a follow-up question about how the user feels about the activity. When ending the conversation, the agent will say goodbye and remind the user to wash their hands and wear a mask.

## 2.2 Empathy Analysis

The empathy analysis module contains three modules to understand the user's mood: stress detection, sentiment analysis, and emotion recognition from text and audio (Winata et al., 2017). These modules are later used in the dialogue manager to respond appropriately without discomforting the user. We compute stress, sentiment, and emotion scores on every user turn, and use them to identify whether the user has an extreme psychological condition or not. We also use the scores to track the user's mood every day and provide suggestions to the user for improving their mental well-being.

## 2.3 Nora Virtual Agent

The Nora virtual agent has a web interface, as shown in Figure 1a, that accepts speech input. Users can see their input and responses in text as well as the automatic speech recognition (ASR) results of their utterances. To improve the interaction, the virtual agent provides sound effects to signal the user when the system starts and stops listening. To make the conversation more natural, Nora uses a text-to-speech (TTS) module to generate a speech response.

## 2.4 ERICA

ERICA is a super-realistic female humanoid developed as a conversational agent to play various roles (Glas et al., 2016). She has facial expressions controlled by a facial expression predictor inside the dialogue manager. We develop a mapping of the emotion category to ERICA's actual facial movement and execute it during her utterance, with examples shown in Figures 3(a), (b),

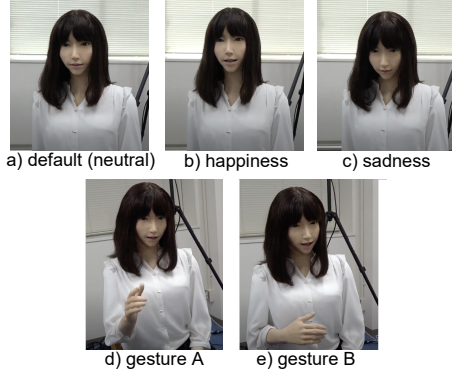a) default (neutral)  b) happiness  c) sadness

d) gesture A  e) gesture B

Figure 3: ERICA's facial expressions and gestures.

and (c). During the user turn, ERICA adopts the default (neutral) face. She also has a lip-motion generation module which is directly controlled by speech signals obtained by the TTS module.

We implement nonverbal behaviors which are triggered based on the turn: body gestures during the system turn, and nodding during the user turn. Body gestures are intended to show openness to users during ERICA's utterance, mainly moving her right hand, as shown in Figure 3(d) and (e). We design four versatile movements and play one of them randomly during ERICA's utterance. During the user turn, ERICA nods to play the role of an active listener until 2.0 s of user silence is detected.

To enhance the naturalness of ERICA's behavior during the conversation, a random gazing model is also introduced. ERICA normally does speaker tracking using Kinect (Inoue et al., 2016, 2020), but since the participant in our case is not on-site, we model gazing behavior as a random uniform sampling of a gaze point nearby the webcam. The gaze point will be randomly changed within a hollow cylinder from the center of the webcam with an outer radius of 0.3 m, inner radius of 0.05 m, and width of 0.2 m. The gaze change decision is taken every 1.5 s.

## 3 Experiments

We conducted a comparative evaluation to see how nonverbal information such as facial expressions and body gestures affect the user experience by asking volunteers to participate in a session with the Nora virtual agent and ERICA.

### 3.1 Experimental Setup

We conducted a simulation of counseling and recruited 19 participants who are fluent in English. In the experiment, a participant accessed the web interface through their web browser and reached the dialogue session page as in Figure 1(a) to have a session with Nora. Then, using a video conference tool, they talk with ERICA just as they would a usual video call.

After finishing the two sessions, we asked participants to evaluate the two systems by choosing which agent is preferred from four different criteria based on their experience during the conversation. Participants were also asked to give an additional comment describing the reason for their choice on each criterion.

### 3.2 Results and Analysis

| Question Item | ERICA | Nora Virtual Agent |
|---|---|---|
| Q1. Overall Experience | 52.6 | 47.4 |
| Q2. Empathy | **68.4** | 31.6 |
| Q3. Attentiveness | **94.7** | 5.3 |
| Q4. User Friendliness | 21.1 | **78.9** |

Table 1: Human evaluation results in terms of the winning rate (%) with participants of $n = 19$. Bold denotes statistically significant (one-sided t-test with $p < 0.1$).

In Table 1, we summarize the experimental results. Overall, ERICA is only slightly preferred (52.6%) over the Nora virtual agent (47.4%) due to its system drawbacks, even though it is perceived to be more attentive and empathetic.

**Q1: Overall Experience** is comparable for several reasons: Although ERICA is regarded as more empathetic and engaging in conversations, users reported that they had a poorer experience, mainly because of the delay in ERICA's response. Moreover, some participants pointed out that the virtual agent is preferable since calling ERICA every day might be troublesome.

**Q2: Empathy** shows that ERICA is perceived as significantly more empathetic thanks to its facial expressions and gestures. Some participants reported that gestures reflected their emotions and thus ERICA was being empathetic, even though her gestures are independent of their emotions.

**Q3: Attentiveness** shows that ERICA is perceived to be significantly more attentive to users because of her nodding, facial expressions, and gestures that mimic human listening behaviors to some extent. Most of the participants agreed that the feedback from ERICA during the user turn, namely, nodding, reduced their anxiety about not being understood.

**Q4: User Friendliness** measures technical or psychological difficulties. The majority of the par-

ticipants reported that ASR accuracy and response time are the drawbacks of ERICA, while some preferred ERICA as she is more human-like and easier to talk to. To enhance the user friendliness, further investigation should be done to handle additional environmental noise in the video call.

## 4 Related Work

One of the major challenges in dialogue systems is how to incorporate empathy, and several papers have explored approaches for end-to-end chatbots (Lin et al., 2020; Ma et al., 2020). Empathetic dialogue systems are attracting more interest in the field of psychiatry as well (Vaidyam et al., 2019), especially those equipped with nonverbal features (DeVault et al., 2014; Rizzo et al., 2011). In addition, Inoue et al. (2016, 2020) utilized ERICA's nonverbal features to make her more empathetic in more generic situations.

## 5 Conclusion

In this paper, we described the implementation of the Nora dialogue system and its application in the android ERICA. A comparison of ERICA against Nora shows that the facial expressions and body gestures of ERICA give a better impression of attentiveness and empathy, even though ERICA has technical drawbacks such as delayed response and worse ASR quality than Nora. These results suggest that nonverbal communication is crucial for machine-to-human conversation as for human-to-human conversation, and special care is needed to design the nonverbal behaviors of empathetic dialogue systems.

## References

Rainer S Beck, Rebecca Daughtridge, and Philip D Sloane. 2002. Physician-patient communication in the primary care office: a systematic review. *The Journal of the American Board of Family Medicine*, 15(1):25–38.

Edmond Pui Hang Choi, Bryant Pui Hung Hui, and Eric Yuk Fai Wan. 2020. Depression and anxiety in hong kong during covid-19. *International Journal of Environmental Research and Public Health*, 17(10):3740.

David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, Gale Lucas, Stacy Marsella, Fabrizio Morbini, Angela Nazarian, Stefan Scherer, Giota Stratou, Apar Suri, David Traum, Rachel Wood, Yuyu Xu, Albert Rizzo,

and Louis-Philippe Morency. 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proc. AAMAS*.

Gretchen N. Foley and Julie P. Gentile. 2010. Nonverbal communication in psychotherapy. *Psychiatry (Edgmont)*, 7(6):38–44.

Dylan F. Glas, Takashi Minato, Carlos T. Ishi, Tatsuya Kawahara, and Hiroshi Ishiguro. 2016. Erica: The erato intelligent conversational android. In *Proc. RO-MAN*, pages 22–29.

Koji Inoue, Divesh Lala, Kenta Yamamoto, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara. 2020. An attentive listening system with android erica: Comparison of autonomous and woz interactions. In *Proc. SIGDIAL*, pages 118–127.

Koji Inoue, Pierrick Milhorat, Divesh Lala, Tianyu Zhao, and Tatsuya Kawahara. 2016. Talking with erica, an autonomous android. In *Proc. SIGDIAL*, pages 212–215.

Zhaojiang Lin, Peng Xu, Genta Indra Winata, Farhad Bin Siddique, Zihan Liu, Jamin Shin, and Pascale Fung. 2020. Caire: An end-to-end empathetic chatbot. In *Proc. AAAI*, volume 34, pages 13622–13623.

Yukun Ma, Khanh Linh Nguyen, Frank Z. Xing, and Erik Cambria. 2020. A survey on empathetic dialogue systems. *Information Fusion*, 64:50 – 70.

Albert Rizzo, Belinda Lange, John G. Buckwalter, Eric Forbell, Julia Kim, Kenji Sagae, Josh Williams, Joann Difede, Barbara O. Rothbaum, Greg Reger, and et al. 2011. Simcoach: an intelligent virtual human system for providing healthcare information and support. *International Journal on Disability and Human Development*, 10(4).

Aditya Nrusimha Vaidyam, Hannah Wisniewski, John David Halamka, Matcheri S. Kashavan, and John Blake Torous. 2019. Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7):456–464.

Genta Indra Winata, Onno Kampman, Yang Yang, Anik Dey, and Pascale Fung. 2017. Nora the empathetic psychologist. In *Proc. INTERSPEECH*, pages 3437–3438.

Genta Indra Winata, Holy Lovenia, Etsuko Ishii, Farhad Bin Siddique, Yongsheng Yang, and Pascale Fung. 2021. Nora: The well-being coach. *arXiv preprint arXiv:2106.00410*.

Sheng Zhi Zhao, Janet Yuen Ha Wong, Yongda Wu, Edmond Pui Hang Choi, Man Ping Wang, and Tai Hing Lam. 2020. Social distancing compliance under covid-19 pandemic and mental health impacts: A population-based study. *International Journal of Environmental Research and Public Health*, 17(18):6692.